14th International Fluid Power Conference (14. IFK)

March 19 – 21, 2024
in Dresden

Sehr geehrte Damen und Herren,

wir freuen uns, Ihnen die Tagungsbände zur 14. Auflage des Internationalen Fluidtechnischen Kolloquiums präsentieren zu können. Als eine der weltweit wichtigsten Tagungen im Bereich der Fluid-Mechatronischen Systemtechnik hat sich das IFK als zentrale Austauschplattform für Experten der Branche in Europa etabliert. Es bietet im internationalen Rahmen Anwendern, Herstellern und Wissenschaftlern die Möglichkeit, Innovationen zu präsentieren und über Entwicklungstrends zu diskutieren.

Der Lehrstuhl für Fluid-Mechatronische Systemtechnik (LFD) der Technischen Universität Dresden organisiert und veranstaltet nun zum 7. mal das IFK, welches sich stetiger Teilnehmer- und Vortragszahlen erfreut. Mitveranstalter ist der Fachverband Fluidtechnik im Verband Deutscher Maschinen- und Anlagenbau e. V. (VDMA). Die Organisation und der Austragungsort wechseln alle zwei Jahre zwischen dem LFD in Dresden und dem Institut für fluidtechnische Antriebe und Systeme (ifas) in Aachen.

Am ersten Tag der Veranstaltung widmet sich das Symposium methoden- und grundlagenfokussierten Beiträgen. Die beiden folgenden Konferenztage bieten einen umfassenden anwendungs- und technologieorientierten Überblick über den neuesten Stand der Fluidtechnik. In dieser Kombination ist das IFK ein einzigartiges Forum zum Austausch zwischen universitärer Grundlagenforschung und industrieller Anwendererfahrung. Eine parallele Fachausstellung bietet die Möglichkeit, sich direkt über Produkte zu informieren und mit Herstellern, Forschern und Anwendern von morgen zu vernetzen.

Der thematische Schwerpunkt der Konferenz liegt auf "Fluid Power - Sustainable Productivity". Die Herausforderungen der neuartigen Systemarchitekturen sind unbedingt zu bewältigen. Die daraus resultierenden Lösungen werden helfen, die technischen und gesellschaftlichen Aufgaben von morgen zu meistern.

Ich wünsche Ihnen viel Spaß beim Lesen der Tagungsbeiträge.



Prof. Dr.-Ing. Jürgen Weber

Dear Sir or Madam,

We are pleased to present the Proceedings for the 14th edition of the International Fluid Power Conference (IFK). The IFK is one of the world's most significant scientific conferences on Fluid-Mechatronic Systems. It offers a common platform for the presentation and discussion of trends and innovations to manufacturers, users and scientists.

The Chair of Fluid-Mechatronic Systems (LFD) at the TU Dresden is organising and hosting the IFK for the seventh time, which has shown high numbers of participants and presentations. Supporting host is the Fluid Power Association of the German Engineering Federation (VDMA). The organization and the conference location alternates every two years between Dresden at the LFD and Aachen at the Institute for Fluid Power Drives and Systems (ifas).

The symposium on the first day is dedicated to presentations focused on methodology and fundamental research. The two following conference days offer a wide variety of application and technology orientated papers about the latest state of the art in fluid power. It is this combination that makes the IFK a unique and excellent forum for the exchange of academic research and industrial application experience. A simultaneously ongoing exhibition offers the possibility to get product information and to have individual talks with manufacturers. The conference is followed by excursions to regional companies and technical sights.

The thematic focus of the conference will be on "Fluid Power – Sustainable Productivity" applications. The challenges of the novel system architectures have necessarily to be overcome. The resulting solutions will help to master the technical and social tasks of tomorrow.

I hope you enjoy reading the conference proceedings.



Prof. Dr.-Ing. Jürgen Weber

Program Committee

Achten, P.	Dr. ir.	INNAS BV, Breda, The Netherlands
Bauer, F.	Dr.-Ing.	HYDAC Technology GmbH, Sulzbach/Saar, Germany
Beck, M.	Dr.-Ing.	Aventics GmbH, Laatzen, Germany
Bock, E.	Prof.	Freudenberg FST GmbH, Weinheim, Germany
Boes C.	Dr.-Ing.	Moog GmbH, Böblingen, Germany
Brand, M.	Dipl.-Ing. (FH)	DMG Mori Seebach, Seebach, Germany
Faß, U.	Dr.-Ing.	Volvo Construction Equipment Germany GmbH, Konz, Germany
Fiedler, M.	Dr.-Ing.	J. Schmalz GmbH, Glatten, Germany
Finzel, R.	Dr.-Ing.	Wacker Neuson Linz GmbH, Hörsching, Austria
Fischer, M.	Dr. Ing.	Argo-Hytos GmbH, Zug, Switzerland
Flaig, A.	Dr.-Ing.	Bosch Rexroth AG, Lohr am Main, Germany
Fuchs, M.	Dipl.-Ing.	Bürkert Werke GmbH, Großröhrsdorf, Germany
Gauchel, W.	Dr.-Ing.	Festo AG & Co. KG, Esslingen, Germany
Geis, C.	Dr.-Ing.	VDMA, Frankfurt, Germany
Huster, G.	Dipl.-Ing.	KraussMaffei Technologies GmbH, München, Germany
Jähne, H.	Dr.-Ing.	HYDRIVE Engineering, Freital, Germany
Klug, D.	Dr.-Ing.	Schuler SMG GmbH & Co. KG, Waghäusel, Germany
Knobloch, M.	Dipl.-Ing. (FH)	HAWE Hydraulik SE, München, Germany
Krieg, M.	Dr.-Ing.	Bosch Rexroth AG, Lohr am Main, Germany
Lasaar, R.	Dr.-Ing.	Liebherr Machines Bulle SA, Bulle, Switzerland

Leonhard, A.	Dr.-Ing.	Parker Hannifin Manufacturing Germany GmbH & Co. KG, Chemnitz, Germany
Lüüs, H.	Dr.-Ing.	Bucher Hydraulics GmbH, Klettgau, Germany
Martens, O.	Dr.-Ing.	Komatsu Mining Germany GmbH, Düsseldorf, Germany
Merkelbach, S.	Dr.-Ing.	Norgren GmbH, Alpen, Germany
Papatheodorou, T.	Dipl.-Ing.	Parker Hannifin Manufacturing Germany GmbH, Bietigheim-Bissingen, Germany
Pfab, H.	Dr.-Ing.	Liebherr-Werk Bischofshofen GmbH, Bischofshofen, Austria
Rahmfeld, R.	Dr.-Ing.	Danfoss Power Solutions GmbH, Hamburg, Germany
Schmitz, J.	Dr.-Ing.	Linde Hydraulics GmbH & Co. KG, Aschaffenburg, Germany
Schmitz, K.	Prof.	IFAS, RWTH Aachen, Aachen, Germany
Schultz, A.	Dr.-Ing.	Magnet-Schultz GmbH & Co. KG, Memmingen, Germany
Weber, J.	Prof.	LFD, TU Dresden, Dresden, Germany

International Advisory Committee

Professor Kalevi Huhtala
Tampere University of Technology, Finland

Professor Bernhard Manhartsgruber
Johannes Kepler University Linz, Austria

Professor Tatjana Minav
Tampere University of Technology, Finland

Professor Victor De Negri
Universidade Federal de Santa Catarina, Brasil

Professor Andrew Plummer
University of Bath, United Kingdom

Professor Kazushi Sanada
Yokohama National University, Japan

Professor Rudolf Scheidl
Johannes Kepler University Linz, Austria

Professor Kim Stelson
University of Minnesota, USA

Professor Katharina Schmitz
RWTH Aachen, Germany

Professor Andrea Vacca
Purdue University, USA

Peer Review beim 14. IFK

Auch in diesem Jahr bieten wir allen Autoren die Möglichkeit eines optionalen Peer Reviews für ihr Paper. Dies stellt sicher, dass die begutachteten Papers hohen wissenschaftlichen Anforderungen genügen und für Förderprojekte oder Promotionsarbeiten die entsprechende Anerkennung erhalten. Diese Papers werden im Tagungsband mit dem Hinweis - **peer reviewed** - gekennzeichnet.

Ein vom Autor zum Review ausgewähltes Paper wird durch insgesamt drei wissenschaftliche Spezialisten einer unabhängigen Begutachtung unterzogen. Diese wird von den Mitgliedern des Programmausschusses sowie zusätzlich vom Lehrstuhl für Fluid-Mechatronische Systemtechnik der TU Dresden oder dem Institut für fluidtechnische Antriebe und Systeme der RWTH Aachen übernommen. Nach dieser ersten Bewertung haben die Autoren die Möglichkeit, ihr Paper falls erforderlich, mit den geforderten Änderungen zu überarbeiten und erneut einzureichen. Nehmen die Gutachter die Änderungen an, wird der Beitrag als - **peer reviewed** - in den Tagungsband aufgenommen.

Dieser aufwändige Begutachtungsprozess dient der inhaltlichen und formellen Qualitätssicherung und wäre ohne die fachliche Unterstützung des Programmausschusses nicht möglich gewesen. Die Organisatoren des IFK bedanken sich bei allen Gutachtern für ihre Unterstützung.

Peer Review at the 14th IFK

This year we once again give all authors the opportunity of an optional peer review of their papers. This guarantees that the examined papers meet the high scientific requirements, and may receive their respective appreciation for funded projects or doctoral theses. Those papers will be marked – **peer reviewed** –.

A paper that has been chosen by the author to be reviewed will be subjected to an independent examination by a total of 3 specialized experts among the Program Committee as well as the Chair of Fluid-Mechatronic Systems (Fluidtronics) at the TU Dresden, or the Institute for Fluid Power Drives and Systems at RWTH Aachen University. After this initial evaluation, the authors have the opportunity to revise their paper as needed and resubmit it with the necessary changes. If the reviewers accept the changes, the paper will be included in the conference transcript as – **peer reviewed** –.

This extensive assessment process serves the purpose of quality assurance in terms of content and form, and would not have been possible without the expert support of the Program Committee. The organizers of the IFK want to thank all reviewers for their support.

Contents

1 Efficiency **15**

1.1 Thermal analysis of the cylinder block of an axial piston pump, the key to monitoring efficiency . . . 17

1.2 Comparison Study of Fully Individualized System Architectures for Electrified Mini-Excavators: Displacement Control (DC) vs Electro-Hydraulic Actuation (EHA) . . . 38

1.3 Run-in behaviour and wear on hydraulic piston seals, evaluation of an endurance test for piston accumulators . . . 51

1.4 Efficiency definitions of hydraulic transformers and first test results of the Floating Cup Transformer (FCT80) . . . 65

2 Fundamentals **79**

2.1 Remaining Useful Life Estimation for Rubber O-Ring under Storage Conditions Considering Dependent Performance Indicators . . . 81

2.2 Development of a Hydraulic Artificial Muscle with High Force Density . . . 94

2.3 Hydraulic pile hammer surrogate model based on physics-informed neural network . . . 104

3 Water Hydraulics **115**

3.1 Holistic Efficiency Measurements of a Mobile Working Machine: Comparison of Conventional Mineral Oils and a Sustainable Water-Based Fluid . . . 117

3.2 Tribological properties of hydraulic cylinder piston sealings in water and oil hydraulics . . . 129

3.3 Numerical Model of Piston/Cylinder Interface with Consideration of Turbulence Effect for Water Hydraulics . . . 144

3.4 Development of digital type tap-water drive flow control valve . . . 158

4 Valves **165**

4.1 Characteristic and oscillation tendency study for different seat geometries of the pilot stage of a two-staged pressure control valve . . . 167

4.2 Dedicated design of the flow angle of free jets for rotary slide valves . . . 183

4.3 Simulation of Gas Leakage on Ball Seat Valves . . . 199

4.4 Development and Tests of a Hydraulic Swivel Drive with Hydrostatic Bearings . . . 211

5 Tribology **219**
5.1 Tribological Design by Molecular Dynamics Simulation - The Influence of Polar Additives on Wall Slip and Bulk Shear 221
5.2 Numerical Study on Abrasive Wear of Reciprocating Seals Under Mixed Lubrication Conditions 232
5.3 Tribological Properties of Different Slipper Designs of an Axial Piston Pump 244
5.4 Fast Computation of Lubricated Contacts: A Physics-Informed Deep Learning Approach 262

6 Materials **275**
6.1 Additive manufacturing of hydraulic components - pressure loss comparison of different self-supporting channel geometries 277
6.2 Bronze cladding on bimetal parts produced by laser deposition brazing . . 290
6.3 On polyoxymethylene composite for sustainable hydraulic valves 298

7 Mobile Applications **309**
7.1 Methodology of System Parameter Optimization for Parallel Electric Hydraulic Hybrid Mobile Machine via Convex Programming 311
7.2 Control of rear-wheel steering for a four-wheel steered agricultural standard tractor 323
7.3 Optimal speed trajectory of electric wheel loaders aiming at extending battery lifetime 335
7.4 Automated System Synthesis for Electrified Mobile Machinery 348

8 Pumps **361**
8.1 Predictive Maintenance for Axial Piston Pumps: A Novel Method for Real-Time Health Monitoring and Remaining Useful Life Estimation 363
8.2 Practical review of reliability methods combined with virtual validation techniques to shift limits of todays hydrostats 375
8.3 A Study on the Effects of Body Deformation on the Performance of External Gear Machines 383
8.4 A novel pulsation compensator for displacement machines 395

9 New and Special Applications-I **407**
9.1 Development of reciprocating air expander for micro-CAES technology . . 409
9.2 Comparative Analysis of Performances of Non-metal Pressurized Reservoirs with Variable Volume 418
9.3 Ship ballasting process time calculation with use of submerged ballast pumps with hydraulic drive supplied from constant pressure hydraulic central loading system on modern product and chemical tankers 430

10 Independent Metering in Mobile Applications **443**
10.1 Compact Fluid Power Control Unit with Independent Metering 445
10.2 Comparison of strategies for unnoticeable mode shifting for independent metering systems in mobile applications 455
10.3 Fault localization for independent metering systems by model-based fault detection 469

11 Sustainable Pneumatics **483**
11.1 Control of a pneumatic system for material strength testing 485
11.2 Product Carbon Footprint of Hydraulic and Pneumatic Components, Challenges in Accounting and Comparability 496
11.3 Exergy analysis for the intermittent air supply in pneumatic machines . . . 505

12 New and Special Applications-II **517**
12.1 Digital redundance for compact subsea electro-hydrostatic actuators using sensor fusion 519
12.2 Development of a generic test rig for the determination of the influence of non-Newtonian fluid properties on the leakage characteristic of rotating displacement pumps 533
12.3 Self-Sensing Micropump detecting air bubbles to improve dosing reliability 544

13 Industrial Control Strategy **551**
13.1 Seamless Integration of Device and Field Data into the System Simulation of a Hydraulic Servo-Press Using AAS and NFC Interface 553
13.2 Development of an open and modular Platform for Hydraulics to increase productivity and flexibility 570
13.3 Physical implentationof a distributed, agent-based control for fluid systems using OPC UA 578

14 Pneumatics **589**
14.1 Sizing of pneumatic drives under energy efficiency aspects 591
14.2 Feasibility Study and Experimental Validation of a Novel Combined Throttling Approach 603
14.3 Reinforcement Learning based PID Controller Design for Mass Flow Controllers 615
14.4 A Trajectory-Specific Approach for Calculating the Holding Force for Surface Grippers 625

15 Control **637**
15.1 Online Learning of Cylinder Velocity Controllers for Excavator Assistance Functions using Local Model Networks 639
15.2 Validation of a hydraulic pulse controller on an off-highway machine 651
15.3 Model Predictive Control of Electro-Hydraulic Systems with multiple degrees of freedom 664
15.4 Data-driven vibration control strategy for hypergravity centrifugal shaking table 674

16 System Design and Architecture-I **685**
16.1 Efficient model-based thermal simulation method demonstrated on a 24-ton wheel loader 687
16.2 A Hydro-Mechanical Vibration Absorber with Adjustable Operating Frequency 699
16.3 Energetic optimization of an existing clamping powerpack by system and control concept analysis and adaption of the hydraulic fluid viscosity 711
16.4 Use of Broadband Silencers in hydraulic circuits to reduce pulsations . . . 717

17 Simulation **729**
17.1 An approach to the evaluation of the energy efficiency of machines based on digital twins and simulation methods . . . 731
17.2 A novel SaaS development platform for fluid power standard drives . . . 743
17.3 Credible simulation: Evaluating the credibility of simulation models and simulation model libraries . . . 754
17.4 Hazard-free steer-by-wire in articulated heavy earth moving machinery using co-simulation model . . . 769

18 System Design and Architecture-II **781**
18.1 Solutions for energy-efficient and easy implementable electrified variable-speed pump drives in mobile applications . . . 783
18.2 Fundamentals of hydraulic transformers . . . 794
18.3 Dynamic valve plate design for an axial piston pump (servo-less pump) . . . 812

19 System Layouts in Mobile Machines **827**
19.1 Electro-hydraulic damping strategies for hydro-pneumatic suspensions . . . 829
19.2 Energy Efficient Excavator Functions based on Electro-hydraulic Variable-speed Drive Network . . . 839
19.3 A comprehensive review of electronic controlled implement architectures for mobile machinery using secondary control . . . 851
19.4 Weight saving potentials of pressure increase in cylinders of mobilemachines kinematics . . . 859

20 Digitalization **869**
20.1 Precise hydrostatic Cylinder Drive with increased Pressure Level for industrial Applications . . . 871
20.2 Novel Engineering and Product Solutions towards Digitalization and Sustainability in Vacuum Handling Automation . . . 883
20.3 Hands-on Approach on developing a Deep Learning Algorithm for State Classification of a Hydraulic Accumulator . . . 895
20.4 Acoustic optimization of a servo-hydraulic pump unit and AI evaluation of the subjective sound perception . . . 907

21 Digital Construction **915**
21.1 Digital assisted collision avoidance for mobile machinery . . . 917
21.2 Object Detection for Autonomous Forest Machinery: A Comparison of State-of-the-Art Instance Segmentation Architectures . . . 937
21.3 Assisted driving Midi-Excavator for augmented performances and improved safety . . . 949

22 Actuators and Sensors-I **961**
22.1 Experimental analysis of energy consumption of piezo actuators used in hydraulic switching valve . . . 963
22.2 Energy harvesting from hydraulic pressure fluctuations using an oscillating piston . . . 975
22.3 Load holding valves with integrated flow sensors . . . 989

23 Hydrogen **1001**
23.1 Hydrogen powered hydraulic Powerpack 1003
23.2 High-Pressure Shut-Off Valve suitable for Hydrogen Applications 1013
23.3 Holistic approach to electro-hydraulic drive solutions for hydrogen piston compressors 1021

24 Actuators and Sensors-II **1035**
24.1 Automation of pneumatic throttle check valves by using novel multi-stable solenoids 1037
24.2 Low energy consumption high flow control system using spool-in-spool design of proportional valve. 1054
24.3 Real-Time Models of Valve Solenoids: An Evaluation of Measurement and Simulation-Based Parameter Identification 1062

Chapter 1

Efficiency

THERMAL ANALYSIS OF THE CYLINDER BLOCK OF AN AXIAL PISTON PUMP – THE KEY TO MONITORING EFFICIENCY

peer reviewed

Roman Ivantysyn*, Jürgen Weber, Alexander Kunze, Wieland Uffrecht

Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden

* Corresponding author: Tel.: +49 351 463-33701; E-mail address: roman.ivantysyn@tu-dresden.de

ABSTRACT

To prepare today's fluid power systems for the future digitalization of the industry, it is necessary to improve the information available regarding the current condition of crucial components of the system. Positive displacement machines, which constitute the core of any hydraulic system, play a vital role in this process. Future smart systems will require more information about the current state of the pump such as power usage and efficiency. Current condition monitoring approaches utilize an array of sensors that need to be sampled at high frequency. The transmission, storage, and post processing of this vast amount of data requires an enormous number of resources, especially if exercised at scale. Previous work conducted at the Institute of Mechatronics Engineering at TU Dresden has demonstrated that measuring the temperature in the lubricating gaps can allow for a deeper insight into the tribological mechanisms in these interfaces. Not only can the gap height, viscous friction and leakage be determined from this information, but also crucial information such as wear level and expected component lifespan can be derived from temperature levels with adequate reference models [1]–[4]

This paper demonstrates that monitoring the thermal condition of the cylinder block is an effective approach to estimate the pump's efficiency. This will be illustrated through both simulation and measurement, in addition to the pioneering measurement of the heat convection coefficient on the cylinder block surface, a critical boundary condition for the simulation.

To measure the temperature of a moving cylinder block, a 160cc axial piston pump was equipped with a telemetric system, which was specially designed and built for this task. Next to the 20 temperature sensors four heat convection coefficient sensors were also carefully placed inside the cylinder block. The resulting measurements did not only validate the simulation but also give a unique insight into the inner workings of a piston pump.

Keywords: Temperature, efficiency, axial piston pump

1. MOTIVATION

The prevailing methodology for condition monitoring relies on a black-box strategy. This method predominantly employs sensors to observe pressure or vibration signals in the frequency domain. Subsequently, this data undergoes processing through artificial intelligence mechanisms to forecast potential breakdowns. While this approach has shown promise, it is contingent on the availability of extensive data sets, which are not universally attainable. The applicability of this data-driven approach is constrained by factors such as the specific manufacturer, operating conditions, user practices, maintenance routines, and the type of hydraulic oil employed. Consequently, this approach remains largely accessible only to the largest corporations capable of amassing substantial datasets, and even for them, it may not be suitable for all pump types and sizes. Moreover, unlike consumer products like automobiles, axial piston pumps are not ubiquitous. This limits the pool of data crucial

for predictive maintenance strategies, further complicating the pursuit of data-intensive predictive maintenance methods.

In light of these limitations, this paper seeks to introduce a novel approach to condition monitoring centered on temperature measurements. Temperature is a readily measurable parameter that can be effortlessly recorded and stored, demanding only minimal bandwidth for data transmission. Within the thermal field of an axial piston pump lies a wealth of information, encapsulating the complex interplay between viscous losses that produce heat and volumetric losses that disperse heat via leakage.

The primary objective of this paper is to demonstrate the existence of a direct correlation between a pump's efficiency, characterized by its losses, and its temperature. To illustrate this relationship, the focus was set on the temperature field of the cylinder block of a 160cc open circuit pump. A cross section of the pump along with the simulated temperature field and the measured correlation between hydromechanics efficiency with its temperature level is shown in **Figure 1**. This paper shows how to simulate this trend accurately and how the measurements were performed. With capabilities to accurately forecast a pump's temperature field, a digital twin can not only deduce the pump's current state but also, given sufficient training data, anticipate looming failures. Beyond merely predicting failures, this method offers invaluable insights into the axial piston machine's current efficiency.

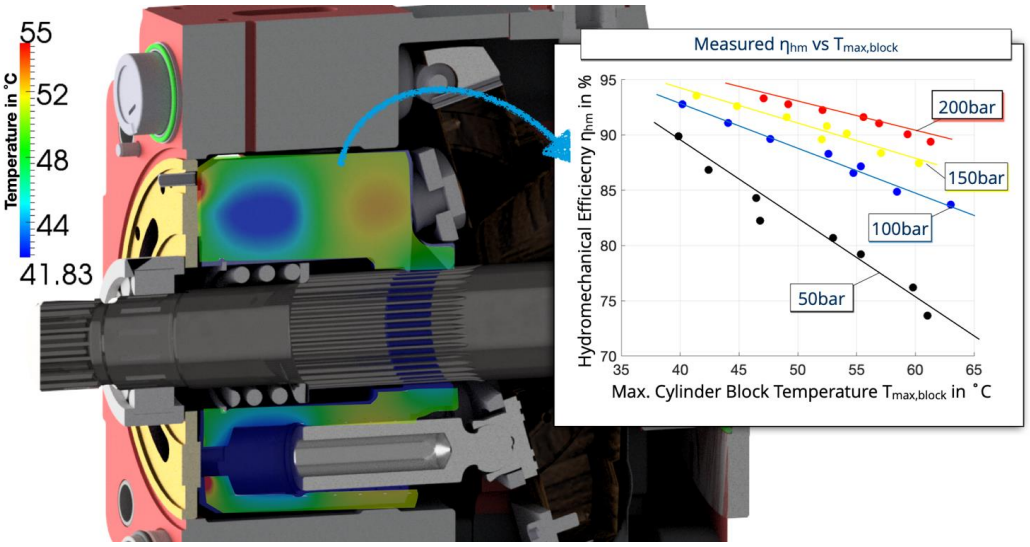


Figure 1: 160cc pump rotating kit, expected temperature field and measured hydromechanic efficiency trend with maximum cylinder block temperature.

2. RESEARCH GOAL AND APPROACH

The primary aim of this research is to devise a methodology to monitor the efficiency of an axial piston pump through the evaluation of the cylinder block's surface temperature. To achieve this goal the following steps are required:

1. Build a comprehensive model of the entire pump in Caspar FSTI to establish the relationship between temperature and efficiency.
2. Validate the model by conducting a pump efficiency test on an unaltered version.

3. Adapt the pump to facilitate thermal evaluations of the rotating cylinder block without altering its efficiency-influencing gap behavior.
4. Highlight prominent thermal facets from the measurements, offering potential industrial benefits.
5. Refine the pump model boundaries based on these measurements.
6. Propose a viable strategy to monitor hydrostatic pump efficiency.

The simulation model was built using Caspar FSTI and serves as a predictive tool, enabling the estimation of losses within all three lubricating gaps of the pump. The ability to simulate the pump's temperature field and efficiency is crucial, given the impracticality of such detailed measurements in industrial settings. This digital twin can serve both as a reference point to ideal conditions and nominal part dimensions and as a comparative tool to gauge efficiency, lifetime and wear-in.

To validate the tool, reference measurements of the pump's efficiency were conducted before any modifications were introduced to the pump. These reference measurements served as a benchmark against which the simulation results were compared. These measurements furnished a comparative framework to ascertain the simulation model's predictive accuracy in terms of efficiency and ensured that any modifications did not adversely impact the pump's performance.

After the validation the pump was modified, including the strategic placement of temperature sensors at locations determined through simulated temperature field analysis. These sensors captured even nuanced temperature gradients across the cylinder block, revealing its thermal dynamics during operation. Coupled FEM analysis, which used the piston load from the gap simulation, were used to estimate the blocks lifetime with various sensor configurations. As part of this task a novel telemetric system was deployed, which was developed by Dr. Uffrecht for high bandwidth measurements in a high speed rotational environment. It was previously used for temperature and pressure measurements in turbine applications [12]–[14].

To enhance the fidelity of the simulation model the measurements will be used to improve the thermal boundary conditions. For the gap simulations it is necessary to estimate the temperature surrounding the pump parts as well as the heat convection coefficients. Conventionally, these heat convection coefficients are estimated using analytical equations for simplified geometries or CFD simulations, as was also done for this pump. Given the inherent variability in literature-based values, there was a compelling need to measure these coefficients directly. A pioneering sensor, previously developed but untested in oil-based applications, was mobilized for this purpose [14]–[17]. This groundbreaking approach not only improves the precision of the gap simulation but also serves as a validation mechanism for the CFD simulations conducted throughout the study.

3. LITERATURE

In the past numerous researchers focused on the thermal analysis as well as the study of the lubricating gaps in hydrostatic machines. The first comprehensive thermal investigations that linked temperature with efficiency have been published by Schlösser and Witt in 1976 [18], [19]. In their studies the stationary thermal behavior of the fluid line temperature was successfully linked with the efficiency of pumps and motors. The concept was confirmed both in theory, using entropy and enthalpy terms, and measurement. The practicality of their findings is lacking the time aspect, e.g. the amount of time it takes for the line temperatures to reach steady state conditions and the distinction of where these losses occur, but nevertheless pave the way for a comprehensive thermal – efficiency model.

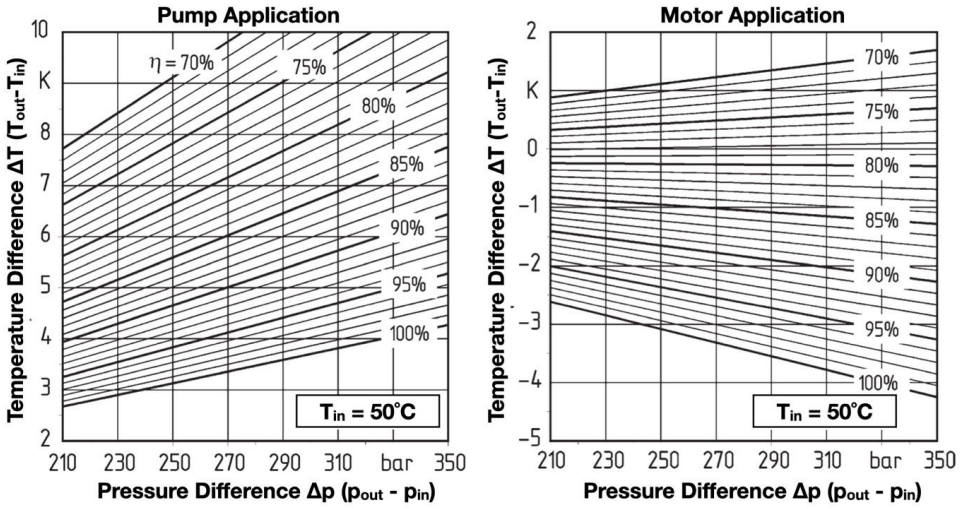


Figure 2: Total efficiency of the hydrostatic machine at steady state line conditions [20]

To understand where these losses occur and under what conditions, a gap analysis is necessary. In the era preceding advanced computational modeling, the fluid gaps of these machines were studied empirically, mostly with highly modified pump components on specialized test rigs. The body of scholars contributing to this domain is vast, hence only a short overview will be given. Test rigs that analyzed the micro movement and pressure distribution gave insight on the necessary gap heights, surface finish and analytical design methods [21]–[28]. As most of these empirical studies were not performed in an actual pump environment, efficiency tests were infeasible. In the past two decades sensors and electronics have allowed for a less invasive measurement of both gap height and temperature in fully functional pumps [9]–[11], [24], [25], [29]–[32]. However, the findings mostly focused on the validation of simulation models, rather than a pump analysis. Most investigations were confined to a singular lubricating gap, missing a holistic analysis of the pump.
In an effort to tackle this problem the author has built a test rig where two out of the three lubricating gaps were monitored both in gap height and temperature [1], [5], [33]. The work gave a unique insight into the thermal conditions in a pump over its entire working regime, demonstrating the unique behavior of each gap, as shown in **Figure 3**.

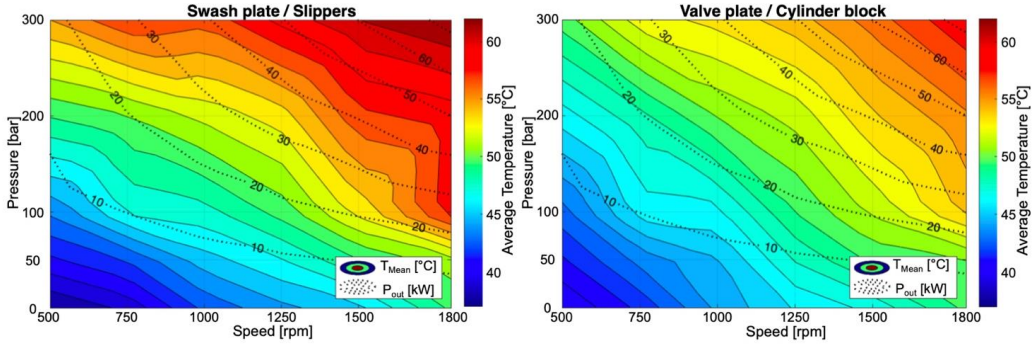


Figure 3: Simultaneous temperature measurements at both valve plate and swash plate [1]

The findings also showed that there is a direct correlation between temperature in the gap and its power loss, which was derived from the measured gap heights (see **Figure 4**). While the trend between

the piston and the cylinder block was captured exclusively in simulation, it mirrored parallel trends, e.g., an apparent pattern across varying pressure levels. This research endeavors to integrate the final missing link: the measurement of the cylinder block's temperature. In doing so, it paves the way for an exhaustive thermo-energetic dissection of the lubricating gaps.

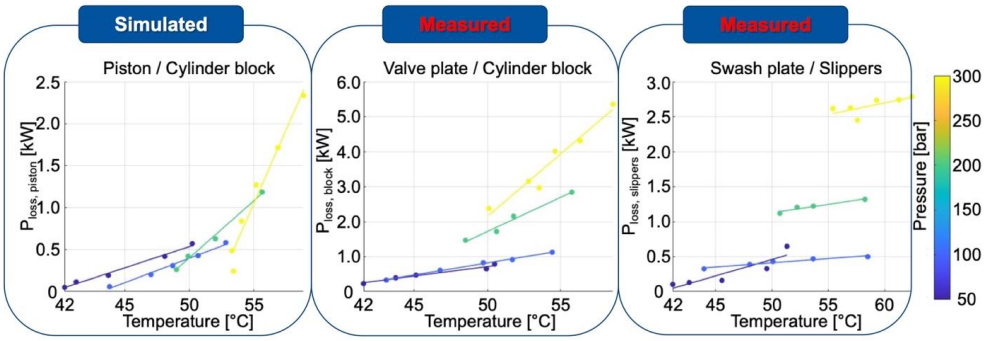


Figure 4: Predicted power loss vs temperature at all three lubricating gaps [1]

4. SIMULATION PREDICTION

Building on the gained knowledge presented in [1], where an entire pump model was built for a 92 cc open circuit pump and compared to measurements both in temperature, gap height and efficiency, a larger 160cc pump from Inline Hydraulics was chosen. This pump was not only used to simulate and validate the piston/cylinder gap simulation model, but also to confirm that the proclaimed temperature-efficiency trends are universally applicable.

4.1. Utilizing Caspar FSTI for advanced axial piston machine simulation

The modelling approach of Caspar FSTI is extensively detailed in [9]–[11], [34]. The credibility of the simulation model is fortified by extensive measurements of gap heights and temperature profiles, independently verified by Schenk [10] and Ivantysyn [5] for the slipper, and temperature measurements for the valve plate by Zecchi [11] and Ivantysyn [2]. The piston/bushing interface was validated using temperature measurements by Pelosi [9], however these measurements were performed for a highly modified singular and stationary bushing, instead of a moving cylinder block. All studies exhibited remarkable agreement between simulation and measurement, instilling high confidence in Caspar FSTI.

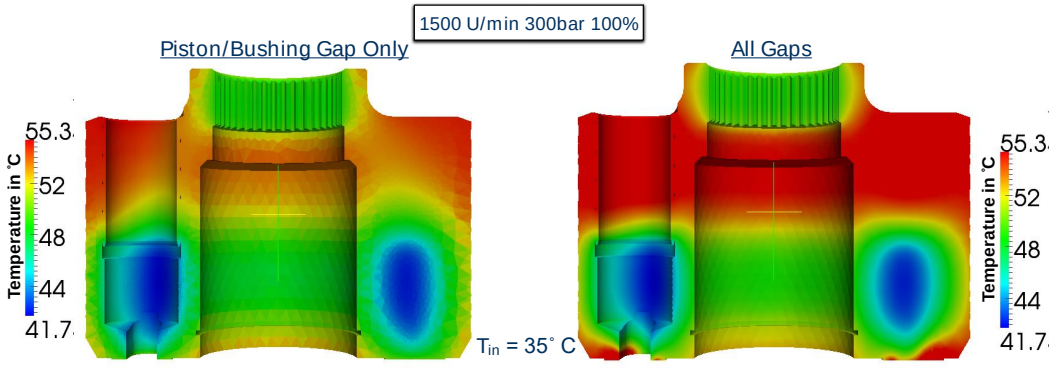


Figure 5: Cylinder block temperature due to piston gap only (left) and considering all 3 gaps (right).

In this study, a 160cc axial piston pump was modelled in simulation for all three main lubricating gaps. Even though the main focus is the cylinder block, a complete pump model is necessary in order to capture all cross interactions and the overall pumps efficiency. The strong influence of the gaps performance on the temperature field of the block is demonstrated in **Figure 5**, where the temperature field is shown with just the piston/bushing gap active (left) and with both other gaps simulated on the right, underlining the necessity to simulate all gaps when trying to accurately predict the block temperature.

Caspar FSTI's capabilities extend to solving the pressure field, temperature field, and micro motion of the entire rotating kit, considering thermal and pressure-induced deformations. This integrated approach yields invaluable data, but more importantly enables the prediction of the power losses in the individual interfaces and the entire pump. The power losses for an example operating point for the 160cc pump are shown in **Figure 6** split up into slipper, valve plate and block losses.

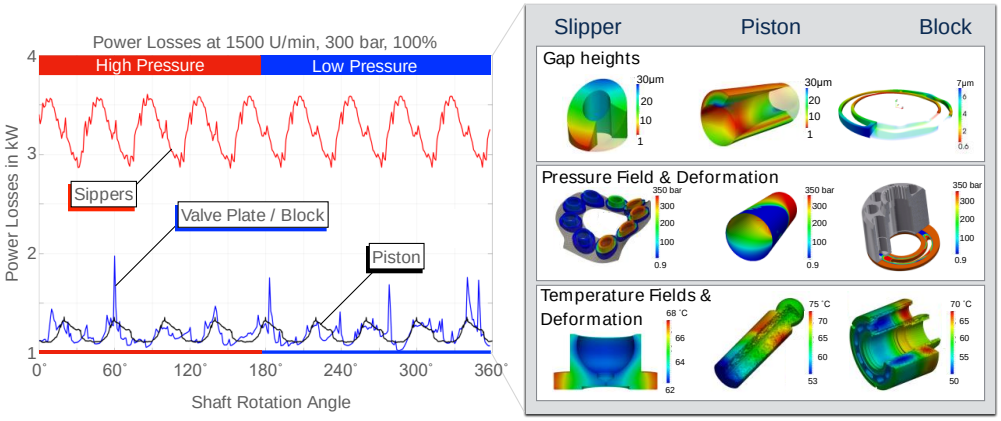


Figure 6: Outputs of the simulation including individual power losses.

For the forthcoming simulation results the following boundary conditions were used: The inlet temperature was kept constant at 35˚C, while the outlet and leakage temperature were set to the measured port temperatures. The integrated Caspar FSTI oil model for HLP46 was used, which is the oil in the utilized test rig. The resulting oil properties can be seen in **Figure 7** below. As depicted the model considers most oil property changes with respect to pressure and temperature. Only exception is the heat capacity of the oil which is kept constant at 2000 J/kg-K and the oil heat conductivity at 130 W/K.

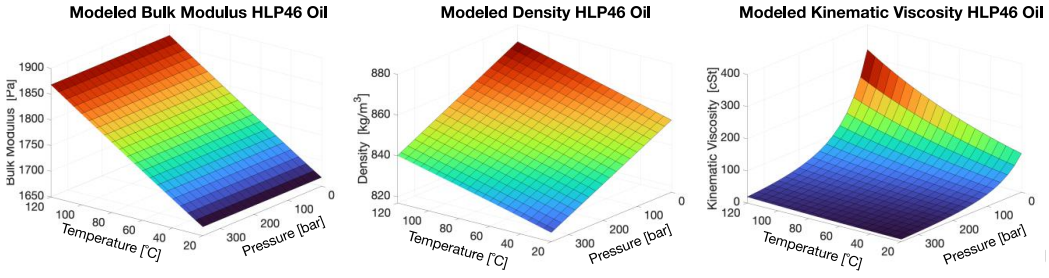


Figure 7: Modeled oil properties of typical HLP46 hydraulic oil.

4.2. Predicting wear to improve thermal forecast

In order to forecast the correct pump efficiency, it is crucial to include the worn-in pump parts, otherwise the simulation will overpredict the losses and temperature levels. Surface measurements are not always available; thus this section will illustrate how the simulation can predict the surface run-in and demonstrate the dramatic change in efficiency an temperature level, due to just a few micrometer of material abrasion. Figure 6 shows that the simulation predicts a much higher power loss at the slipper/swash plate interface for the given operating condition. The inputs were nominal part dimensions of the slipper along with a perfectly flat surface assumption. With these assumptions the simulation predicted contact during the high-pressure stroke not only for this pressure and speed level. **Figure 8** (left) shows the predicted slipper wear potential for various operating conditions. The speed increases from left to right, indicating that at low speed and high pressure as well as at high speed the slipper tends to wear in. In the plot on the top right, the average contact pressure in the fluid gap between slipper and swash plate is shown over one shaft rotation for 2100 rpm and 300 bar. The simulation predicts contact during the high-pressure stroke between 0-180°. On the bottom right, contact pressure, fluid height and the pressure in the gap are shown for φ = 117°, illustrating the contact location.

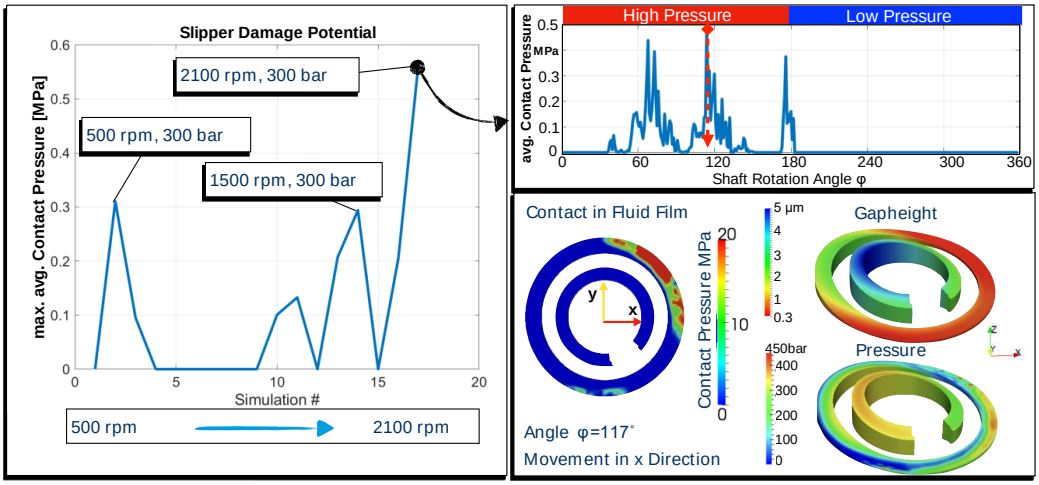


Figure 8: Slipper contact prediction.

The x-direction illustrates the movement direction, while the y-direction is the radial outwards direction, in which the centripetal force acts. As a result of this force slipper is tilted towards the inside edge, as indicated by the 3D gap height in the top right corner. Due to inadequate hydrodynamic pressure build-up, the gap height is very low during this point, causing a contact pressure on the leading edge of the slipper. This slipper has two sealing lands, the outer is the main seal, where the pressure drop from the pocket occurs, while the inner land is for hydrodynamic and structural purposes. Before the wear-in the slipper orientation is not beneficial for the inner edge, causing a low hydrodynamic built up. After the wear the slipper tilt changes, leaning more on the trailing edge.

This is illustrated by the 3D pressure field on the bottom of **Figure 9**. This figure shows the contact pressure for three wear stages on the top half: No wear, wear after 2 iterations, and contact prediction after the final wear profile is reached. The bottom right graph shows the material removal on the main sealing land as predicted by iterative simulation approach compared to a measured profile using a profilometer. The predicted wear matches reality quite well, indicating a 2 µm wear of at the outer edge of the sealing land. The used wear model, which uses the contact pressure and predicts the

material removal in µm, is described in detail in [3].

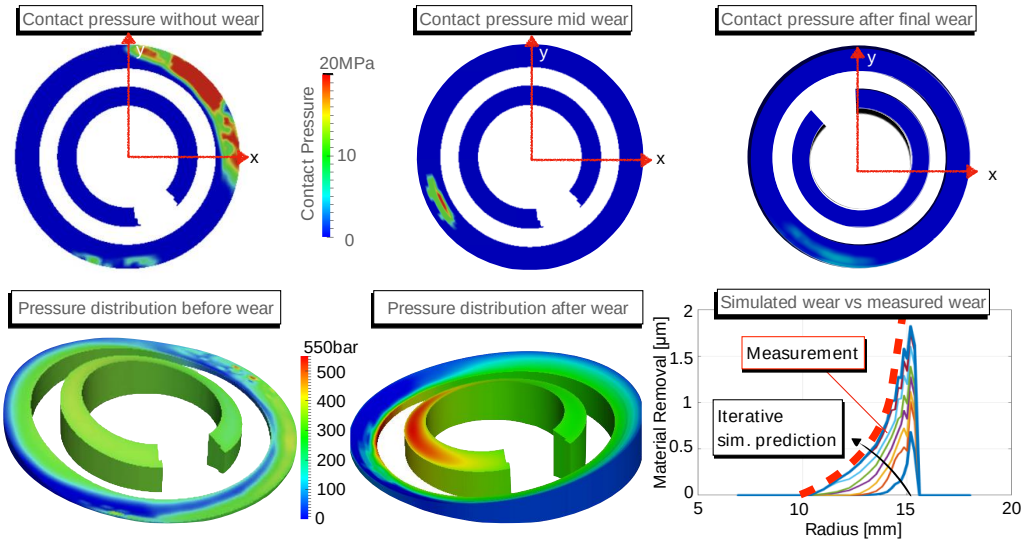


Figure 9: Wear prediction at the slipper using Caspar FSTI at 2100 rpm and 300 bar

On the bottom of Figure 9 is the 3D gap height with the pressure field for the non-worn slipper and the slipper with wear at 1500 rpm and 300 bar. The less then 2 µm rounding of the sealing land edge causes a shift in orientation, allowing the inner land of the slipper to build up a significant amount of pressure. This stabilizes the slipper, increases gap height to more efficient levels and prevents further wear. This change in orientation and gap height produces less viscous friction, reducing the powerlosses at 1500 rpm and 300 bar by 3 kW and the torque loss by more than 50%, as shown in **Figure 10**. The temperature level of the swash plate is greatly affected by the reduced power losses, dropping by more than 100 ˚C, which in turn reduces the leakage and retrospectively the entire cylinder block temperature. The temperature field of the swash plate is shown in **Figure 11**. This example demonstrates the importance to adequately model each gap, as they are all connected either directly or indirectly through the surrounding oil in the case.

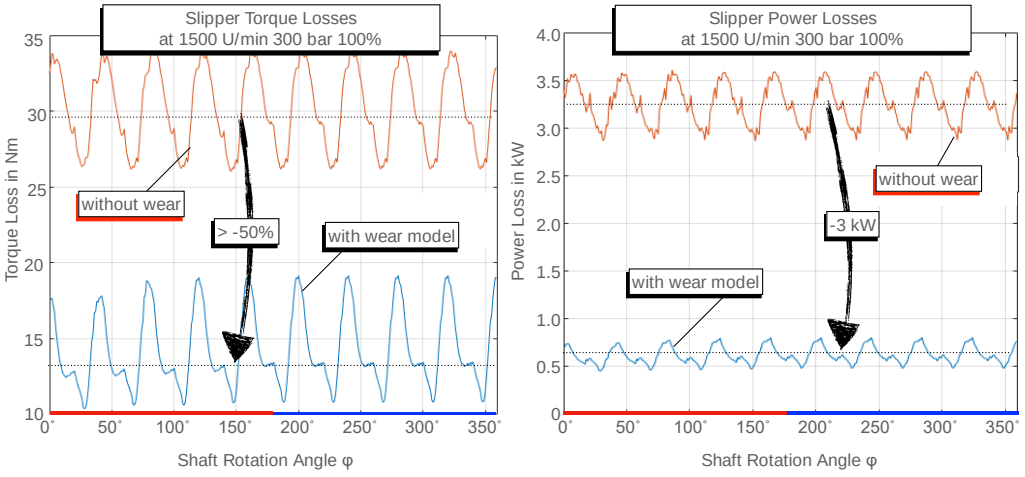


Figure 10: Simulated torque and power losses with and without wear modeled.

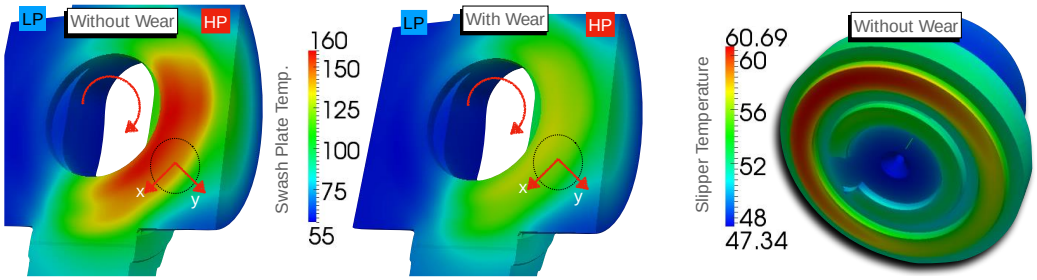


Figure 11: Predicted temperature with and without wear model at 1500 rpm and 300 bar.

Similar to the shown slipper case study, each surface pairing will be modeled, and wear was either predicted or measured and then accounted for. After all wear is accounted for, the entire operating range of the pump is simulated and analyzed. The results are presented in the following chapter, correlating the pump's efficiency with the cylinder block temperature.

4.3. Simulated correlation between efficiency and cylinder block temperature

Block temperature and pump efficiency are closely linked. For instance, elevated temperatures lead to reduced oil viscosities (refer to Figure 7), which could potentially increase leakage. Simultaneously, this reduced viscosity might decrease the oil's load-holding capacity, resulting in smaller gap heights. These lower gap heights, in turn, enhance viscous friction and diminish leakage, thereby raising the temperature. Furthermore, increased temperatures also cause thermal deformation of components, affecting the gap height as well. Despite these complexities, the trends shown subsequently are consistent and corroborated by the measurements detailed in chapter 5.2.

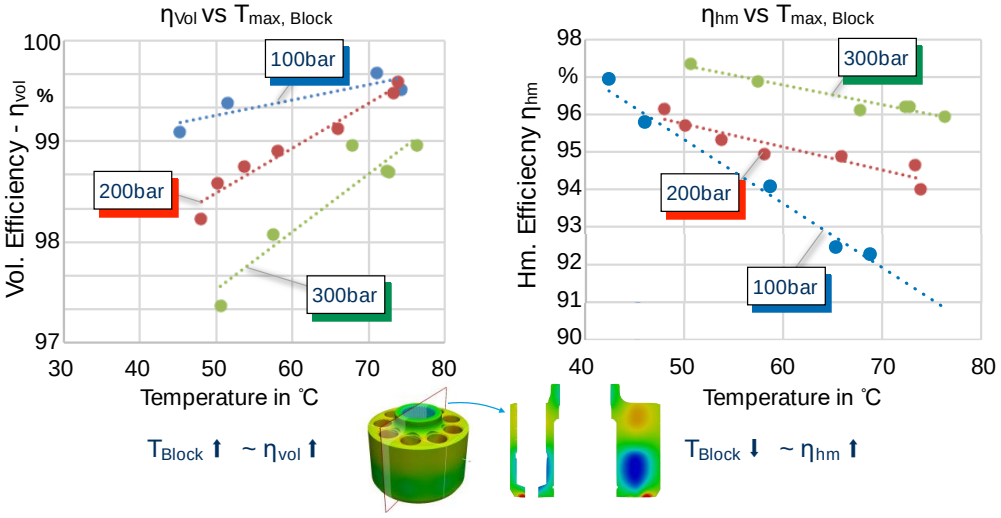


Figure 12: Simulated correlation between efficiency and cylinder block temperature

Figure 12 presents the simulation results for the maximum cylinder block temperature on the x-axis versus the volumetric (left) and hydromechanic efficiency (right) after accounting for all surface run-in. Each circle represents one operating condition, and they are color-coded to a pressure level. Blue dots show the temperature level and the partial pumps efficiency (η_{vol} or η_{hm}) for operating conditions at 100 bar, red 200 bar and green at 300 bar. The dotted lines show the linear fit for each pressure level, which demonstrates a distinct temperature corresponding to specific efficiencies. To facilitate

a clearer visualization, the data points have been plotted on a linear scale. However, it's noteworthy that the volumetric efficiency aligns more accurately with a logarithmic fit than a linear one. In terms of trends, the volumetric efficiency exhibits an upward trajectory, particularly at higher pressures. This pattern can be attributed to the cooling effect of the leakage flow. Essentially, greater leakage translates into more heat being carried away from the lubricating gaps, thus lowering the overall temperature of the part. Consequently, a higher volumetric efficiency, indicative of lower leakage, results in elevated temperatures.

Conversely, the trend observed with hydromechanical efficiency differs from that of volumetric efficiency. Higher hydromechanical efficiency is associated with reduced power losses due to friction, thus leading to lower temperatures.
The simulated pump efficiency compared to the measured efficiency of the pump can be seen in **Figure 13**. The simulated efficiency is based on the worn-in pump parts and the worn surface shapes. Left is the total efficiency at 100 bar and 250 bar plotted across the pumps speed range. The measured values are in the solid line and color coded to the pressure. Overall, the simulated efficiency matches the trend quite well, albeit it is higher than the measured. This is to be expected as not all loss sources are accounted for. When comparing the trend there is a close match between simulation and measurement for both pressure levels. The simulation correctly predicts the drop off at 250 rpm and 250 bar, as well as the trend with speed at the both pressure levels. This low-speed region showed the most significant changes when comparing simulation with and without wear. The differences in the efficiency magnitudes can be explained by the right side of the figure. Here the partial efficiencies at 100bar are shown for both measurement and simulation. As can be seen the volumetric efficiency is matched quite well. The discrepancy results in the hydrodynamic efficiency, which directly correlates to friction and torque losses. The simulation model predicts the losses in the lubricating gaps of the rotating group only, neglecting additional frictional losses in the roller bearings, the setting system and due to the oil churning in the housing. This claim can be supported by the fact, that the difference between the measured and simulated hydromechanical efficiency increases with speed, as the churning losses increase with the exponential of the speed, as shown in published models and measurements [35]–[37]. According to the CFD simulations (see Figure 21) the churning losses for this pump at 2100 rpm are close to 5.5 Nm. This amounts to about 2% additional losses in the hydromechanical efficiency at 100 bar, which is in agreement with the model of Bing [36]. Factoring in these churning losses in the hydromechanical efficiency result in the dotted line shown in the right figure.

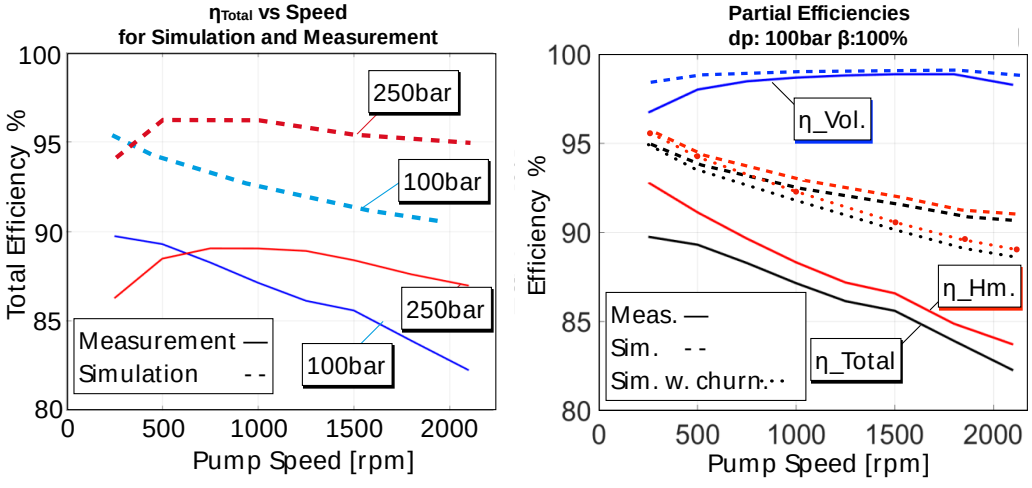


Figure 13: Simulations including wear vs measured efficiencies.

5. MEASUREMENT

To validate the predicted temperature-efficiency trends, it is necessary to measure the cylinder block temperature in a regular pump setting, without modifying crucial pump components, as has been done in the past [9], [29], [38]. To accomplish this, a special telemetric system was used, which can accurately measure a large number of sensors simultaneously. The system was originally developed to measure temperature and pressure in high speed turbine applications within an accuracy of ≦0.1K. The system was adapted to measure 24 thermocouples and 4 heat convection coefficient sensors, which are described in more detail in section 5.3.

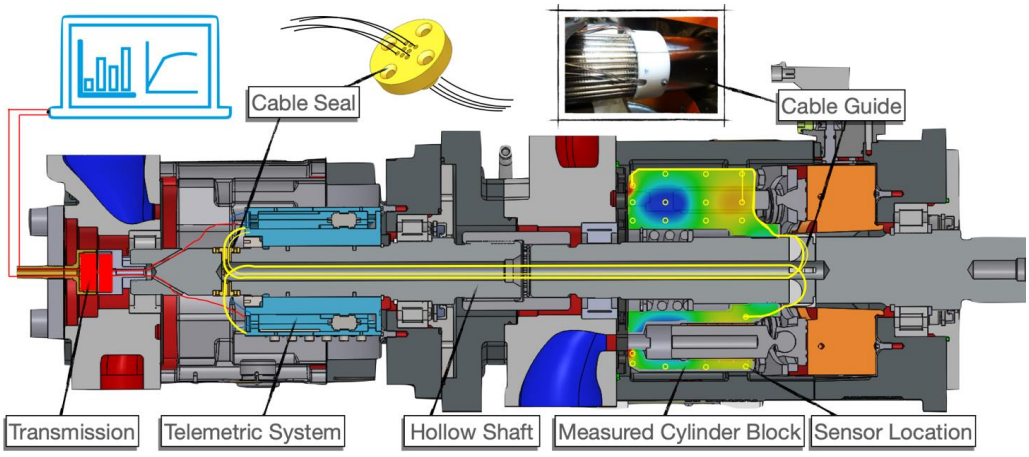


Figure 14: Measurement Set Up

The set up can be seen in **Figure 14**, where a tandem pump arrangement was used to mount and house the telemetric system to the 160cc open circuit pump, while the test pump is shown on the right, with the tested cylinder block highlighted. The sensor locations and temperature field are illustrated, with the exact sensor locations shown in Figure 16. The sensor cables are guided from the cylinder block through hollow shafts using 3D printed parts to guide the cables during pump assembly. In order to seal the case properly the cables were individually guided through a cable seal, that was then sintered shut. The used thermocouples were of type K with metal mantle, which allowed them to be sintered into the cylinder block. The telemetric system converts the analog temperature signals to digital bits that can be transmitted wirelessly to the acquisition system. The power is also transmitted through this brushless system. This set up allows for a very reliable and accurate measurement, not only of temperature but also other signals. In future publications simultaneous pressure and temperature measurements will be shown, that were recorded using this set up.

Next to the cylinder block's temperature, the pump is also monitored holistically with an array of sensors, as shown in **Figure 15**. Each fluid line is equipped both with pressure and temperature sensors, where the temperature is recorded with industrial as well as high accuracy thermistors. The housing of the pump is equipped with 5 temperature sensors and a thermal camera recorded the temperature distribution. As the run-in process and the local temperature are in correlation, the pump was also monitored for particles emitted. The suction line was also equipped with a transparent section that allows for visual inspection of air bubbles, to ensure for a cavitation free operation.

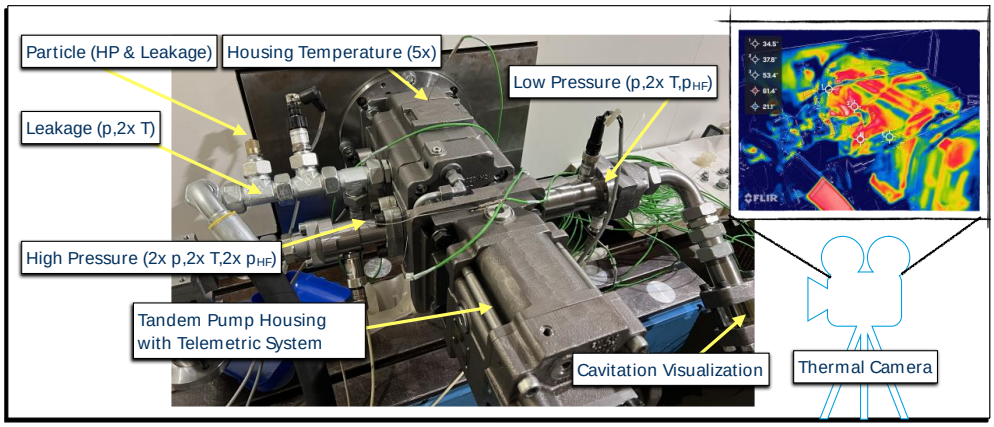
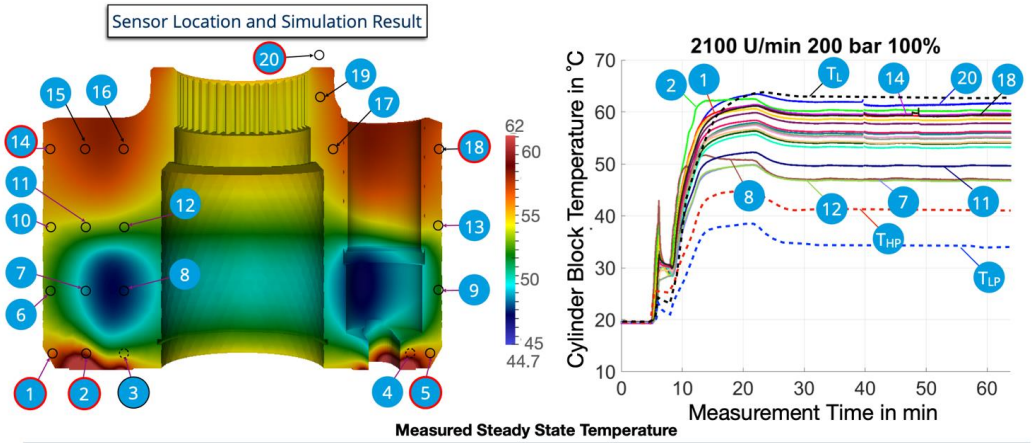


Figure 15: External sensors on the test pump.

5.1. Cylinder block temperature measurements

Figure 16 showcases a sample result for a high-speed operating condition. Here, the simulated outcomes (on the left) can be compared to the measurements (on the right and in the table). For the specific operating parameters of 2100 rpm, 200 bar, and 100% displacement, the pump remained inactive for over 24 hours, essentially reverting to room temperature before initiating the test. Following this, there was a gradual ramp-up, with speed escalating from 0 to 2100 rpm within approximately 10 seconds and pressure surging from 0 to 200 bar within a similar timeframe.



Measured Steady State Temperature

T_{LP}	T_{HP}	T_L	S1	S2	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
35.0	42.1	62.8	60.4	61.3	59.6	55.1	47.9	48.0	54.3	57.0	50.7	47.8	56.3	60.7	58.9	57.3	55.3	60.5	56.1	62.7

Figure 16: Measured and simulated cylinder block temperatures at 2100 rpm, 200 bar and 100% disp.

One notable observation is the swift temperature response in the cylinder block, registering changes in 15 s. However, attaining thermal equilibrium across the entire pump—where all components stabilize, and the desired inlet temperature is met (with a margin of ±1K)—takes close to an hour. It should be noted that this time frame corresponds to a cold start, condition changes when the pump reached operating temperatures usually reach thermal equilibrium within 10-20 min. All measurements were aimed to be taken at 35˚C inlet. Small deviations were later corrected to make

the results comparable between operating conditions.

Looking at the measured temperature distribution (right) and the simulated (left) in Figure 16, a good match between simulation and measurement can be seen. The predicted 15K temperature difference in simulation matches almost the 13.5 K measurement. The sensors were placed 5 mm away from the bottom cylinder block surface, as FEM fatigue calculations predicted a negative trend drilling the holes any closer. This could explain the difference in the maximal temperatures. The highest temperature is registered at sensor 2. These high temperatures at the bottom of the block predict higher viscous losses at the valve plate interface and the other high temperatures are at the top of the cylinder block (sensor 14 & 18), close to the top of the piston / bushing pairing. The piston has a high inclination at these speeds and pressures with its lowest gap height at the top of the bushing, explaining the hot spot. The low spots are at sensor 7 and 8, which are the closest to displacement chamber, which acts in a cooling manner. Sensor 20 was placed outside of the block directly in the leakage oil around the block. It is even hotter than the block, predicting that the hottest region is in fact in the gap and that the block is heated also from the surrounding oil and not just from the gaps.

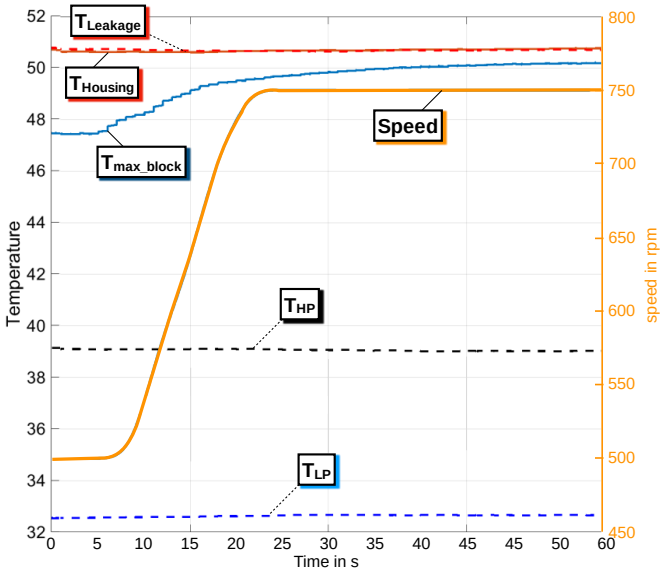


Figure 17: Temperature response rate of the cylinder block

To demonstrate that the cylinder block is better suited for the thermal monitoring as compared to the fluid temperature or the housing, **Figure 17** should be inspected. It shows a speed change from 500 to 750 rpm during warm conditions and the corresponding fluid temperature in the ports, housing temperature and the block temperature. The maximal block temperature (T_{max}) has an instant reaction to the speed change, rising 3˚C in 10s. Contrarily, other fluid and body temperature readings, such as those of the inlet, outlet, housing, and leakage, remain largely consistent in the given time frame despite the speed variation. This emphasizes the significance of the cylinder block temperature and underscores the advantage of in-pump measurements.

Preliminary studies suggest that this immediate temperature spike can potentially be harnessed to predict steady-state temperatures, paving the way for a more industrially relevant monitoring approach. Nevertheless, given that the simulation resides in a steady state, achieving this state in real-time measurements is paramount for meaningful comparisons.

5.2. Predicting the pumps efficiency with temperature measurement

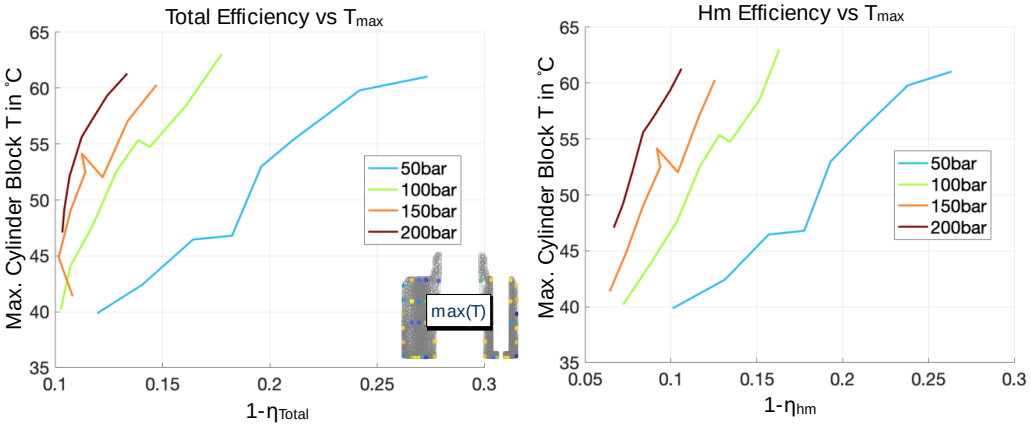


Figure 18: Measured trend between maximum block temperature and total (left) and hydromechanical (right) efficiency at 35˚C inlet.

Figure 18 shows the measured relationship between efficiency and block temperature at 35˚C inlet temperature (as is true for all following graphs). Left in the figure is the maximum cylinder block steady state temperature T_{max} and the total efficiency loss ($1-\eta_{total}$) for each pressure region. Although the pattern isn't strictly linear, distinct zones emerge for each pressure level. At 50 bar, the temperature increase appears to have a linear correlation with efficiency loss, suggesting that it's primarily driven by viscous losses. However, as pressures elevate, volumetric losses start to play a more prominent role. The linear relationship between viscous losses and the temperature can be confirmed by the right graph in the figure, which shows T_{max} vs $1-\eta_{hm}$. There is a clear linear trend for each pressure level. Linear trends are easily modeled and can be very helpful in prediction. Quadratic relationships can be helpful if the entire modeling region is known. The volumetric efficiency and the temperature seem to have such a quadratic relationship, as can be confirmed by the minimal surface temperature and the volumetric efficiency in **Figure 19**.

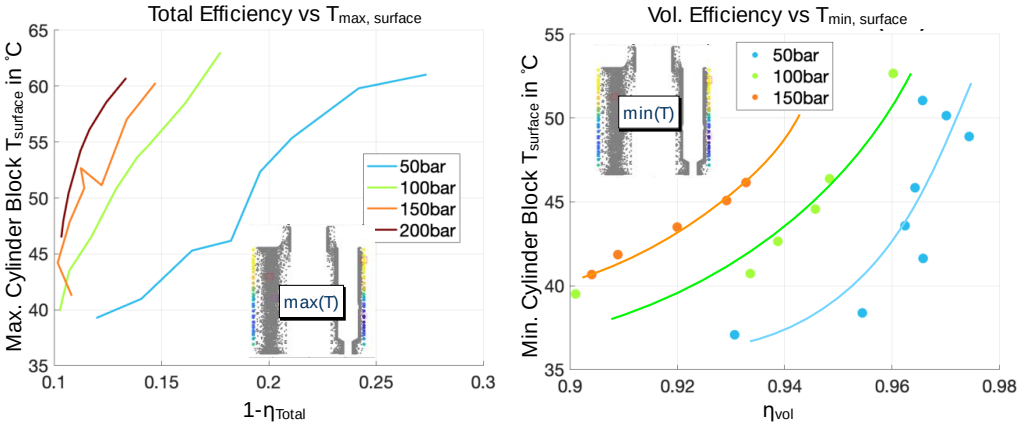


Figure 19: Total (left) and volumetric (right) efficiency trends with max and min block *surface* temperature.

This quadratic relationship of temperature and volumetric efficiency is shown on the right side of the figure for three pressure levels. This trend can be explained by the relationship between leakage flow increases and the gap height, as the flow increases with the height squared. Moreover, the figure

underscores that merely gauging the surface temperature suffices to forecast the pump's total efficiency. This is exemplified in the left graph, where the peak surface temperature is plotted against 1-ηtotal, revealing a similar pattern to what's observed in Figure 18. While the measured efficiencies and temperatures are different in magnitude at the current simulation stage, they still match in trend, confirming the simulated predictions. It is worth mentioning that while the maximum surface temperature can also be correlated with the volumetric efficiency, it was better and clearer trend utilizing the minimum temperature. When correlating the T_{max} with the volumetric efficiency the trend is comparable to the one shown in Figure 20.

Another interesting beneficial correlating the pumps efficiency with the block surface temperature, is the prediction of the pump's speed limit for each pressure level. In **Figure 20** the plot of mean surface temperature against the loss in volumetric efficiency (1-η_{vol}) reveals a noticeable deviation from the quadratic trend at specific temperatures for all pressure levels. Typically, temperature rises with an increase in volumetric efficiency. Yet, beyond a particular temperature threshold, despite a continuing rise in temperature, the volumetric efficiency starts to decline. Detailed analysis highlighted that this phenomenon emerges when the pump operates at exceedingly high speeds, nearing its maximum permissible speed.

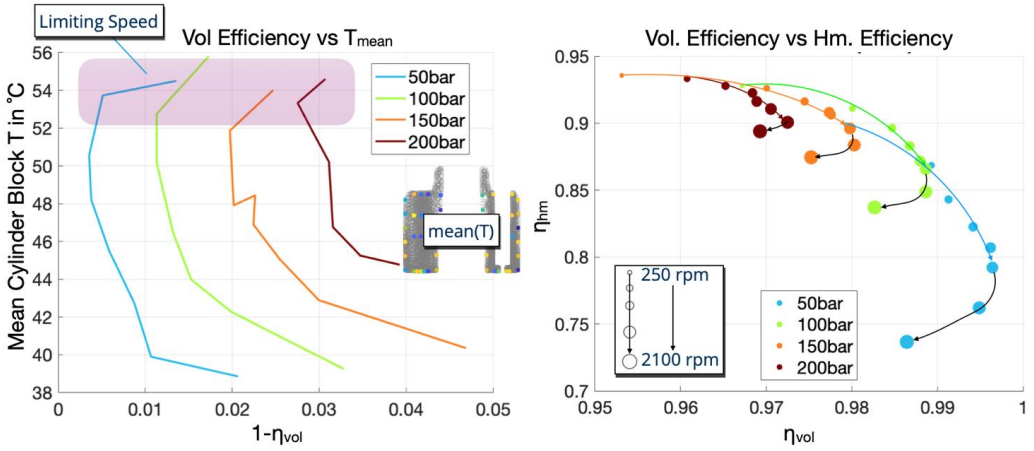


Figure 20: Limiting pump speed as per temperature and efficiency.

This observation is more pronounced when comparing volumetric efficiency to hydromechanical efficiency, as depicted on the right side of Figure 20. The plot shows each measurement point color coded with pressure and the size of the dot related the pump speed from 250-2100 rpm. During normal operation the volumetric efficiency η_{vol} increases with speed while the hydromechanical efficiency η_{hm} decreases. As can be seen this happens in a quite predictable fashion. The trend breaks at high speed when the η_{vol} drops while the η_{hm} drops as well. Notably, the speed at which this occurs varies with the pressure level: it's more immediate at lower pressures compared to higher ones. It will be further investigated if this trend can be detected with the temperature only, but initial trends suggest that it is detectible by comparing inlet, outlet and block temperature rise. This will give the installation of a block temperature a further incentive, being able to predict speed limits in open circuit pumps can be a crucial feature.

5.3. Measuring the heat convection coefficient

In Caspar's thermal modeling, key boundary conditions encompass the oil temperature, operational parameters, and the heat convection coefficient between the component's surface and the oil. While the oil temperature in the outlet and leakage channels can be directly gauged, it can also be estimated

based on pressure and anticipated efficiency, as noted in [39]. Caspar's simulation can ascertain efficiency using an initial approximation, which is then iteratively refined with each subsequent simulation. However, the heat convection coefficient remains a largely unexplored variable in hydraulic pump modeling due to a lack of prior measurements. Historically, computational fluid dynamics (CFD) tools were utilized to estimate this coefficient. Yet, such predictions were grounded in numerous assumptions, and in the absence of direct measurement, a validation was not available.

Uffrecht et. al. [16], [17], [40] were able to develop a miniature sensor, which was able to measure this coefficient. The sensor consists of a thermistor that can simultaneously be used as a heating element, all while being insulated from the surrounding material. The equation of for the heat transfer is given in Eq. 1 below. With a known current the heat loss Q is known, using an insulated material the area A is known as well. To determine the heat convection coefficient h is necessary to measure the temperature ΔT, which is the body surface temperature T and the surrounding Temperature T_{inf}. The sensor accomplishes this by acting as both the heat source and the temperature sensor itself. When powered only a short time the temperature can be measured, and by continuously powering the sensor it will heat up.

$$Q = h \cdot A \cdot (T - T_{inf}) \tag{1}$$

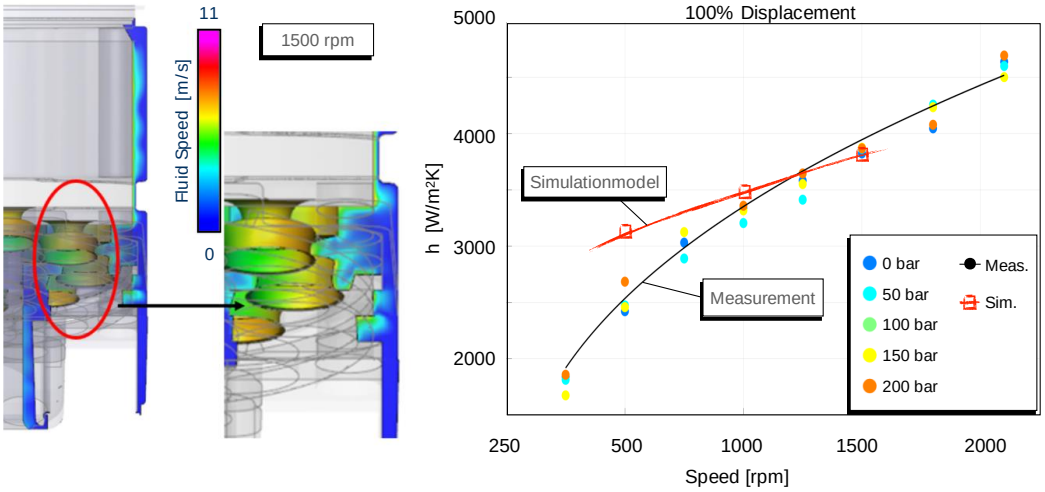


Figure 21: Measured heat convection coefficient of the cylinder block outer surface at 100% displacement vs simulation.

Four of such sensors were placed on the cylinder block wall, two at the outer diameter and two at the neck. The idea was to confirm the different heat convection coefficients at different radii and therefore at different fluid speeds. The neck region is also surrounded by moving pistons, which cause a significant turbulence as predicted by the CFD model shown in **Figure 21** on the left. The CFD simulations were made with Simerics MP+. The model prediction at 100% displacement at three different speeds are shown on the right in red. The measurements at 5 different pressure levels are also shown. It can be seen, that the CFD simulation and the measurement do match in certain regions, however there is a slight difference in slope due to a worse match at low speeds. Nevertheless, these measurements confirm that CFD tools are capable of predicting the heat convection coefficient, as small deviations in the h-coefficient don't have a significant impact on the result. Some CFD models however precited values a full magnitude above these levels, hence it was necessary to validate these numbers.

Figure 21 reveals that pressure variations minimally impact heat convection, with measurements clustering closely for each speed. It also verifies that the sensor delivers reliable results and works well at different thermal loadings of the block.

Figure 22 illustrates how the swash plate angle and block surface influence the convection coefficient. Here the results from 3 of 4 installed sensors are shown, unfortunately did one of the sensors placed on the outer block wall stopped working shortly after commissioning the pump. The 3 remaining sensors are shown in different colors. In green and red are the sensors placed at the neck and in blue the sensor on the outer surface. A dotted line represents 50% swash plate angle, while 100% represents full displacement. The displacement seems indifferent to the heat convection of the outer surface, as evidenced by the overlapping data lines, attesting to the measurement's precision. Both Neck1 (red) and Neck 2 (green) display a noticeable effect from the swash plate angle, with 'h' rising correspondingly. This can be explained by the fact that a lower angle increases the total surface area of the piston exposed to the surrounding oil, increasing drag and turbulence and therefore the heat convection. This effect was first shown in [36], [41] for the churning losses.

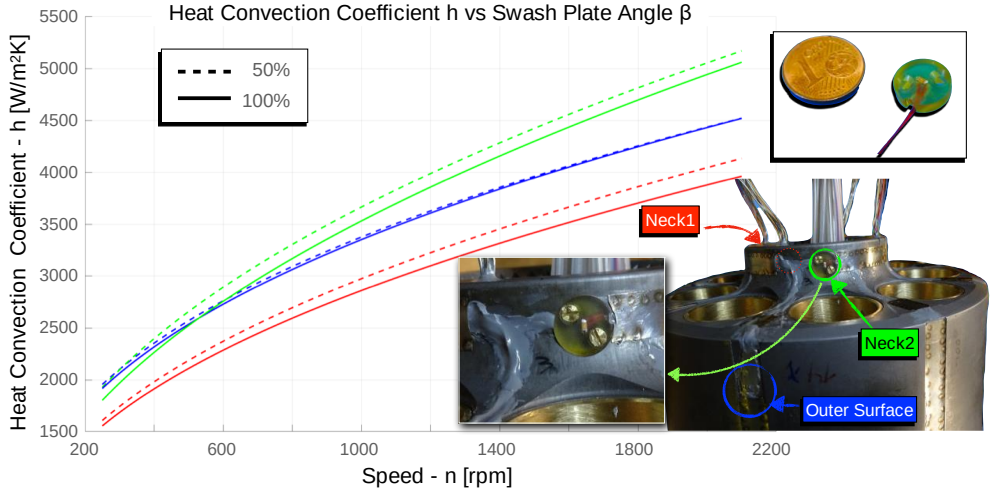


Figure 22: Measured h of the cylinder block at different locations and displacements.

The results of the Neck1 sensor with respect to the outer surface sensors show the expected results. As the rotational radius decreases the fluid speed also decreases, which leads to smaller heat convection coefficients, however the overall trend remains the same just with an offset. The Neck2 sensor diverges in its trend and notably exceeds the magnitudes of the other two sensors. The cause for this deviation is due to a small collection of glue right before the sensor, which was used to guide and protect the cables, see bottom of Figure 22. This caused the oil flow to become turbulent and therefore increase h. Plans are underway to replicate these tests using four new sensors and special care will be taken to avoid surface topological changes with the cable guidance, aiming to validate this hypothesis.

6. CONCLUSION AND OUTLOOK

The intricate relationship between the steady state temperature of a cylinder block in an axial piston pump and its total, volumetric and hydromechanical efficiency forms the core of this investigation, revealing insights that echo and solidify findings previously derived from simulation analyses. Harnessing the capability of the Caspar FSTI—a validate simulation software renowned for its ability

to predict losses in axial piston pumps—the paper revealed a correlation consistent with those previously detected for a 92cc pump from an alternative manufacturer – then in simulation only [1]. This resonant correlation stands testament to the versatility and universality of the observed temperature-efficiency relationship across different pump capacities and potentially, manufacturing intricacies.

A prominent feature of this study was the strategic placement of 24 temperature sensors across 20 distinct positions within a cylinder block, guided strictly by insights drawn from the simulation model. This approach was instrumental in capturing nuanced temperature profiles, thus allowing for an enriched understanding of the internal thermal dynamics. Further enhancing the credibility and accuracy of the empirical process was the incorporation of a specialized telemetric system. This system, known for its precision, enabled real-time transmission of data from both the temperature sensors and the four additional sensors dedicated to measuring heat convection coefficients.

The measurement matches the simulation quite well at many operating conditions, predicting both hot and cold spots, as well as the temperature range within a reasonable accuracy. The empirical findings, when compared with the simulated trends, not only validate the latter's credibility but also shed light on previously uncharted terrains of axial piston pump temperature behavior. For instance, a noteworthy revelation pertains to the pump's behavior upon approaching its limiting speed, providing a fresh perspective into the operational constraints, and associated thermal implications.

In the next steps of this research endeavor the measurements will be repeated with all new parts, which were not previously run-in at the manufacturers end-of line test. This should give additional insight into the thermal effects during the pumps first hours and effect of wear on efficiency and temperature. Furthermore, the cylinder block will be equipped with even more sensors such as pressure and temperature inside the piston chamber.

The ability to predict efficiency using steady-state surface temperatures, combined with the block temperature's rapid response to the pump's operational changes, lends credence to the notion that this method is viable for real-time condition monitoring in axial piston pumps. Moreover, this can be achieved without compromising the block's durability or the overall integrity of the pump. In line with this, a wireless sensor is currently under evaluation and will be tested soon. This sensor can be effortlessly mounted on the surface without necessitating any wired connections. Both power and data transfers are executed wirelessly through a straightforward receiver located inside the pump housing.

Lastly, a focal area for further exploration is the relationship between temperature and efficiency trends. Current findings rely on steady state measurements which are time-consuming. However, preliminary investigations suggest that the steady state temperature may be estimated using an exponential function based on a mere 15 seconds of data. If validated, this approach could significantly enhance the applicability of the research results in real-world scenarios.

NOMENCLATURE

A	Area	m^2
β	Swash plate angle	°
ΔT	Differential temperature	K
φ	Shaft angle	°
h	Heat convection coefficient	W/m^2K
HP	High pressure	
n	Speed of the pump	rpm
η_{vol}	Volumetric pump efficiency	
η_{hm}	Hydromechanical pump efficiency	
η_{total}	Total pump efficiency	
LP	Low pressure	
T	Temperature	°C
T_{in}	Inlet temperature	°C
T_{max}	Maximum temperature of the cylinder block	°C
T_{min}	Minimum temperature of the cylinder block	°C
T_{mean}	Mean temperature of the cylinder block	
p	Pressure	bar
p_{HF}	High frequency pressure transducer	
P_{out}	Output power	kW
Q	Heat Flow	W

REFERENCES

[1] A. Shorbagy, R. Ivantysyn, F. Berthold, und J. Weber, „Holistic analysis of the tribological interfaces of an axial piston pump - Focusing on pump ’ s efficiency“, in *IFK2022*, Aachen, 2022.

[2] R. Ivantysyn, A. Shorbagy, und J. Weber, „An approach to visualize lifetime limiting factors in the cylinder block/valve plate gap in axial piston pumps“, in *ASME/BATH 2017 Symposium on Fluid Power and Motion Control, FPMC 2017*, 2017. doi: 10.1115/FPMC2017-4327.

[3] R. Ivantysyn, A. Shorbagy, und J. Weber, „Investigation of the Wear Behavior of the Slipper in an Axial Piston Pump by Means of Simulation and Measurement“, in *12. IFK 2020*, 2020.

[4] R. Ivantysyn und J. Weber, „Investigation of the Thermal Behavior in the Lubricating Gap of an Axial Piston Pump with Respect to Lifetime“, in *11. IFK 2018*, 2018.

[5] R. Ivantysyn und J. Weber, „“Transparent Pump” – An approach to visualize lifetime limiting factors in axial piston pumps“, in *ASME 2016 9th FPNI Ph.D Symposium on Fluid Power*, Florianapolis, Brazil, 2016.

[6] R. Ivantysyn, A. Shorbagy, und S. Schmatz, „Smart Pump – Decentralised control for vessel engine“, 2018.

[7] S. Horn, R. Ivantysyn, und J. Weber, „Validated efficiency improvements by implementation of structures on the slipper surface of an axial piston pump“, in *IEEE GFPS 2022*, Naples, 2022.

[8] U. Wieczorek, „Simulation Of The Gap Flow In The Sealing And Bearing Gaps Of Axial Piston Machines“, *Proceedings of the 1st FPNI PhD Symposium*, S. 493–507, 2000.

[9] M. Pelosi, „An Investigation on the Fluid-Structure Interaction of Piston/Cylinder Interface“, Purdue University, 2012.

[10] A. Schenk, „Predicting Lubrication Performance Between the Slipper and Swashplate in Axial Piston Hydraulic Machines“, Purdue University, 2014.

[11] M. Zecchi, „A novel fluid structure interaction and thermal model to predict the cylinder block/valve plate interface performance in swash plate type axial piston machines“, Purdue University, West Lafayette, IN, 2013. Zugegriffen: 20. August 2014. [Online]. Verfügbar unter: http://gradworks.umi.com/36/13/3613547.html

[12] G. Schroeder und W. Uffrecht, „A new test rig for time-resolved pressure measurments in rotating cavities with pulsed inflow“, in *ASME Turbo Expo 2010*, 2010, S. 1–10.

[13] W. Uffrecht und E. Kaiser, „Influence of force field direction on pressure sensors calibrated at up to 12,000 g“, *J Eng Gas Turbine Power*, Bd. 130, Nr. 6, S. 1–8, 2008, doi: 10.1115/1.2966390.

[14] B. Heinschke, W. Uffrecht, A. Günther, S. Odenbach, und V. Caspary, „Telemetric measurement of heat transfer coefficients in gaseous flow - First test of a recent sensor concept in a rotating application“, *Proceedings of the ASME Turbo Expo*, Bd. 6, S. 1–10, 2014, doi: 10.1115/GT2014-26239.

[15] W. Uffrecht und A. Günther, „Electro-thermal measurment of heat transfer coefficients“, in *ASME Turbo Expo 2012*, 2012.

[16] W. Uffrecht, A. Günther, und V. Caspary, „Kleine Thermistoren zur messung von wärmeübergangskoeffizienten“, *Technisches Messen*, Bd. 79, Nr. 12, S. 549–558, 2012, doi: 10.1524/teme.2012.0230.

[17] G. Eschmann, A. Kuntze, W. Uffrecht, und S. Odenbach, „Measurement if heat transfer coefficients un gaseous flow - first test of a recent sensor concept for stationary and oscillating flow“, in *ASME Turbo Expo 2014*, 2014, S. 1–11.

[18] W. M. J. Schlösser und K. Witt, „Thermodynamisches Messen in der Ölhydraulik“. VDMA, 1976.

[19] K. Witt, „Thermodynamisches Messen in der Ölhydraulik : Einführung und Übersicht. “, *Oelhydraulik und Pneumatik*, Bd. 20, Nr. 6, S. 416–424, 1976.

[20] H. J. Matthies und K. T. Renius, *Einführung in die Ölhydraulik*. 2014. doi: 10.1007/978-3-658-06715-1.

[21] K. T. Renius, „Experimentelle Untersuchung an Gleitschuhen von Axialkolbenmaschinen“, *O + P : Zeitschrift für Fluidtechnik*, Bd. 17, Nr. 3, S. 75–80, 1973.

[22] K. T. Renius, „Untersuchungen zur Reibung zwischen Kolben und Zylinder bei Schrägscheiben-Axialkolbenmaschinen“, *VDI Forschungsheft*, 1974.

[23] D. S. Wegner und F. Löschner, „Validation of the physical effect implementation in a simulation model for the cylinder block / valve plate contact supported by experimental investigations“, in *IFK2016*, 2016, S. 275–287.

[24] J. Kim und J. Jae-Youn, „Measurement of Fluid Film Thickness on the Valve Plate in Oil Hydraulic Axial Piston Pumps (Part I)“, *KSME International Journal*, Bd. 17, Nr. 2, S. 246–253, 2003, [Online]. Verfügbar unter: http://www.dbpia.co.kr/Journal/ArticleDetail/3227503

[25] J. M. Bergada, J. Watton, und S. Kumar, „Pressure, Flow, Force, and Torque Between the Barrel and Port Plate in an Axial Piston Pump“, *J Dyn Syst Meas Control*, Bd. 130, Nr. 1, S. 011011, 2008, doi: 10.1115/1.2807183.

[26] C. J. Hooke, „The lubrication of overclamped slippers in axial piston pumps—centrally loaded behaviour“, *Proc Inst Mech Eng C J Mech Eng Sci*, Bd. 202, Nr. 4, S. 287–293, 1988, doi: 10.1243/PIME_PROC_1988_202_121_02.

[27] C. J. Hooke, „The effects of non-flatness on the performance of slippers in axial piston pumps“, *Proc Inst Mech Eng C J Mech Eng Sci*, Bd. 197, Nr. 4, S. 239–247, 1983, doi: 10.1243/PIME_PROC_1983_197_104_02.

[28] N. D. Manring, C. L. Wray, und Z. Dong, „Experimental studies on the performance of slipper bearings within axial-piston pumps“, *J Tribol*, Bd. 126, Nr. 3, S. 511–518, 2004, doi: 10.1115/1.1698936.

[29] T. KAZAMA, T. TSURUNO, und H. SASAKI, „Temperature Measurement of Tribological Parts in Swash-Plate Type Axial Piston Pumps“, *Proceedings of the JFPS International Symposium on Fluid Power*, Bd. 2008, Nr. 7–2, S. 341–346, 2008, doi: 10.5739/isfp.2008.341.

[30] P. Achten und S. Eggenkamp, „Barrel tipping in axial piston pumps and motors“, *Proceedings of 15:th Scandinavian International Conference on Fluid Power, 15th Scandinavian International Conference on Fluid Power, Fluid Power in the Digital Age, SICFP'17, June 7-9 2017 - Linköping, Sweden*, Bd. 144, S. 381–391, 2017, doi: 10.3384/ecp17144381.

[31] S. Wegner, „Experimental investigation of the cylinder block movement in an axial piston machine“, in *FPMC2015 -9529*, 2015.

[32] H. Xu *u. a.*, „The direct measurement of the cylinder block dynamic characteristics based on a non-contact method in an axial piston pump“, *Measurement (Lond)*, Bd. 167, Nr. April 2020, 2021, doi: 10.1016/j.measurement.2020.108279.

[33] A. Shorbagy, R. Ivantysyn, und J. Weber, „An experimental approach to simultaneously measure the temperature field and fluid film thickness in the cylinder block/valve plate gap of an axial piston pump“, in *Turbulence, Heat and Mass Transfer 9*, 2018.

[34] U. Wieczorek und M. Ivantysynova, „Computer aided optimization of bearing and sealing gaps in hydrostatic machines—the simulation tool caspar“, *International Journal of Fluid Power*, Bd. 3, Nr. 1, S. 7–20, 2002, doi: 10.1080/14399776.2002.10781124.

[35] Y. Li, B. Xu, J. H. Zhang, und X. Chen, „Experimental study on churning losses reduction for axial piston pumps“, *11th International Fluid Power Conference*, S. Vol. 272-280, 2018.

[36] X. Bing, „Modeling and Analysis of the Churning Losses Characteristics of Swash Plate Axial Piston Pump“, *2015 International Conference on Fluid Power and Mechatronics (FPM)*, S. 22–26, 2015, doi: 10.1109/FPM.2015.7337078.

[37] W. Chunhui, „Analysis of churning Losses of Axial Piston Pump Rotating Parts Based on the Moving Particle Semi Implicit Method“, *Int J Heat Mass Transf*, 2021.

[38] L. Olems, „Investigations of the Temperature Behaviour of the Piston Cylinder Assembly in Axial Piston Pumps“, *International Journal of Fluid Power*, Bd. 1, Nr. 1, S. 27–38, 2000, Zugegriffen: 19. August 2014. [Online]. Verfügbar unter: http://www.tandfonline.com/doi/abs/10.1080/14399776.2000.10781080

[39] J. Schlösser, „Thermondynamisches Messen des Gesamtwirkungsgrades and hydrostatischen Antrieben“, *O + P : Zeitschrift für Fluidtechnik*, S. 285–287, 1973.

[40] W. Uffrecht, B. Heinschke, A. Günther, V. Caspary, und S. Odenbach, „Measurement of heat transfer coefficients at up to 25,500g - A sensor test at a rotating free disk with complex telemetric instrumentation“, *International Journal of Thermal Sciences*, Bd. 96, S. 331–344, 2015, doi: 10.1016/j.ijthermalsci.2015.03.006.

[41] J. Li, Ying, Xu, Bing, Zhang, „Experimental study on churning losses reduction for axial piston pumps“, in *IFK2018*, 2018.

COMPARISON STUDY OF FULLY INDIVIDUALIZED SYSTEM ARCHITECTURES FOR ELECTRIFIED MINI-EXCAVATORS: DISPLACEMENT CONTROL (DC) VS ELECTRO-HYDRAULIC ACTUATION (EHA)

peer reviewed

Timir Patel[1]*, Leonardo dos Santos Franquilino[1], Andrea Vacca[1], Charlie Young[2]

[1]*Maha Fluid Power Research Center, Purdue University, Lafayette, IN 47905, USA*
[2] *Bobcat Company, Bismarck, ND 58504, USA*

* Corresponding author: Tel.: +1 224 499 4497; E-mail address: pate1401@purdue.edu

ABSTRACT

Energy consumption and overall installed power are key parameters for evaluating different technology for the hydraulic actuation system of electrified off-road vehicles. This paper presents a study on these parameters considering the case of all the functions of a 5-ton mini-excavator. The hydraulic system architectures considered for this study are the current commercial solution (load sensing system) and two alternative high-efficiency primary controlled architectures that minimize energy loss: the Displacement Control (DC) system that uses a single electric prime mover and a variable displacement pump for each actuator, and an Electro-Hydraulic Actuator (EHA) system that uses an electric prime mover and a fixed displacement pump for each actuator. For this study, a basic sizing for the two alternative systems is performed based on available commercial components. Both the proposed systems and the conventional LS system are simulated in Simcenter Amesim to determine the energy distribution over a digging duty cycle. The baseline system is also validated based on experimental measurements. The results show a 50% reduction in energy consumption for both the individualized systems, but the overall displacement of the pump units increases by a factor of 3.75. Furthermore, the installed power for the EHA system is 8 times higher than the baseline. Although the study does not suggest a specific architecture for the reference application, the results can assist decision making processes for selecting sub-function primary controlled actuation in future electrified mini-excavators.

Keywords: Electrification, Excavator, Displacement Control, Electro-Hydraulic Actuation, Installed Power

1. INTRODUCTION

The rising demand for battery operated off-road vehicles (ORVs) that can avoid local pollutant emissions is pushing towards the proposal of actuation architectures that can maximize energy efficiency, so that overall energy consumption and battery up-time can be maximized. Advantages of hydraulic actuation, including high-power to weight ratio, resistance to shocks as well as load holding features [1,2] makes fluid power technology particularly suitable for off-road applications such as excavator, independently from the prime mover technology. However, typical metering control concepts used in commercial machines, such as Load Sensing (LS) systems, suffer from high throttle loss and inability to recover energy during overrunning load conditions [3]. Zimmerman et al. [4] showed how in a 5-ton LS mini-excavator operating a digging cycle, only 31% of the pump shaft energy is provided to the actuator. Therefore, there is an opportunity to adopt more energy-efficient architectures, that could be applied over a wide range of ORVs. There are several solutions proposed

by research institutes or companies that could achieve this goal, with different cost/efficiency levels. Among these, significant are the solutions proposed by Pellegri et al. [5], based on the promising concept of digital displacement pump, outlining a potential fuel savings of 43% compared to a baseline 16-ton excavator; or the common-pressure rail strategy proposed by Heybroek et al. [6] on a 30-ton excavator, improving the fuel efficiency in the 34-50% range. A comprehensive list of solutions is presented in the review paper [1]. Certainly, among the most efficient architectures, there are the pure primary controlled ones that avoid throttle loss and allow energy recuperation during overrunning load [3]. As laid out by Weber et al [7], primary control can be implemented using different function-individualization strategies: (i) the use of a common prime mover with separate pumps for each actuator (commonly referred to as Displacement Control – DC); and (ii) the use of separate motor-pump units for each actuator respectively (commonly referred to as Electro-Hydraulic Actuator – EHA). There are some key differences between the two systems, predominantly in the areas of power distribution method, flexibility, compactness, and installed capacity of components. Both architectures have been extensively studied at the authors' center: DC allowed to save 40% fuel in [8] in a 5-ton excavator (further improvement was achieved later through hybridization); EHA permitted to reach efficiencies of up to 80% on the boom-bucket system of a compact loader [9].

There is a common misconception that simply applying electro-hydraulic actuators (EHA) or electro-mechanical actuators (which have comparable efficiency to EHA [10]), is in general the best solution for future sustainable electrified machines [11]. One important aspect of implementing a fully individualized system in an electrified ORV relates to the installed power of the electric components machine. However, most of the past studies like the abovementioned ones put focus only on energy consumption, which cannot serve as the only parameter to determine the optimal architecture for the actuation system. Moreover, these studies focus on the individualization of the most relevant functions without considering the secondary functions such as tracks, blade, etc. This paper aims at studying both aspects of energy efficiency and installed power that come with the implementation of either a DC or an EHA architecture for all the hydraulic functions in an excavator, to highlight the possible limitations of the fully-individualization approach. For this purpose, the paper considers the case of a 5-ton excavator, and determines the sizing of the system considering the extreme case of all the vehicle functions individualized. The energy efficiency of two individualized solutions is evaluated in simulations and compared against the baseline LS system considering a 90º digging cycle where measurements were available. The required pump shaft power is used to estimate the required installed power of the electric drives for each one of the considered architectures. The overall results on efficiency gains vs installed power are finally discussed.

The authors acknowledge that the full application of the DC and EHA architecture to all functions represents an extreme scenario. Their application can be cost effective only on selected functions, such as in Casoli et al [12] where the individualization concept was considered only on the boom, arm and bucket functions of an excavator, or Padovani et al. [13], where the focus was put only at the boom. Another approach for increasing cost-effectiveness is the pump-switching layout, as proposed by Busquets et al. [14], who combined the DC architecture with a pump-switching layout, which allowed controlling all the eight functions of an excavator with only four multi-quadrant pumps. Still, the considerations made in the paper are useful for the decision-making process of sub-functions individualization in an ORV. Moreover, in the context of ORV electrification, there is an argument for using electro-mechanical drives for the rotary actuators. But for the scope of this research, and for providing consistent installed power comparison between baseline and individualized systems, the rotary functions, namely the swing and tracks, are considered as hydraulic actuators.

2. REFERENCE EXCAVATOR SYSTEM

The hydraulic system of a 5 ton, 47 hp excavator (Bobcat E50) is chosen as the reference case. As is typical in many compact excavators, the system layout is based on a multi-actuator load-sensing (LS) post compensated architectures. The hydraulic functions of the vehicle include the standard digging functions: boom, arm, bucket, swing, and the secondary functions: tracks, blade, offset and auxiliary ports, used for backfilling holes and machine positioning. The hydraulic circuit is shown with a simplified schematic in Figure 1. The post compensated design allows to incorporate the so-called 'flow sharing' feature, which reduces the flow to the actuators during events of pump flow saturation to allow simultaneous operation of the working functions [3]. This feature allows minimizing the pump displacement and therefore the prime mover torque (for the reference case an engine).

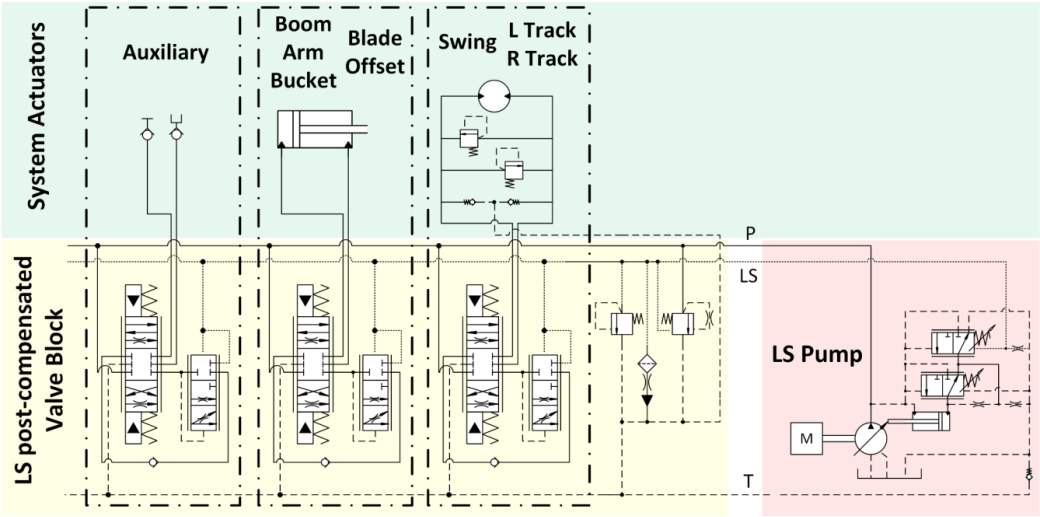


Figure 1: Simplified Hydraulic Circuit of Reference Excavator

3. FULLY INDIVIDUALIZED ARCHITECTURES

The pump-controlled hydraulic systems considered in this study consist of a closed-circuit architecture with the prime mover driving either a variable (DC) or a fixed (EHA) displacement pump. A charge circuit is used to compensates the flow difference resulting from differential areas for linear actuators, pump volumetric losses, and other leakages. Pilot-operated check valves (POCV) are used for connections between the charge and main circuit. Furthermore, load-holding (LH) valves, either mechanical or solenoid operated, can be used to avoid sudden lowering of the actuators [15].

3.1. Displacement-controlled (DC) System

The circuit of the DC architecture for the excavator system is shown in Figure 2. The most notable feature is the use of separate variable-displacement pumps for each of the nine actuators and their connection to a single electric drive system, which is driven at a constant speed. The low-pressure lines of all the actuators are connected to the charge circuit, which includes a charge pump, accumulator, low-pressure relief valve and a cooling system for the fluid. The accumulator in this case is sized large enough to account for the differential flow of the cylinders, which means that the charge pump is sized to only account for the other losses in the system. This leads to lower parasitic losses and a higher overall efficiency of the system. The relieving pressure for the charge circuit is set considering requirements for the swash plate regulation system for the pumps and the pressure

drops across the POCVs.

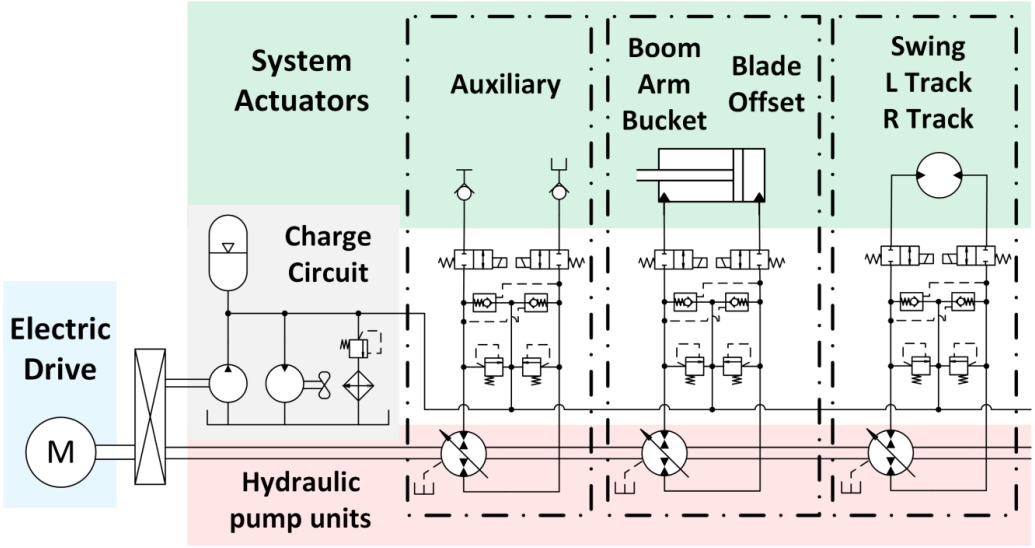


Figure 2: Hydraulic Circuit of Fully Individualized Displacement Controlled (DC) Excavator

3.2. Electro-Hydraulic Actuation (EHA) System

Figure 3 represents the hydraulic circuit of the EHA architecture for the excavator system. Here, the actuators are connected to individual fixed-displacement pumps, each powered by a separate electric drive. The term “EHA” is often used to indicate self-contained and compact electro-hydraulic linear functions [7], although the components of an EHA system can also be spatially distributed for a more convenient implementation that does not require changes in the mechanical structure of the ORV. This architecture is employed for the linear actuators of the system, as shown in Figure 3, where the charge circuit is connected to the pump drain port and an accumulator. The pressure level of the accumulator is therefore, restricted by the allowable drain line pressure for the pump. Furthermore, the self-contained design is able to incorporate a cooling system in the charge circuit [16]. In contrast, the charge circuit for the rotary actuators is equipped with a common charge pump and relieving valve as shown in Figure 3. This design, even with additional parasitic losses and an extra electric drive, is selected considering the flushing requirements of the hydraulic motors. Fixed displacement units, like inexpensive external gear units, can be utilized for the EHA system implementation. Such units commonly have a minimum continuous speed of operation. For this reason, the addition of the by-pass valve (BPV), as presented in [9] is used to handle the low-speed operation of each actuator.

3.3. Comparison between LS and Individualized (DC/EHA) Architectures

The LS system uses the metering concept for controlling the actuator velocity, which leads to significant throttling losses during simultaneous operation. Figure 4 (a) shows an illustrative simultaneous actuation of two functions, the boom and arm, for the LS architecture. The power plot describes the distribution of energy when both actuators are commanded to operate at 50% of their respective maximum rated velocities. From the plot, it can be seen that the pump pressure is always governed by the highest load actuator. The meter-in loss corresponds to throttling of the pump flow in the metering valve, which signifies the margin maintained by the LS pump. The lower load actuator, in this case, the arm, has an additional loss of pressure energy corresponding to the

compensator operation. Lastly, the meter-out throttling loss are a result of the restriction in the return line of the valve to tank. A note here is that the return flow to tank is less than the pump flow due to the differential area of the cylinder. The mentioned losses add up to a significant amount of the total hydraulic energy, which greatly reduces the efficiency of the system. For individualized (DC/EHA) architectures in a similar operation, the power plots are shown in Figure 4 (b). Since these systems have independent drive units, the pump pressures follow the load pressures on their respective actuators and do not involve any load interference. Furthermore, the throttle-less action of the systems leads to minimal loss of energy, with the ability to recover potential energy, a feature not available in LS systems. However, the POCV, LH valves, BPV (with EHA only), and the charge pump result in a small amount of lost energy in the system.

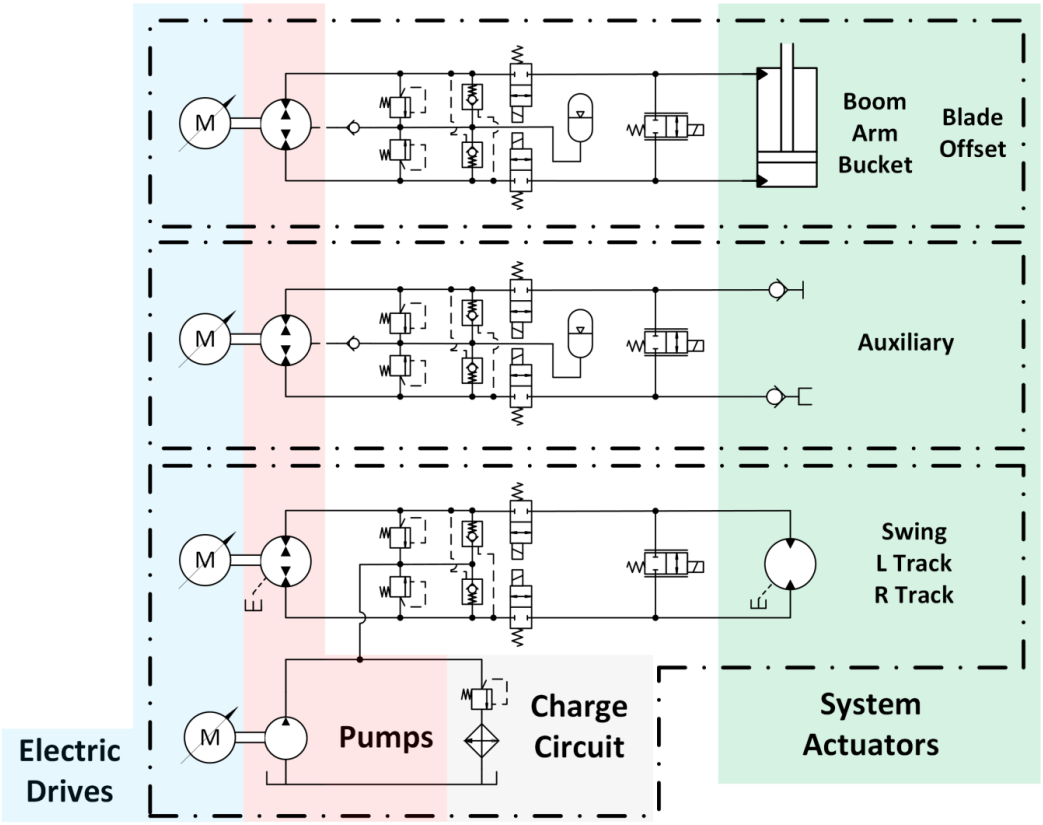


Figure 3: Hydraulic Circuit of Fully Individualized Electro-Hydraulic Actuated (EHA) Excavator

An advantage with the LS architecture is that by studying the operating duty cycles of the machine, the pump can be sized to provide a fraction of the actuators maximum flow during simultaneous actuation. Therefore, to achieve the maximum velocity of an actuator, the simultaneous operation of other actuators can be limited. In other words, flow saturation is imposed on the system to downsize the prime mover and reduce the installed power on the machine. This can be seen in the power plot in Figure 4 (a), where the pump maximum flow is sized close to the required flow to the two actuators at 50% velocity. In contrast, the pump units in the DC/EHA systems have to be sized individually for the required maximum flow of their corresponding actuators. This results in a significant increase in the installed capacity of pumps for such systems as shown in Figure 4 (b).

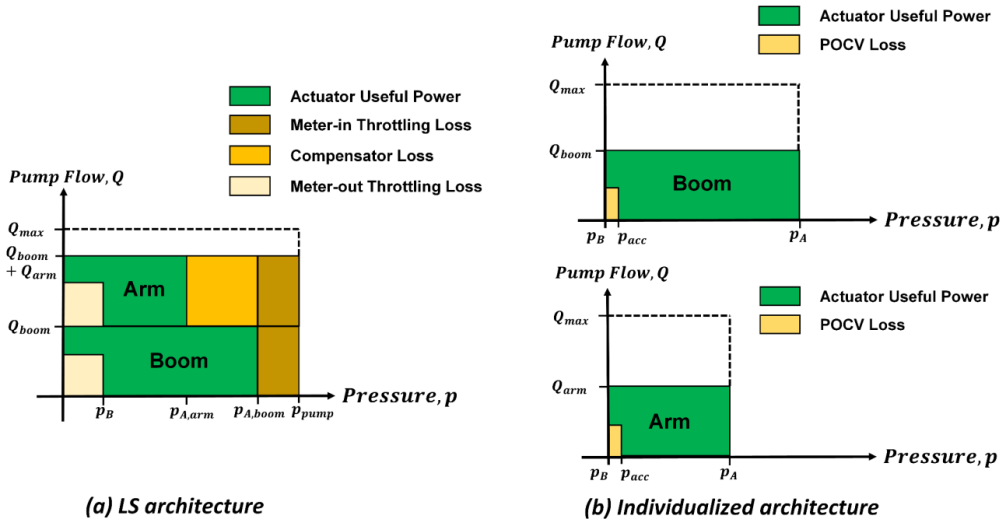


Figure 4: Hydraulic Power Distribution for LS (a) and Individualized DC/EHA (b) architectures, with two actuators operated at 50% of maximum rated velocity (resistive extension mode)

4. SYSTEM SIZING AND SIMULATION MODEL

For the proposed DC and EHA systems, the pumps corresponding to the various functions are selected according to the requirements of the reference machine. Moreover, to allow providing realistic results, each pump is chosen among commericially available units. Swash-plate type axial piston machines designed for closed circuit applications and external gear machines capable of four-quadrant operation are chosen for the DC and EHA systems respectively. From the reference machine, the maximum required speed during extension/retraction (for linear actuators) or maximum angular velocity (for rotary actuators) and the limiting load conditions governed the selection of the pump displacement for each actuator. An important consideration of the pump sizing is that with the current commercial technology, electric motors are capable of driving pump-motor units at much higher speeds than conventional engines, and therefore, the individual pump displacements for the DC/EHA systems can be reduced considerably compared to the baseline LS system. Table 1 lists the pump specifications for each actuator for both the proposed systems. Compared to the baseline LS system, the sum of individual pump displacements for both the DC and EHA systems is about 3.75 times higher. The valves in the systems, namely the POCV, LH, BPV and relief valves, are also sized based on the flow and pressure requirements of the actuators using commercial components. In the DC system, the charge circuit relief pressure is set to 25 bar with a 10-litre bladder-type accumulator and an 8 cc/rev charge pump. The EHA system charge circuit (for the rotary actuators) uses a 5 cc/rev pump with a relief setting of 10 bar. The lower relief setting is selected due to the absence of a swash-plate regulation system in this system. For the linear actuators with the self-contained EHA systems, 5-litre accumulators are used and their pre-charge pressures are set such that the maximum pressure does not exceed 5 bar in the charge circuit, due to the constraint of the allowable drain line pressure for commercial external gear machines.

The sizing for the electric drives in the individualized systems depends on the study of the duty cycles, with simultaneous actuation of functions under specific load conditions affecting the rated power of the installed electric drives (motors). The digging cycle is considered as the most aggressive cycle for the excavator involving the swing, boom, arm and bucket functions, during which the secondary

functions are not operated. Certain considerations can be made for this cycle for systems employing a common prime mover for all actuators, in this case, the LS and DC systems. In the case of the reference excavator, the LS system employs a torque limiter to set a constraint on the maximum torque acting on the variable displacement pump. This method is utilized to limit the installed power of the electric drive for the DC system, while considering the digging duty cycle for the participating functions and the maximum operating limits of the secondary functions. In contrast, for the EHA system, the separate electric drives for the actuators need to be sized for their corresponding operating load limits. The torque limiter is useful in setting the constraint for the maximum torque on electric drives, but the limiting load conditions for each actuator can be used to select the actual rated torque and therefore, power for the corresponding electric drives.

Table 1: Individual Pump Unit Specifications for DC and EHA excavators

	Displacement-Control (DC)			Electro-Hydraulic Actuation (EHA)			
Actuator	Pump Displace-ment [cc/rev]	Maximum continuous speed [rpm]	Maximum continuous pressure [bar]	Pump Displace-ment [cc/rev]	Maximum continuous speed [rpm]	Minimum continuous speed [rpm]	Maximum continuous pressure [bar]
Boom	25	3400	300	26.7	3000	350	300
Arm	30	3400	300	29.06	3000	350	300
Bucket	25	3400	300	21.99	3000	350	280
Swing	18	4000	300	14.53	3500	350	290
Offset	30	3400	300	29.06	3000	350	260
Blade	35	3400	300	34.56	3000	350	300
Auxiliary	25	3400	300	26.7	3000	350	260
Right Track	25	3400	300	26.7	3000	350	300
Left Track	25	3400	300	26.7	3000	350	300

Lumped parameter models for the baseline LS architecture, the DC and EHA architecture were implemented in Simcenter Amesim for the analysis of the energy consumption. The selected approach allows simulating the performance of each architecture considering both steady-state and dynamic component characteristics. To this end, following elements were taken into consideration:

- Volumetric and hydromechanical efficiency maps for both fixed displacement pumps (gear pumps) and variable displacement pumps (swash plate pumps)
- Realistic line lengths/diameter, using values derived from the baseline reference vehicle. Reasonable line lengths are necessary to properly consider capacitance and resistance effects.
- Steady-state characteristic curves for the valves (i.e. flow vs pressure drop at different openings) from off-the shelve components
- Dynamic behaviors in terms of valve opening/closing, stroke/destroke time for the pumps.
- Unitary efficiency (ideal) electric drives and mechanical transmission in the DC and EHA systems. This simplification is also made to make the results agnostic with respect to the technology of the electric system.

As model inputs, the actuator force/torque from baseline measurements is applied to the hydraulic actuators instead of a mechanical model of the excavator bodies. Finally, an “operator model” is designed to ensure tracking the actuator displacement measurements for each one of the three system models. The modelling of the three architectures only involved the digging functions of the excavator, boom, arm, bucket and swing, since the 90º digging duty cycle has been selected for the energy

analysis in this research. To assess the validity of the simulation approach, the LS system model with the four functions was previously validated using measurements from the reference machine for the digging cycle. The energy losses in the system are computed using the approach described by Zimmerman et al. [4].

5. RESULTS

5.1. Excavator Working Cycle

A typical 90º digging cycle, which involves the swing, boom, arm, bucket functions, for the reference excavator is selected for simulation and energy analysis of the three systems. The cycle involves the filling of dirt in the bucket, rotation of about 90º, dumping the load onto a pile, and returning to the starting position. A representative snapshot of the cycle for 21 seconds of operation is shown in Figure 5, which shows the normalized position (left) and the experimentally measured hydraulic force of the actuators (right) in the reference LS system.

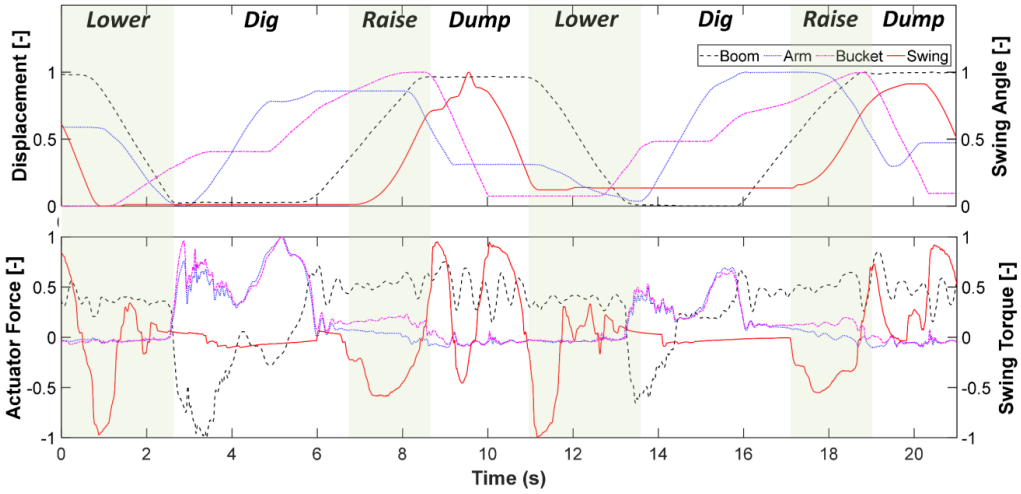


Figure 5: Actuator Displacements (top) and Hydraulic Forces (bottom) during a 90º digging duty cycle

5.2. Energy Analysis

Using the simulation results of the LS, DC and EHA systems with the digging cycle, the distribution of energy at the pump shaft is presented in this section. The losses in the three systems are classified as described in section 3.3. Firstly, the energy distribution for the LS system is shown in Figure 6. Since the simulation results are taken from pressure measurements at the actuator chambers, the actuator work done includes the transient and friction losses. Furthermore, the net actuator work includes the positive and negative (recoverable) work done by the actuator. This means that the negative work, not recovered by the LS system, is considered as part of the meter-out throttling loss. The energy distribution in Figure 6 shows that the useful or net work done by the LS system is 25% of the total pump shaft energy, and about 38% of the energy is lost in the valve block. The recoverable or negative work accounts for about 11% of the total energy. It is also shown that the swing and boom actuator work are only about 5% and 3% respectively, because they form a large portion of the total recoverable (negative) energy, which is lost due to throttling.

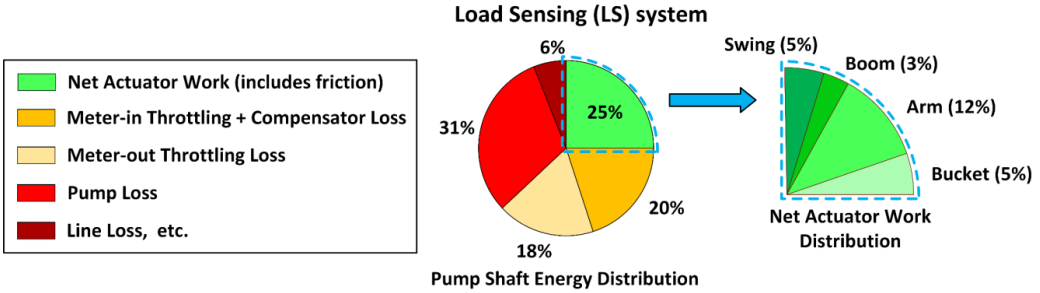


Figure 6: Energy Distribution for LS excavator during the 90º digging cycle

For the individualized architectures, the throttling and compensator losses are eliminated due to the use of independent pumps and the negative work is able to be recovered through mechanical or electrical energy recovery. The energy distribution for the DC and EHA systems are shown in Figures 7 and 8 respectively. The energy distribution of the pump shaft(s) shows an efficiency of 61% and 64.5% for the DC and EHA systems respectively. For the DC system, the energy consumption excluding the charge circuit losses can be split for the working actuators, whereas the complete distribution for the EHA system can be split because of independent charge systems. The relatively low efficiency of the boom actuator results from the pump losses occurring during motoring mode, when the work done is negative. The line loss distribution is the result of the actuator placement, with the swing placed closest to the pump and bucket the farthest. For the EHA system, the by-pass valve (BPV) induces some throttling in the system at low-speed actuation.

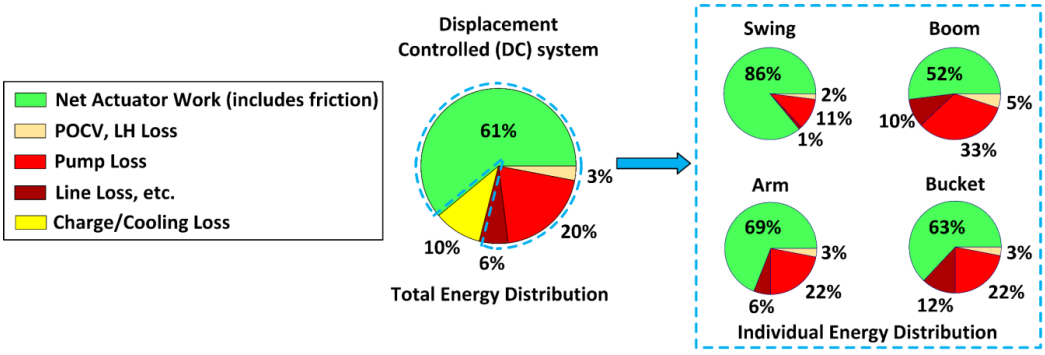


Figure 7: Energy Distribution for DC excavator during the 90º digging cycle

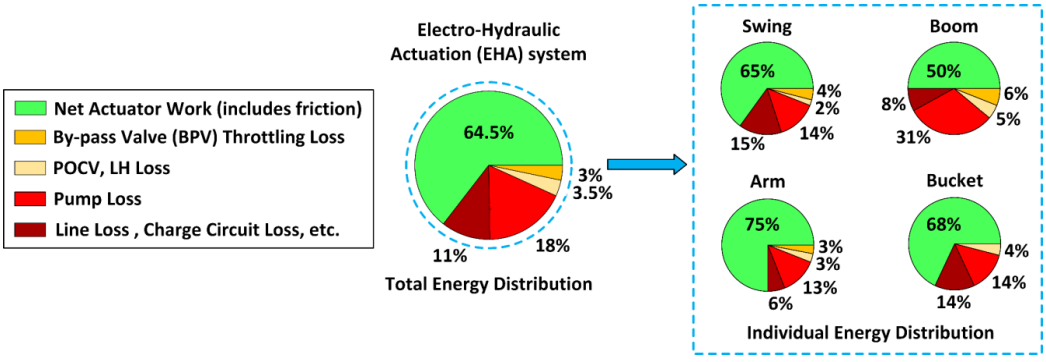


Figure 8: Energy Distribution for EHA excavator during the 90º digging cycle

5.3. Installed Power Analysis

Based on the simulation results, the power requirements for the electrified LS, DC and EHA systems for the digging cycle can be obtained. The operating points of the pump shaft(s) for the LS and individualized architectures is shown in Figure 9, where the LS and DC systems have a common shaft for the variable pumps, whereas the EHA system has independent pump shafts driven by separate electric motors. The figure shows that the operating torque range is lower than the LS system, which is due to lower energy consumption and higher speed capabilities of electric motors, and therefore, a lower displacement of the pumps for the same flow requirements. In contrast, for the EHA system with fixed displacement pumps, the shaft torque only depends on the pressure differential and is therefore, in general higher than the DC system as can be seen for the swing and arm actuators. Also, during negative (overrunning) work in, for example, the boom in simultaneous operation with the arm doing positive work, the energy recovery is mechanical in the DC system and therefore, reduces the torque on the electric drive. On the other hand, the form of recoverable energy is always electrical for the EHA system, which results in further conversion losses and no reduction in shaft torque.

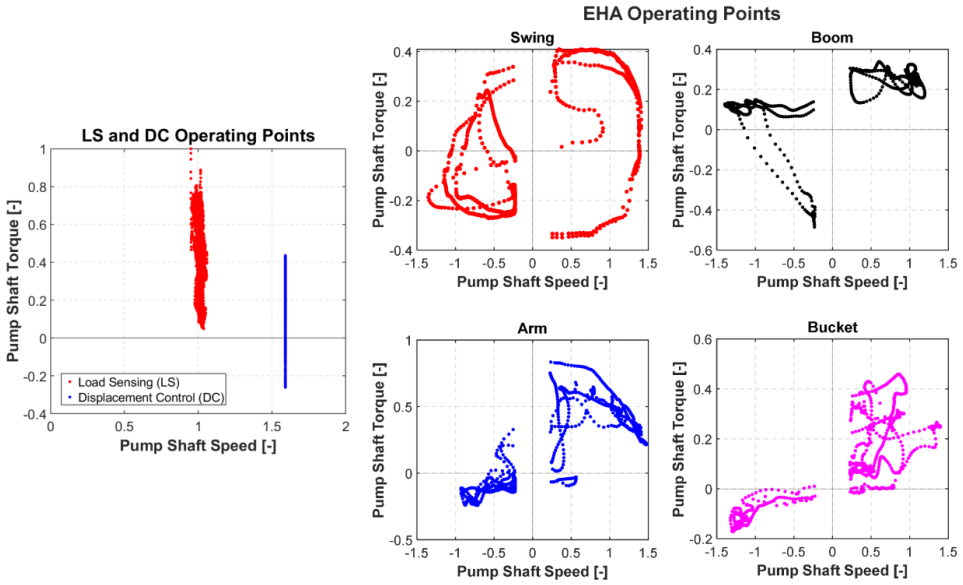


Figure 9: Operating Points (Normalized) for Pump Shaft(s) Power in LS, DC and EHA excavator during the 90º digging cycle

Based on the analysis of the duty cycles, the appropriate electric drives for the DC and EHA systems can be selected to provide adequate power during different phases of the cycles. In the scope of this study, only the digging cycle is considered for this purpose. Instead, the electric drives sizes for the secondary functions are chosen based on specific assumptions and knowledge of the reference machine. Based on the assumptions in section 4, the theoretical installed power of the electric drives relative to the reference LS system is listed in Table 2. Even though the DC system consumes less power during the digging cycle compared to the LS system, the installed power needs to be adequate for limiting operating conditions of single- or multi-actuator operation. The fact that a torque limiter is present in the reference machine to limit the maximum power of the LS system is used to limit the maximum torque for the DC system, so that the power of the electric drive is equal to the reference system. For the EHA system, since the actuators have separate electric drives, the installed power for the digging functions is 2.87 times higher than the LS installed power based on the simulation results

shown in Figure 9. For the remaining actuators, their limiting operating conditions and maximum torque limit of the LS system are utilized to obtain the power of the electric drives. The use of LS torque limit instead of power limit like the DC system is because the torque for fixed displacement pumps depends only on pressure differential. Therefore, the total installed power of the EHA excavator is 8.26 times the LS installed power. This shows that even if the energy consumption during digging for the two individualized excavator systems is similar, the EHA system requires a significantly higher installed power and number of electric drives compared to the DC system, which requires the same installed power as the LS system. However, the advantage of using EHA drives stems from the compactness of integrated systems and modular designs.

Table 2: Relative Theoritical Installed Power of the Electric Drives

System	Maximum Torque [-]	Maximum Speed [-]	Power [-]
LS	1	1	1
DC	0.625	1.6	1
EHA – Swing	0.43	1.37	0.54
EHA – Boom	0.5	1.37	0.69
EHA – Arm	0.8	1.37	1.1
EHA – Bucket	0.43	1.37	0.54
EHA – Blade	1	1.37	1.4
EHA – Offset	0.69	1.37	0.95
EHA – Auxiliary	0.64	1.37	0.88
EHA – Right Track	0.79	1.37	1.08
EHA – Left Track	0.79	1.37	1.08
EHA – Charge Pump	0.003	1.37	0.004
EHA – Total			8.264

6. CONCLUSION AND OUTLOOK

This paper presented a comprehensive study to assess the pros and cons, in terms of energy consumption and the installed power, of fully individualized hydraulic systems in an electrified mini-excavator. The reference case of an electrified load sensing (LS) system, as in the baseline vehicle, is considered against two fully individualized hydraulic architectures based on the Displacement Control (DC) and the Electro-Hydraulic Actuation (EHA) concept. While the DC system uses variable displacement pumps for each actuator with a common electric drive, the EHA system uses fixed displacement pumps with separate electric drives for each actuator. A breakdown of the energy losses in the LS, DC and EHA was given and appropriate sizing for the components was conducted in the proposed systems. Simulation models for the hydraulic systems of the architectures were built in Simcenter Amesim and a 90º digging cycle was simulated to analyse the energy consumption of the three systems. The digging cycle had four working actuators, swing, boom, arm and bucket, and a breakdown of the system power losses showed that the pump shaft energy efficiency increased by about 59% and 61% compared to the reference LS system for the DC and EHA systems respectively. However, the installed capacity of the pumps for both the systems increased significantly. In the DC system, the power of the electric drive installed was comparable to the reference LS system. In contrast, for the EHA system, with the increase in the number of electric drives, the total installed power of the units increased to about 8.3 times the LS installed prime mover power. In conclusion, this study shows that the full individualization of centralized hydraulic circuits of excavators can

notably reduce the energy consumption for both DC and EHA concept, but at the cost of significantly higher installed power of the components. While both DC and EHA increase the overall pump capacity, the EHA concept also increases the installed power of the prime movers by 8 times. Hence, a trade-off exists between the energy efficiency and the level of individualization for electrified hydraulic architectures.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Siemens Digital Industries Software for the use of Simcenter Amesim in the development of the simulation models for the study.

NOMENCLATURE

BPV	By-Pass Valve	
DC	Displacement Control	
EHA	Electro-Hydraulic Actuation	
ICE	Internal Combustion Engine	
LH	Load Holding Valve	
LS	Load Sensing	
ORV	Off-Road Vehicles	
POCV	Pilot Operated Check Valve	
p_A	Pressure in cylinder piston-side chamber (A)	[bar]
p_{acc}	Pressure in the accumulator	[bar]
p_B	Pressure in cylinder rod-side chamber (B)	[bar]
p_{pump}	Pressure at pump delivery port	[bar]
Q_{max}	Maximum flow rate of the pump	[L/min]

REFERENCES

[1] Quan Z, Ge L, Wei Z, Li YW, Quan L (2021) A Survey of Powertrain Technologies for Energy-Efficient Heavy-Duty Machinery. Proceedings of the IEEE 109:279–308

[2] Hagen D, Padovani D, Choux M (2020) Guidelines to Select Between Self-Contained Electro-Hydraulic and Electro-Mechanical Cylinders. 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), 547-554

[3] Vacca A, Franzoni G (2021) Hydraulic Fluid Power: Fundamentals, Applications, and Circuit Design. Hoboken, NJ: Wiley

[4] Zimmerman J, Pelosi M, Williamson C, Ivantysynova M (2007) Energy Consumption of an LS Excavator Hydraulic System. Proceedings of the ASME 2007 International Mechanical Engineering Congress and Exposition. Vol. 4, 117-126

[5] Pellegri M, Green M, Macpherson J, MacKay C, Caldwell N (2020) Applying a multi-service digital displacement® pump to an excavator to reduce valve losses. 12th International Fluid Power Conference 59-68

[6] Heybroek K, Sahlman M (2018) A hydraulic hybrid excavator based on multi-chamber cylinders and secondary control–design and experimental validation. International Journal of Fluid Power, 19(2), 91–105

[7] Weber J, Beck B, Fischer E, Ivantysyn R, Kolks G, Kunkis M, Lohse H, Lübbert J, Michel S, Schneider M, Shabi L, Sitte A, Weber J (2016) Novel System Architectures by Individual Drives.

10th International Fluid Power Conference 29-62

[8] Zimmerman J, Busquets E, Ivantysynova M (2011) 40% Fuel Savings by Displacement Control Leads to Lower Working Temperatures - A Simulation Study and Measurements. Proceedings of the 52nd National Conference on Fluid Power, Vol.1, 693-701

[9] Qu S, Fassbender D, Vacca A, Busquets E (2021) A high-efficient solution for electro-hydraulic actuators with energy regeneration capability. Energy, Elsevier, 216(C)

[10] Haack S, Flaig A (2022) Sustainable Hydraulics for Industrial and Mobile Application. Presented at the 13th International Fluid Power Conference, Aachen, Germany

[11] Pischinger S (2022) Future CO_2-Neutral Propulsion and their Impact on Fluid Technical Systems. Presented at the 13th International Fluid Power Conference, Aachen, Germany

[12] Casoli P, Scolari F, Minav T, Rundo M (2020) Comparative energy analysis of a load sensing system and a zonal hydraulics for a 9-tonne excavator. Actuators 9, 2: 39

[13] Padovani D, Fresia P, Rundo M, Altare G (2022) Downsizing the electric machines of energy-efficient electro-hydraulic drives for mobile hydraulics. Journal of Physics: Conference Series, 2385 012028

[14] Busquets E, Ivantysynova M (2015) A Multi-Actuator Displacement-Controlled System with Pump Switching - A Study of the Architecture and Actuator-Level Control. JFPS International Journal of Fluid Power System 8(2) 66-75

[15] Figge F, Schmitz K (2023) Simulative Comparison Approach of Electro-Hydrostatic Actuator Circuits in Excavators. Proceedings of the BATH/ASME 2023 Symposium on Fluid Power and Motion Control FPMC2023-111798

[16] Padovani D, Ketelsen S, Hagen D, Schmidt L (2019) A self-contained electro-hydraulic cylinder with passive load-holding capability. Energies 12, 2: 292

RUN-IN BEHAVIOUR AND WEAR ON HYDRAULIC PISTON SEALS – EVALUATION OF AN ENDURANCE TEST FOR PISTON ACCUMULATORS

peer reviewed

Tobias Schulze*,[1], Vladimir Boyko[1], Michael Lenz[1], Gonzalo Barillas[2], Mert van Dawen[2], Ejnar Jørgensen[3], Erik Garde[4]

[1] *TUD Dresden University of Technology, Chair of Fluid-Mechatronic Systems (Fluidtronic), Helmholtzstrasse 7a, 01069 Dresden, Germany*
[2] *Freudenberg Sealing Technologies GmbH, Ascheröder Strasse 57, 34613 Schwalmstadt, Germany*
[3] *Lind Jensens Maskinfabrik A/S, Kroghusvej 7, 6940 Lem, Denmark*
[4] *Vestas Wind Systems A/S, Hedeager 42, 8200 Aarhus, Denmark*

* Corresponding author: Tel.: +49 351 463-42603; E-mail address: tobias.schulze2@tu-dresden.de

ABSTRACT

Hydraulic accumulators are essential components in both industrial and mobile hydraulic systems, serving various purposes from energy storage to shock absorption and energy recovery. In hydraulic pitch systems of wind turbines, piston accumulators provide significant advantages, including reliability, resilience to centrifugal forces and temperature fluctuations, as well as simple monitoring. Ensuring a proper seal between the gas and oil sides of a piston accumulator and understanding its wear characteristics are crucial for a reliable operation of the system. A precise determination of wear often requires the use of measured values.

The current paper presents the results of a 5,000-hour endurance test conducted for piston accumulators under load conditions typical for wind turbine applications and reveals insights into the run-in behaviour and wear process of the seals. During and following the test, parameters such as sealing geometry and surface roughness of inner accumulator tubes were measured. The accordance of the measured wear with Archard wear model behaviour is assessed. It could be shown that error-free sealing conditions can be expected for a period of about 20 years.

Keywords: Hydraulics, Wind Turbine, Piston Accumulator, Seal, Sealing, Surface Roughness, Wear

1. INTRODUCTION

1.1. Hydraulic systems in wind turbines

The current political requirements for CO_2 reduction make a worldwide expansion of renewable energy sources necessary [1]. This is apparent in the massive expansion of wind power plants. In the last five years alone, the global installed wind power capacity has increased by 53% to 906 GW in 2022 [2]. Wind power currently accounts for 7.3% of the world's total electricity generation, which is around 50% more than in 2018 [3].

This results in increased sales of piston accumulators, typically used in hydraulic pitch systems of wind turbines for adjusting turbine blade angles and for emergency stops to turn the blades out of the wind. However, accumulators pose a significant safety risk in turbine systems. A study of offshore turbines identified the pitch system as the most unreliable subsystem, with accumulators contributing to 10.7% of its failures [4]. Gas leakage is a major issue with hydraulic accumulators, leading to a

loss of energy storage capacity. Monitoring gas leakage in such accumulators can significantly reduce the risk, so that many signal- and model-based methods for fault detection in piston accumulators have been developed in the recent years (see e.g. [5-7]). Some manufacturers offer monitoring systems for piston accumulators, such as laser or cable position measurement systems and limit switches [8].

The piston seal therefore plays a crucial role, as the functionality of the entire wind turbine depends on it.

1.2. Piston accumulators

Hydraulic accumulators are essential components of hydraulic systems in industrial settings and mobile machinery. They find application in a wide range of scenarios. As an energy storage they take up a volume of liquid under pressure and release it again when required, for example to cover extra flow demand in case of fluctuating requirements or in the event of failure or leakage. Furthermore, they are used to reduce pressure peaks and volume flow, while also serving as a hydropneumatic spring element.

This paper investigates the sealing system and wear behaviour of a new product of piston accumulators from Lind Jensens Maskinfabrik A/S, as shown in **Figure 1**. Main parts are the gas connection (1), the tube (2), the piston (3) with piston seals and guide rings, and the oil connection (4). The accumulator was equipped with an ILR2250-100 “micro epsilon” laser measurement system (5) and a sight glass (6) to measure the position of the piston. The sealing system is composed of a polyurethane (Freudenberg 98 AU 928, hardness 98 Shore A) profile ring (*piston seal*), an NBR preload element (*energiser*), as well as gas-side and oil-side guide rings made of phenolic thermoset material. The countersurface of the sealing on the inside of the tube is a roller burnished pressure vessel steel (EN 10028) with the measured initial average roughness of $Rz = 0.758 \pm 0.095$ µm and $Ra = 0.077 \pm 0.008$ µm (see also **Figure 10**).

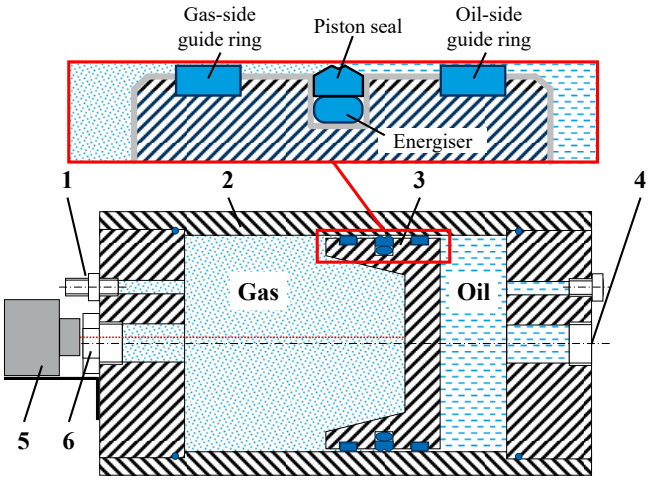


Figure 1: Schematic structure of the investigated piston accumulator

Piston-type accumulators, comprising a cylinder with a sealed, freely floating piston that separates gas (usually nitrogen) from oil, offer several advantages over bladder and membrane accumulators [9]. These advantages include the ability to operate at both high and low temperatures, and higher flow rates, resulting in higher gas compression ratios and better volume utilization rates. Additionally, manufacturing piston accumulators with various nominal volumes is particularly simple, as it can be

achieved by adjusting the tube length. In comparison to bladder and membrane accumulators, manufacturing costs of piston accumulators are relatively high. The disadvantage of piston accumulators is the high inertia of the piston, making them unsuitable for rapid charge-discharge processes. This can be partially mitigated by using lightweight materials for the pistons, such as aluminium. Within the piston accumulator, the gas and oil side are separated by the piston sealing. This imposes specific requirements on the piston seals to prevent the stick-slip effect and reduce friction, which, in turn, limits their potential application to high operating pressures. The sealing has to ensure the proper function of the accumulator throughout its whole operating life, so wear and aging of the seals need to be verified.

2. ENDURANCE TEST STAND

In order to characterise the wear behaviour of piston accumulator seals under operating conditions typical for the wind power industry and to be able to make a prediction on the expected service life of the seal, an endurance test was carried out. In total, 600,000 working cycles (corresponding to 5,000 hours of operation) were performed. The test rig consists of four individual pressure-regulated 20-liter piston accumulators (see **Figure 2, a)**) with the specifications described in **Table 1**.

The piston accumulators were mounted horizontally, which corresponds to a usual situation in a wind turbine (see **Figure 2, b)**). During operation, the oil chamber pressure p_{Oil}, the gas chamber pressure and temperature p_{Gas} and T_{Gas}, as well as the piston position of one accumulator $x_{Pis,1}$ were continuously recorded. The oil type used was the Mobil DTE 10 Excel 32. For oil conditioning and maintaining a constant operating temperature, an additional cooling circuit was included.

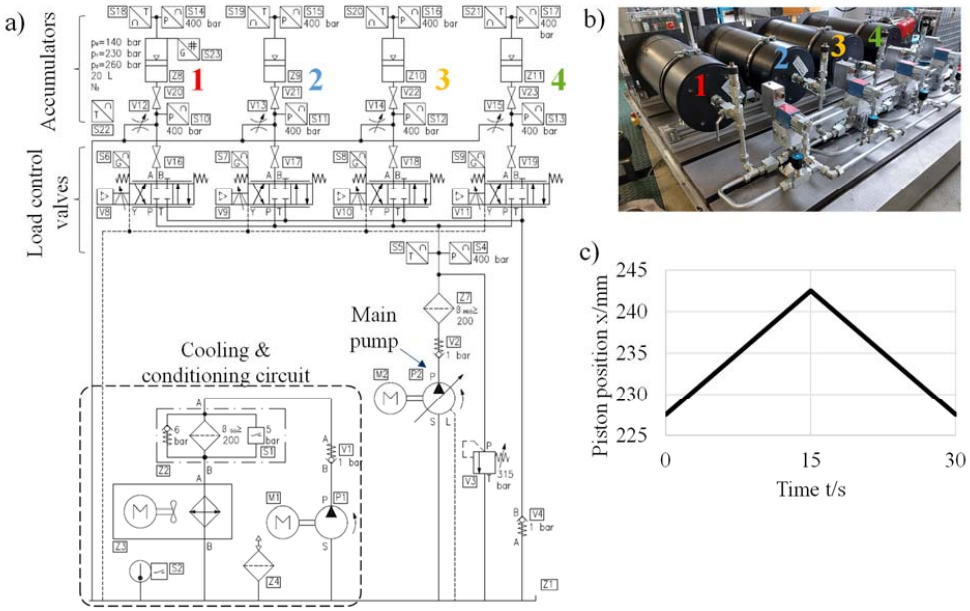


Figure 2: a) Hydraulic circuit of the test rig; b) Photo of the test rig; c) Piston target position

The piston had to maintain a central position within the accumulator, performing a 15 mm stroke from this point, as shown in **Figure 2, c**). To ensure a symmetrical load during the testing of the four cylinders, the test cycle was adjusted to a 30-second cycle with an equal volume flow for both charging and discharging.

Table 1: Geometry and operating conditions of the accumulators

Geometry/Condition	Value
Piston diameter D	240 mm
Piston stroke S	350 mm
Sliding speed $\dot{x}$	1 mm/s
Precharge pressure p_0	140 bar
Working pressure p_1 … p_2	230 … 260 bar
Oil temperature T_{Oil}	40 … 50 °C
Ambient temperature T_{Amb}	18 … 23 °C

3. EVALUATION METHOD

The wear behaviour of the accumulator was mainly determined by measuring the geometry of the piston seal, the surface roughness of the tube and by evaluating the measured process values of the test stand. All four accumulators were disassembled at intervals of 200,000 cycles for measurements of seal geometry and inner tube surface roughness. Following each measurement, the pistons were reinstalled in their original orientation to avoid any rotational variability. To determine the run-in behaviour, additional measurements were conducted at accumulator 1 after 1,500, 3,000 and 4,500 cycles (resp. 12.5, 25 and 37.5 hours of operation).

The process values were recorded for each 100th cycle, evaluated according to their absolute or average value and compared throughout the 600,000 cycles. In particular, the following values were evaluated:

Table 2: Evaluated process values of the test stand

Process value	Description
Precharge pressure	Before each disassembly process, the precharge pressure of the nitrogen was documented and compared to the value from the previous start-up
Piston position	The absolute position of the piston was measured with a laser measurement system. A change of the piston position would indicate gas leakage.
Pressure difference (gas and oil side)	The pressure on the gas and oil side of each accumulator was measured; their difference is an indicator for changing friction forces that might be caused by increased wear on the sealing setup of the piston
Temperature	Oil- and gas temperature were measured at each accumulator in order to monitor and evaluate temperature influences

To determine changes in the seal geometry, several geometrical measurements were taken. The cross section height and width of the piston seals, energisers, and guide rings were measured with a micrometer (resolution: 10 µm, error: 4 µm) at 12 points around the section (0°, 30°, 60° … 330°) after each test stage following the manufacturer's markings on the seals in order to ensure a high degree of reproducibility of the measurement. Additionally, the cross section profile of the sealing was measured with a profilometer Keyence VR-6000 (error: 4 µm). The weight of the seals and guide rings was measured with an analytical balance scale (Kern ABT 120 scale, linearity ±0.2 mg). The seal's inner diameter was determined using an 8-point assessment with a DMG DMU Eco80 linear CNC-machine. All values were determined in the dismantled, unstressed state with oil removed from the surface. **Figure 3** provides an overview of the seal setup at the piston and the measured variables.

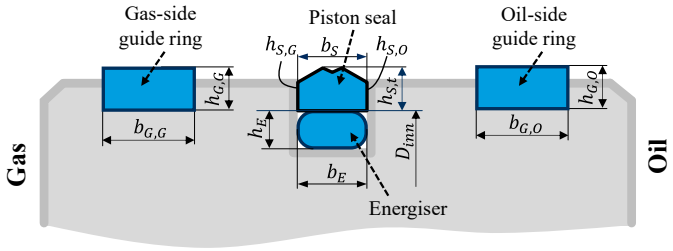


Figure 3: Seal setup and measured geometrical values

To give a brief insight into the measurement procedure and the occurring deviations, **Table 3** shows the measurements for the piston seal height at accumulator 1 after 400,000 cycles. It can be seen that the standard deviation of the micrometer measurements is 17 µm. In order to reduce the influence of random measurement errors, all further evaluation of geometric dimensions is based on the mean value of all 12 measuring points. The results are verified through additional measurement with the profilometer. It shows the decrease in the cross section height (compared to a factory-new piston seal) with a much lower standard deviation of 3.4 µm.

Table 3: Measurements for piston seal at accumulator 1 after 400,000 cycles

Micrometer		Profilometer	
Measurement point	Total height h_t	Measurement point	Decrease in total height Δh_t
0°	4.030	0°	0.079
30°	4.020		
60°	4.025	60°	0.082
90°	4.000		
120°	4.005	120°	0.084
150°	4.010		
180°	4.015	180°	0.084
210°	3.985		
240°	3.980	240°	0.074
270°	3.985		
300°	3.990	300°	0.077
330°	4.020		
Mean value $\bar{h}_t$	4.005	Mean value $\overline{\Delta h_t}$	0.080
Std. deviation σ	0.0170	Std. deviation σ	0.0034

4. RESULTS

4.1. Run-in process

The factory-new sealing setup on the piston is expected to show a run-in process during the first operating hours, with increased wear and more rapidly changing process values, until the friction couples between the seal, guide ring, and tube have developed their individual stabilized friction state. Therefore, especially the measured cross section height of the sealing lips is expected to decrease more rapidly during these first operating hours.

To prove this assumption and determine the required time for the run-in process, additional measurements on accumulator 1 were conducted during the beginning of the experiments after 1,500,

3,000, and 4,500 cycles. **Figure 4** shows the width, total cross section height, and mass of the piston seal of accumulator 1. It can be seen that after 1,500 cycles (equivalent to 12.5 hours of operation), the run-in process is mostly completed, as there are only minor changes in the geometrical measurements. Changes in the cross section width show an increase by 18 µm after 1,500 cycles (20 µm after 3,000 cycles, 22 µm after 4,500 cycles respectively). The total cross section height h_t decreased by 63 µm after 1,500 cycles (66 µm after 3,000 cycles, 75 µm after 4,500 cycles respectively). The mass of the sealing increases by 48 mg after 1.500 cycles (53 mg after 3,000 cycles, 62 mg after 4,500 cycles respectively), which is attributed to swelling overlaying a possible reduction due to abrasive wear. These values are confirmed by further measurements after 200,000, 400,000 and 600,000 cycles, as the further changes are small.

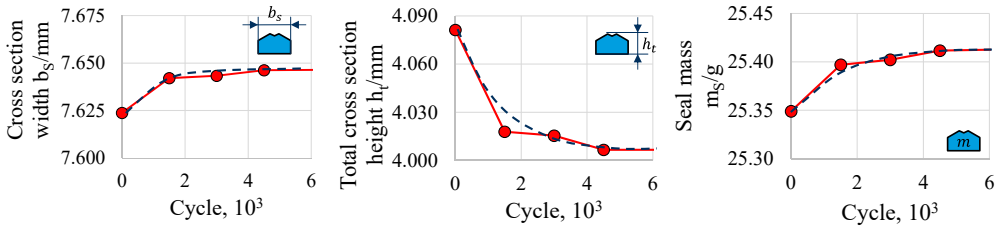


Figure 4: Evaluation of the run-in process at accumulator 1

4.2. Characterization of friction

An additional value to evaluate the stabilized friction couple on the seal is the resulting pressure difference between gas and oil side of the seal $p_d = p_{Oil} - p_{Gas}$. Given that no other forces are acting on the piston and inertia forces are minimal, p_d is proportional to the frictional force F_{fr} acting on the sealing system. The comparison between different cycles provides an indicator for changing friction properties over the course of the experiment, as the trajectory is controlled to remain uniform. **Figure 5** shows the measurement results for accumulator 1 for cycle 201,000. It presents the measured piston trajectory (actual distance from the laser measurement system to the piston, see position '5' in **Figure 1**), the gas pressure p_{Gas}, gas temperature T_{Gas}, as well as p_d. The pressure difference shows a characteristic curve with a slight increase throughout the piston movement and a sign change when the piston reverses.

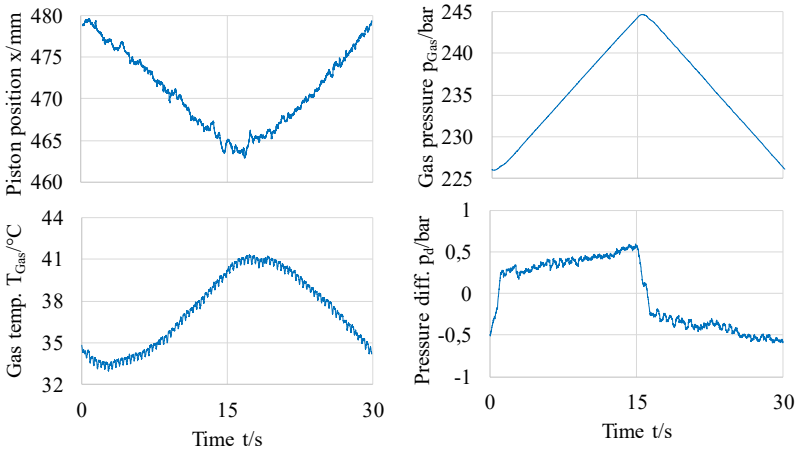


Figure 5: Measurements for cycle 201,000 at accumulator 1

For the presented measurements at cycle 201,000, it is assumed that the run-in-process is completed and the friction couple has been fully stabilized. However, at cycle 200,000, the seal was inspected for wear, which involved full dis- and reassembly. Comparing the values of the pressure difference p_d throughout all cycles, it was observed that after each reassembly p_d reaches significantly higher values as compared to immediately before inspection.

Figure 6 compares the pressure difference in all accumulators shortly after the inspection events and again after additional 100,000 cycles. Inspections of all accumulators were carried out after 4,500, 200,000, 400,000 and 600,000 cycles. The figure shows that shortly after reassembly (displayed in red colours) a higher value of pressure difference was measured in most cases. Each accumulator shows a characteristic shape for the course of the pressure difference throughout the respective cycle shortly after reassembly. The stabilized state after additional 100,000 cycles (displayed in blue colours) is constant throughout the different cycles and repeats over the reassembly events.

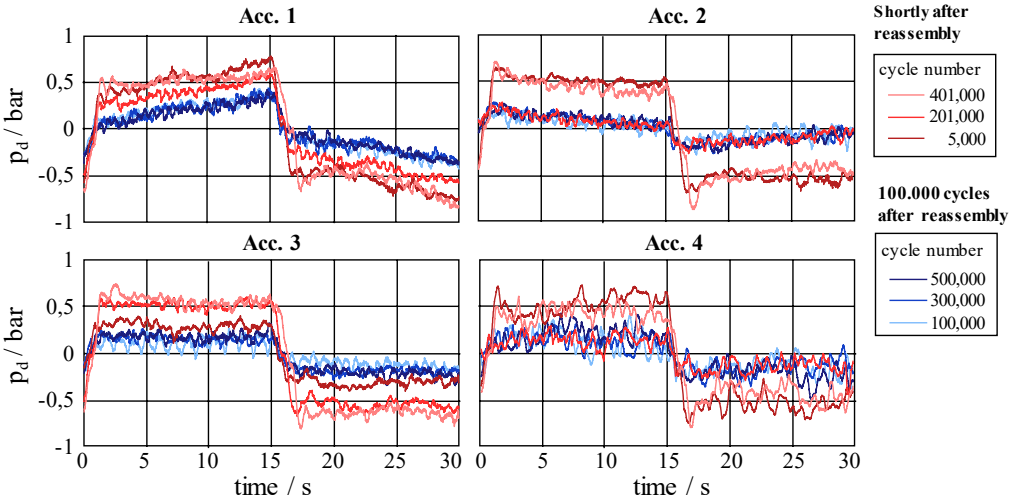


Figure 6: Pressure difference in accumulators shortly after inspection and after stabilization

4.3. Precharge pressure

To evaluate the potential internal and external leakage of the working gas and hydraulic oil, the precharge pressure change before and after the last 200,000 cycles was measured at the same gas temperature of T_{Gas} = 19.5 °C. A slight increase in precharge pressure by 0.25% on average could be observed, indicating a minor internal oil leak from oil side to gas side. Applying the isothermal equation of state for ideal gases, the permeated oil volume can be estimated at 40 ml.

Table 4: Precharge pressure changes in the last test stage (400,000 to 600,000 cycles)

Parameter	Acc. 1	Acc. 2	Acc. 3	Acc. 4
Start precharge pressure $p_{0,400}$/bar	140.25	140.31	140.20	140.14
End precharge pressure $p_{0,600}$/bar	140.54	140.66	140.51	140.45

Since a visual inspection of the accumulators after disassembling did not reveal significant traces of oil, the measured increase can also be partially attributed to the inaccuracy of the pressure sensors. No oil mist could be observed while draining the gas.

4.4. Changes in the sealing

From the measurement results of the piston seal presented in **Figure 7**, a noticeable change in the cross section height is evident, which can be attributed to the initial run-in process. After approximately 4,500 cycles, the seal height continued to gradually decrease at a notably diminished rate. The height on the oil side showed a faster decline compared to the gas side (0.44% average reduction vs. 0.27% on the gas side), presumably attributable to chemical reactions occurring due to contact with the oil, thus accelerating the wear. The total height reduction after 600,000 cycles amounted to 0.06 mm or 0.32%.

Regarding the sealing width, it exhibited an initial increase during the first 200,000 cycles and subsequently reached a near-constant state, showing minimal variation. Even after 600,000 cycles, the width increase remained below 0.05 mm. The sealing weight also showed a similar pattern, increasing during the first test stage due to oil impregnation and remaining nearly constant after 200,000 cycles. The total weight increase after 600,000 cycles is estimated at 0.2 g. The inner diameter of the seals underwent a reduction following the run-in period and subsequently changed within a range of 0.1 mm.

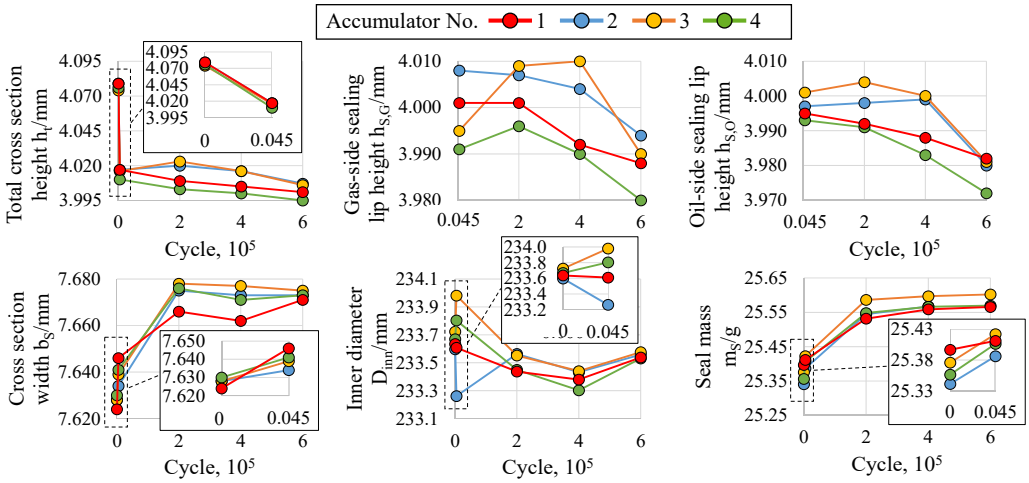


Figure 7: Measurement results of the piston seal

Estimating the remaining service life based on available data can be challenging and requires further studies, such as accelerated life-time tests to determine wear. However, by assuming a critical seal height change of 12% (manufacturer specification), after which a gas leakage becomes likely, and extrapolating the total cross section height h_t in **Figure 7** linearly, the remaining service life can be estimated to be approximately 20 years.

In addition to the manual measurement of the sealing lip height, a profilometer Keyence VR-6000 was used to determine the cross section profile of the piston sealing. The profile shows wear on the sealing lips with rounder edges compared to the factory new and unused profile (see **Figure 8**). By evaluating the depth of the groove between both sealing lips, the measurements shown in **Figure 7** are confirmed and a reduced sealing height of about 0.08 mm after 600,000 cycles was measured. This measurement verifies the previous values measured with the micrometer which also determined a decrease by 0.08 mm.

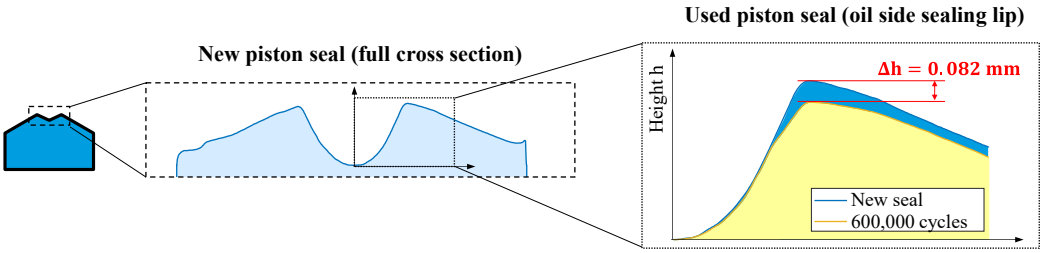


Figure 8: Cross section of the sealing lips and evaluated groove depth

Both oil-side and gas-side guide rings experienced a decrease in height and width (s. **Figure 9**). Similar to the seal ring, the deformation of the guide ring on the oil side was more pronounced compared to the gas side. On average, the height of the oil-side guide ring decreased by 1.77%, while the gas-side guide ring exhibited a slightly lower average height decrease of 1.19%. This difference may also be due to measurement inaccuracies. The mass of the oil-side guide ring showed a significant increase, approximately 1.26% higher than the gas-side ring, which could also potentially be attributed to the oil absorption.

The measurements performed on the energiser of the accumulator 1 did not detect any changes in the geometry.

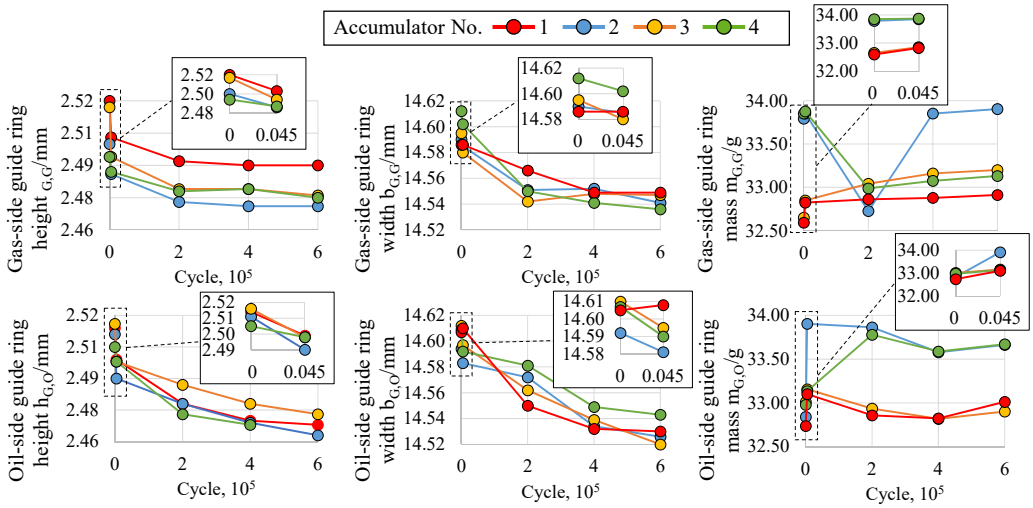


Figure 9: Measurement results of the gas-side (upper row) and oil-side (lower row) guide ring

4.5. Roughness of the inner tube surface

In addition to measuring the seal geometry, the roughness values *Ra* and *Rz* of the inner surface of the accumulators in the area of the piston movement were measured. These measurements were taken every 200,000 cycles using a profilometer MahrSurf PS10 at four points around the circumference. Furthermore, the full peak height *Rpkx*, the kernel roughness depth *Rk*, and the full valley depth *Rvkx* were determined for the accumulator 1 before and after the test. **Figure 10** demonstrates the change in the roughness values over the entire duration of the experiment.

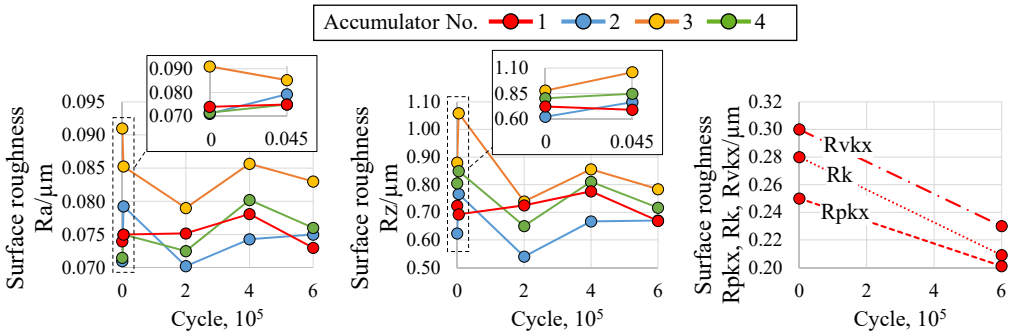


Figure 10: Change in surface roughness over the test period

No significant reduction during the test can be read from the usual values *Ra* and *Rz*. The fluctuations in the value are attributable to the measurement deviations of the instrument. However, a significant reduction in the peak, kernel, and valley depth can be observed, indicating the run-in and polishing effect.

5. WEAR EVALUATION AND LIFETIME PREDICTION

The observed evolution of cross section wear is assessed by means of a simplified Archard [10] wear model. The regular Archard model predicts the volume of material lost to wear V_W with the equation

$$V_{\mathrm{W,A}} = k \int \frac{F_\mathrm{N}}{H} \mathrm{d}s, \tag{1}$$

F_N being the normal force acting in the contact, s the sliding distance between the friction partners, H the hardness of the softer material, and k an empirical constant.

While the literature reports a limited applicability of the Archard model for elastomer friction on hard surfaces, a variety of more elaborate models for elastomer wear, often modifications of the Archard model, have been proposed to overcome its shortcomings – see e.g. [12] for a non-comprehensive overview. These studies are typically concerned with deriving general models for predictive simulation in the context of elastohydrodynamics (EHD). For this experimental assessment, the original model's simplicity is considered favourable. This is because micromechanical processes are not observable, and typical model parameters such as sliding velocity and load are kept constant in the experimental setup. Therefore, parameters of refined models are either inaccessible or irrelevant. Model simplicity should also be considered in relation to its potential application in a condition monitoring context, where extensive parameter identification may not be feasible and a limited number of signals must suffice.

A proportional friction law $F_\mathrm{fr} \propto F_\mathrm{N}$ is assumed, since the sliding velocity as a further parameter is kept constant during the experiment. The influence of variations in temperature, which was controlled to stay constant, is neglected. With furthermore constant hardness H, and observing $p_\mathrm{d} \propto F_\mathrm{fr}$ as stated in Section 4.1, the model reduces to an expression

$$V_\mathrm{W} = K \int p_\mathrm{d} \, \mathrm{d}s \tag{2}$$

The result is an abstracted model postulating wear volume as proportional to friction work.

Since the sliding distance was not measured directly for all four accumulators, it is derived from p_Gas by supposing ideal gas behaviour such that

$$s = \frac{s_0 p_{\mathrm{Gas},0}}{T_{\mathrm{Gas},0}} \cdot \frac{T_{\mathrm{Gas}}}{p_{\mathrm{Gas}}}, \tag{3}$$

where 0 denotes the state variables at the moment of initial charging of the accumulators, and s is measured from the rear wall of the cylindrical gas volume.

The evaluation of the integral V_{W} with as yet unknown constant K is shown in **Figure 11**. The values are broadly comparable between the accumulators, although accumulator 2 has the markedly lowest. The figure highlights the inspection events at 4,500 and 200,000 cycles. As reported in Section 4.2, the pressure difference tends to increase clearly after inspections, leading to an increased value of the per-cycle integral. **Figure 11** shows how this comparatively high value stabilizes or decreases only slowly over a couple 1,000 to up to several tens of thousands cycles (between a few hours to several days of operation) and then changes rather abruptly back to a value that is similar to the previously stabilized friction state before the disassembly. This indicates a sudden change towards a more favourable tribological regime and is observed for all four accumulators for each inspection with exception of accumulator 3, at 4,500 cycles, where the value is lower after inspection, and accumulators 2 and 4 at 200,000 cycles, where a higher value appears right after disassembly but decreases more rapidly. The number of cycles for stabilizing differs between the accumulators and appears to be random, indicating a spontaneous transition between the stable tribological regimes.

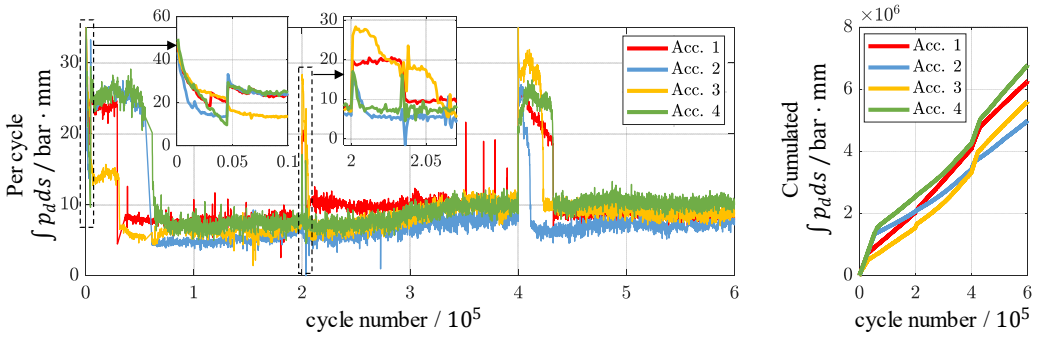


Figure 11: Evaluation of wear integral

The representation of the per-cycle wear integral confirms the spontaneous change towards a much lower value after about 1,000 up to 10,000 cycles after inspection. Due to its diverse influencing factors, this behaviour can unfortunately not be discussed in further detail within the limited space of this publication.

During the favourable equilibrium, the per-cycle integral, which according to the models is a measure for wear, is markedly highest for accumulator 1, in particular after the inspection at 200,000 cycles. Accumulators 3 and 4 exhibit comparable rates during those periods, while it is notably lower for accumulator 2. During the periods of increased per-cycle value immediately after inspections, the accumulated integral increases much faster, and in particular the phases after the first and the third inspection contribute significantly to the overall value of the wear measure attained during the test.

In **Figure 12**, the computed wear integral is exemplarily depicted in relation to the observed changes in the sealing lip height presented in **Figure 7**. The wear model according to Equation (2) would postulate a linear dependency, with the same slopes K for all test specimens for each of the wear locations.

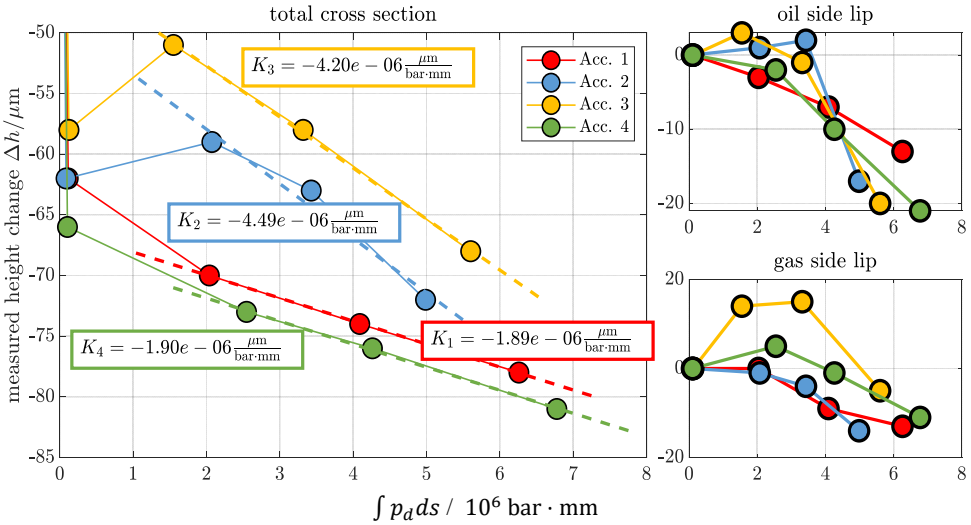


Figure 12: Wear integral vs. observed wear

The presented results cannot be interpreted to support this on a global level for neither of the wear locations discussed here. Possible explanations are:

a) the model being overly simplistic and hence inapplicable for the tribological system;

b) as discussed in Section 4.4, the quantitative assessment of the sealing wear during the tests is complicated by the superposition of wear and swelling, and by the low total amount of observed wear and associated measurement error. The resulting uncertainties complicate the model assessment;

c) the measured cross section and sealing lip heights being imperfect proxies for the volume of worn material from Equation (3), the quality of which could not be comprehensively assessed over the experiment; however **Figure 8** gives an indication.

The total cross section diagram does indicate a partial relation between the model and the measured results for the higher cycle numbers. If only the data points pertaining to the inspections at 200,000, 400,000 and 600,000 cycles are considered, linear fits of high significance, with values of $R^2 = 0.9997$, 0.9686, 0.9992 and 0.9988 for the four accumulators respectively, are achieved. There is a striking accordance of the slopes for accumulators 1 and 4, and those for 2 and 3. A possible interpretation is the model failing for the run-in process up to the 200,000 cycles inspection, at which point the wear process has stabilized sufficiently for the model to capture its evolution. The two different slopes may be explained by accumulators 2 and 3 having undergone swelling during the initial phase of the experiment to an increased degree, to the point where the cross section height increased during operation despite wear. In this interpretation, two distinct friction pairings have formed, with different associated wear constants K. It is acknowledged that more data points for each accumulator would be desirable in order to allow for a confident statement regarding the model's suitability for the later phase of the experiment, which at this point must remain speculative to a degree. The graphs for the two single sealing lip heights do not show a similar indication of partial accordance with the model. However, as stated in Section 4, the associated measurements are associated with higher uncertainty. Further tests with higher cycle numbers for a larger, more significant quantity of observed wear, as well as a revised experimental setup in order to exclude uncertainties, could provide a basis to confidently evaluate the applicability of the model or find a

suitable alternative. This is, however, of limited feasibility considering the required high cycle counts, and motivates an accelerated testing scheme.

6. CONCLUSION AND OUTLOOK

The conducted experiments deliver results for wear in piston accumulators on the example of an application-driven test cycle. It was shown that the run-in process of the sealing is completed after only few cycles and about 12 hours of operation. Extensive measurements of the sealing geometry confirm a safe operation of the piston accumulator throughout its desired live cycle.

The wear models according to Holm-Archard and Fleischer were applied to the measurements. The evolution of the wear integral indicates a wear intensive tribological regime after assembly, including after inspections, which after several thousands to tens of thousands of cycles changes spontaneously to a stable, more favourable regime. With the available experimental results, only a limited assessment of the model applicability is possible. It clearly fails during the early stages of the experiment. The wear measure of total cross section height appears to follow the model behaviour for higher cycle numbers > 200,000, but the available amount of data points, together with measurement uncertainties, make a verdict difficult.

While extended tests would be desirable to fully assess the evolution of wear in the component, the long cycle times lead to a prohibitively long duration. This motivates the design of an accelerated test procedure. To provide meaningful results for the component under its operation conditions, such tests must preserve a comparable tribological regime, characterized by the Hersey number and the dynamic stiffness of the sealing material, which both depend on the sliding velocity in the contact. An increase in velocity could be compensated by targeted control of temperature – and hence viscosity – and pressure, so as to maintain both quantities in the same range as under operating conditions.

NOMENCLATURE

Symbols

b	Cross section width	mm
D	Diameter	mm
e^*	Energy density of the tribological system	J
F	Force	N
h	Cross section height	mm
H	Material hardness of the softer friction partner	-
k	Empirical wear coefficient	-
m	Mass	g
p	Pressure	bar
Ra	Average roughness	µm
Rk	Kernel roughness depth	µm
$Rpkx$	Full peak height	µm
$Rvkx$	Full valley depth	µm
Rz	Ten-point mean roughness	µm
S	Piston stroke	mm
T	Temperature	°C
V	Volume	l
W	Wear volume	mm³
x	Piston position	mm

Indexes and Abbreviations

0	Initial/precharge state
1	Minimum working pressure
2	Maximum working pressure
d	Difference between oil- and gas-side
fr	Friction force
G,G	Gas-side guide ring
G,O	Oil-side guide ring
Gas	Gas side
Inn	Inner diameter
N	Normal force
Oil	Oil side
S	Piston seal
S,G	Gas-side sealing lip
S,O	Oil-side sealing lip
t	Total

REFERENCES

[1] European Commission (2020) Stepping up Europe's 2030 climate ambition. Investing in a climate-neutral future for the benefit of our people, COM/2020/562

[2] Statista Research Department (2023) Cumulative installed wind power capacity worldwide from 2001 to 2022, online: https://www.statista.com/statistics/268363/installed-wind-power-capacity-worldwide/ (accessed on 15 October 2023)

[3] Statista Research Department (2023) Share of electricity generation from wind energy sources worldwide from 2010 to 2022, online: https://www.statista.com/statistics/1302053/global-wind-energy-share-electricity-mix/ (accessed on 15 October 2023)

[4] Carroll J, McDonald A, McMillan D (2015) Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. Wind Energy, pp. 1107-1119

[5] Pfeffer A, Glück T, Kemmetmüller W, Kugi A (2015) State of charge estimator design for a gas charged hydraulic accumulator. Journal of Dynamic Systems, Measurement and Control, 137(6)

[6] Liniger J, Soltani M, Pedersen H C, Carrol J, Sepehri N (2017) Reliability based design of fluid power pitch systems for wind turbines. Wind Energy, 20(6), June, pp. 1097-1110

[7] Liniger J, Sepheri N, Soltani M, Pedersen H C (2017) Signal-based gas leakage detection for fluid power accumulators in wind turbines. Energies, 10(3), Mar., pp. 1-18

[8] Hydac International (2023) Accumulator Technology. Product Catalogue EN 3.000.18/04.21

[9] Gebhardt N, Weber J (2020) Hydraulik – Fluid-Mechatronik. Springer, Berlin

[10] Archard J (1953) Contact and rubbing of flat surfaces. Journal of applied physics, 24(8), pp. 981-988

[11] Hakami F, Pramanik A, Basak A K, Ridgway N (2019) Elastomers' wear: comparison of theory with experiment. Tribology International, pp. 46-54.

EFFICENCY DEFINITIONS OF HYDRAULIC TRANSFORMERS AND FIRST TEST RESULTS OF THE FLOATING CUP TRANSFORMER (FCT80)

Robin Mommer*, Sef Achten, Jeroen Potma, Jasper Achten, Peter Achten

INNAS B.V., Nikkelstraat 15, 4823AE Breda, The Netherlands

* E-mail address: rmommers@innas.com

ABSTRACT

The FCT80 is a new hydraulic transformer, based on the floating cup principle. The transformer has valve plates with three ports, and is controlled by means of setting the rotational position of these valve plates. Contrary to hydraulic pumps and motors, there are no standardised efficiency definitions for this type of integrated transformers. In this paper, general definitions for efficiency and power loss for hydraulic transformers are proposed, which also take the compressibility of the oil into account. This paper also describes and presents the first test results of the overall efficiency of the FCT80.

Keywords: hydraulic transformer, efficiency, floating cup

1. INTRODUCTION

It is well known, that fluid power systems are not very energy efficient. In [1], the authors estimate an average overall hydraulic system efficiency of 22%. Three significant factors for a low overall system efficiency are:

- dissipative control by means of proportional valves
- inability to recuperate energy
- inefficient operating conditions for the main (variable displacement) supply pump

Common pressure rail (CPR) systems and secondary control provide a solution to these three issues [2]. Several studies have concluded that the introduction of such system architectures can reduce the energy consumption of hydraulic systems by 30% to over 50% compared to load-sense systems [3–7]. Hydraulic transformers can be essential for the realisation of CPR-systems.

Hydraulic transformers provide a non-dissipative way to control the power delivery to the individual work functions. Furthermore, transformers can recuperate energy, which can then be used by other loads or stored in hydraulic accumulators. Additionally, by using accumulators, the energy source (the supply pump) is decoupled from the loads. Therefore, the inefficiently operated, variable displacement supply pump, can be replaced by a more efficient (and less expensive) constant displacement pump. Finally, because of the drastic reduction in losses, less cooling capacity will need to be installed to control the fluid temperature.

While several prototype hydraulic transformers have been developed and tested [8–10], there currently is no hydraulic transformer that is commercially available. Therefore, there is no standard definition for the efficiency of these machines. In this study, a definition for the power loss and efficiency of hydraulic transformers is derived that also takes into account the compressibility of the oil.

Furthermore, an attempt is made to divide these losses into flow and torque losses, based on the theoretical displacement of the transformer. This analysis provides a good theoretical understanding of the basic operation of a hydraulic transformer, as well as give some indication of where losses in such a device can be expected. However, it will be shown that this theoretical division cannot provide an accurate representation of the different losses in an actual transformer.

The derived definitions for overall power loss and efficiency are used to map the performance of a new hydraulic transformer prototype. This 80 cc/rev Floating Cup type hydraulic transformer has been designed and developed by Innas [11].

2. EFFICIENCY DEFINITION

2.1. Hydraulic power

The classical definition of hydraulic power is simply the product of pressure p and flow rate Q. Several studies have suggested that this definition is no longer sufficient, and the hydraulic power should also include a small amount of compression energy [12–15]. According to [12], the hydraulic power P in a mass flow, including compression energy, can be approximated using (1).

$$P = pQ\left(1 + \frac{p}{2\overline{K}_s}\right) \tag{1}$$

with $\overline{K}_s$ the average isentropic bulk modulus of the oil. Please note that (1) describes the hydraulic power in the fluid with respect to absolute zero pressure. Since the pressure in a hydraulic system will never drop down to absolute zero, there is always a certain amount of hydraulic power present in the oil. Therefore the power consumption of a hydraulic component is always defined as a change in power, as will be shown next.

2.2. Motors and pumps

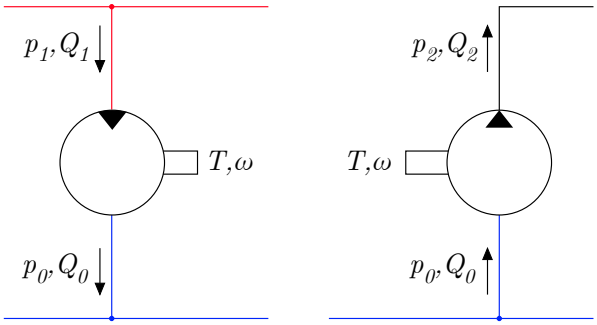
Figure 1a shows a hydraulic motor without an external leak port, that is connected to a supply pressure p_1 and an output pressure p_0 ($p_1 > p_0$). Since the output pressure is not absolute zero pressure ($p_1 > 0$), the hydraulic power that is consumed by this motor, P_{in}, is found to be the difference between the hydraulic power that is available at the input port and the power that remains in the fluid that leaves the motor at the output port.

$$P_{in} = p_1 Q_1\left(1 + \frac{p_1}{2\overline{K}_s}\right) - p_0 Q_0\left(1 + \frac{p_0}{2\overline{K}_s}\right) \tag{2}$$

The motor converts this hydraulic power into mechanical power in the form of torque and rotary speed (T, ω, respectively). In [14], the authors show that, at least for a hydrostatic machine without external leak flow, the hydraulic power can be defined with respect to any reference pressure level. Choosing p_0 as the reference pressure results in (3).

$$P_{in} = \hat{p}_1 Q_1\left(1 + \frac{\hat{p}_1}{2\overline{K}_s}\right), \qquad \text{with } \hat{p}_1 = p_1 - p_0 \tag{3}$$

For a pump, the process of power conversion is reversed: mechanical power is converted to hydraulic power. **Figure 1b** shows a hydraulic pump without an external leak port. The hydraulic output power of this pump, P_{out}, is given by (4).



(a) motor (b) pump

Figure 1: A hydraulic motor converts hydraulic power into mechanical power, a pump converts mechanical power into hydraulic power.

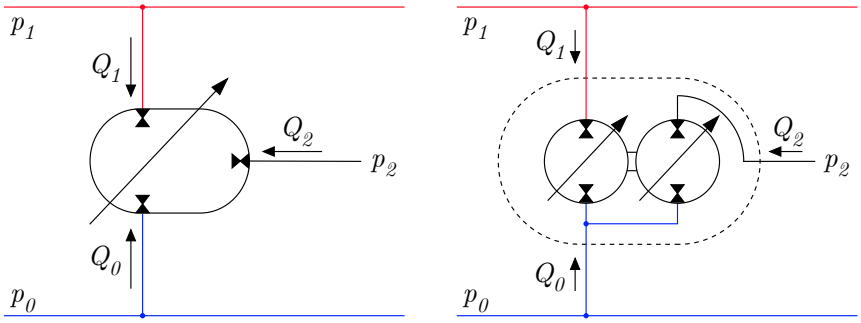
$$P_{out} = \hat{p}_2 Q_2 \left(1 + \frac{\hat{p}_2}{2\bar{K}_s}\right), \qquad \text{with } \hat{p}_2 = p_2 - p_0 \tag{4}$$

where p_2 and Q_2 are the pressure and flow rate at the discharge side of the pump.

2.3. Hydraulic transformers

Traditionally, the mechanical input power for a pump is supplied by an internal combustion engine, or an electric motor, while the output power of a motor can be used at a work function. The basic idea of a hydraulic transformer is to use hydraulic power to generate mechanical power, that is then converted back to hydraulic power again. In that sense, a hydraulic transformer can be seen as a combination of a motor and a pump, as is illustrated in **Figure 2b**.

Figure 2a shows the symbol for a hydraulic transformer. To function correctly, such a device is part of a CPR-system. The CPR-system consist of a high-pressure rail (HPR), at pressure p_1, and a low-pressure rail (LPR), at pressure p_0. A transformer uses the pressure difference between these two rails to control the power that is send to a load via the third port, at pressure p_2. The direction of flow rate Q at the different ports is not fixed; it can flow in either direction. In this study, flow rates are defined positive when oil flows into the transformer.



(a) hydraulic transformer (b) combined pump/motors

Figure 2: A hydraulic transformer (a) can be interpreted as a combination of two pump/motors (b), and is connected to a CPR-system and a load.

Note that the only difference between the input power for a motor, from (3), and the output power for a pump, from (4), is the direction of the flow rate. Since the transformer is able to deliver power to the load ($Q_2 < 0$), as well as recuperate energy from the load back to the CPR-system ($Q_2 > 0$), there is no clear definition of input or output power. Instead, we get the following equilibrium of power:

$$\hat{p}_1 Q_1 \left(1 + \frac{\hat{p}_1}{2\overline{K}_s}\right) + \hat{p}_2 Q_2 \left(1 + \frac{\hat{p}_2}{2\overline{K}_s}\right) = P_{loss} \tag{5}$$

in which P_{loss} is the total amount of power loss within the transformer. Since the efficiency η is the ratio between output and input power, we need to differentiate between the two modes of operation.

$$\eta = \frac{P_{out}}{P_{in}} = \begin{cases} -\dfrac{\hat{p}_2 Q_2 \left(1 + \frac{\hat{p}_2}{2\overline{K}_s}\right)}{\hat{p}_1 Q_1 \left(1 + \frac{\hat{p}_1}{2\overline{K}_s}\right)}, & \text{when } Q_2 \leq 0 \\ -\dfrac{\hat{p}_1 Q_1 \left(1 + \frac{\hat{p}_1}{2\overline{K}_s}\right)}{\hat{p}_2 Q_2 \left(1 + \frac{\hat{p}_2}{2\overline{K}_s}\right)}, & \text{when } Q_2 > 0 \end{cases} \tag{6}$$

This definition implies that, in order to calculate the efficiency of a transformer, we need to measure the flow rate at ports 1 and 2, and the pressure at ports 0, 1, and 2.

2.4. Previous efficiency definitions

There are different definitions of the efficiency of hydraulic transformers to be found in literature [9,10,16,17]. Apart from the fact that these studies do not include the compressibility of the oil, there are some other notable differences.

In these studies, the transformer is isolated from the rest of the CPR-system. The absolute pressure level (p_i) is used instead of the pressure level with respect to the LPR-pressure ($\hat{p}_i$). In doing so, the flow rate at port 0 is considered to have usable hydraulic power as well and needs to be included in the definitions for power loss and efficiency. In cases where $p_2 < p_1$, we find a positive flow rate into port 0 ($Q_0 > 0$) so this is considered input power. In cases where $p_2 > p_1$, the direction of the flow rate at port 0 can change ($Q_0 < 0$). In these cases, the hydraulic power at port 0 is considered an output power.

In [17], a second definition is given where the author considers the transformer to be part of a hydrostatic transmission. The derived efficiency definition is therefore similar to (6). This approach is based on the idea that the transformer cannot extract any more energy from the oil once it is at pressure level p_0, so this is the minimum energy state of the CPR-system.

3. DIVISION OF LOSSES

To better understand where the power loss from (5) is coming from, the combined pump/motor interpretation shown in **Figure 2b** is used again. In this section we assume that the transformer is delivering power to a load ($Q_1 > 0$, $Q_2 < 0$, $\omega > 0$). The machine that is connected to p_1 therefore acts as a motor, and the machine connected to p_2 acts as a pump. The derivations in this section are made under the assumption of a constant bulk modulus and a linear compression ratio. In other words, areas in pV-diagrams are approximated by simple triangles and rectangles.

3.1. Theoretical single piston displacement

Figure 3 shows the ideal pV-diagram of a single piston in both of the machines, including compressibility effects as also derived in [12–15]. In these figures, ΔV_i describes the volume that a single piston displaces from port i.

$$\Delta V_1 = V_{1,max}\left(1 - \frac{\hat{p}_1}{\overline{K}_s}\right) - V_{1,min} \tag{7}$$

$$\Delta V_2 = V_{2,min} - V_{2,max}\left(1 - \frac{\hat{p}_2}{\overline{K}_s}\right) \tag{8}$$

where $V_{i,min}$ and $V_{i,max}$ are the minimum and maximum geometric piston volume (for the current displacement) of machine i. Please note that $\Delta V_2 < 0$ in order to agree with the sign convention that was introduced in the previous section.

3.2. Flow rate loss

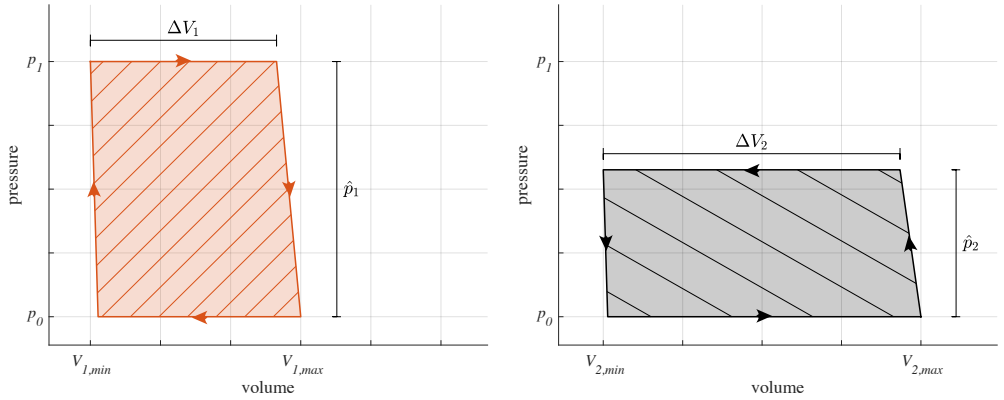
The displacement volume per piston per revolution, as described in (7) and (8), can be used to derive a theoretical flow rate $Q_{i,th}$.

$$Q_{i,th} = \frac{\omega z_i}{2\pi}\Delta V_i = \omega D_i \tag{9}$$

where z_i is the number of pistons and D_i is the displacement of the full machine at port i. When measuring the flow rate for an actual transformer, the flow rate at port i will not be ideal.

$$Q_i = Q_{i,th} + Q_{i,loss} = \omega D_i + Q_{i,loss}, \qquad \text{with } Q_{i,loss} \geq 0 \tag{10}$$

where $Q_{i,loss}$ is the difference between the measured flow rate Q_i and the ideal flow rate. If there is a lot of leakage at the motor, the flow rate over the motor is much higher than the theoretical flow rate at the current rotor speed. The hydraulic power in this leaked oil is not used to propel the pump, and is thus lost. Similarly, if there is a lot of leakage at the pump, there will be less fluid leaving the



(a) ideal motor (b) ideal pump

Figure 3: Ideal pV-diagram of a single piston in case of a motor (a) and a pump (b). The arrows show the direction through which the diagram is followed for a positive rotational speed. N.B. The bulk modulus has been decreased significantly to show the effect of compressibility.

transformer than the theoretical flow rate. Some of the torque that was used to propel the pump has been converted to hydraulic power that then has leaked back to reference pressure p_0 instead of leaving the transformer to move a work function. This is also a loss.

3.3. Torque loss

The area that the pV-diagram of a piston encircles can be used to calculate the torque T_i of machine i. From **Figure 3**, the torque is found to be:

$$T_i = \frac{z_i}{2\pi}\int pdV = \frac{z_i}{2\pi}\hat{p}_i \Delta V_i \left(1 + \frac{\hat{p}_i}{2\bar{K}_s}\right) = \hat{p}_i D_i \left(1 + \frac{\hat{p}_i}{2\bar{K}_s}\right) \tag{11}$$

When a transformer is in steady state, i.e. the transformer is not accelerating, the total torque on the shaft will be 0. Assuming that the transformer has some internal torque loss, we get the following equilibrium.

$$\hat{p}_1 D_1 \left(1 + \frac{\hat{p}_1}{2\bar{K}_s}\right) + \hat{p}_2 D_2 \left(1 + \frac{\hat{p}_2}{2\bar{K}_s}\right) = T_{loss} \tag{12}$$

For an ideal transformer ($T_{oss} = 0$), the motor torque is equal to the additive inverse of the pump torque. When the motor and pump are set to a certain displacement, and the CPR-system is at a certain pressure difference $\hat{p}_1$, the only variable in (12) is the output pressure $\hat{p}_2$. This means that in order to balance the torque on the rotor, the output pressure changes.

3.4. Comparing losses

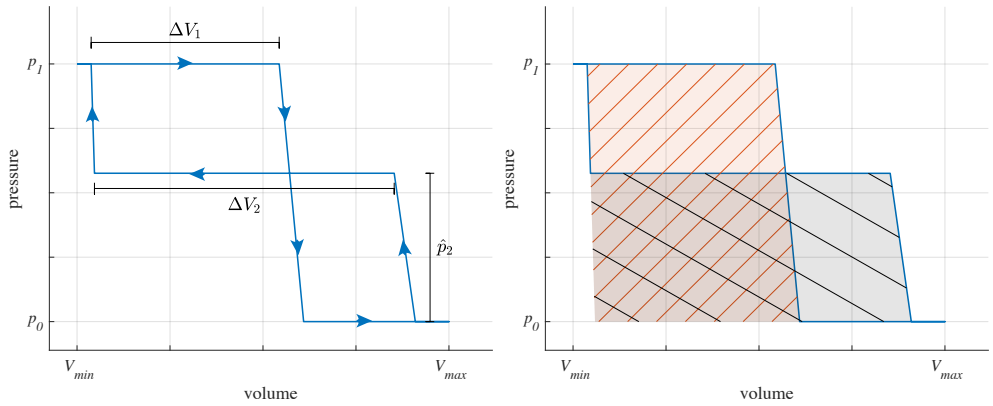
Since the flow rate loss and the torque loss have different units, it is difficult to compare the size of the two. Substitution of (10) in (5) results in (13).

$$\begin{aligned} P_{loss} &= \hat{p}_1(\omega D_1 + Q_{1,loss})\left(1 + \frac{\hat{p}_1}{2\bar{K}_s}\right) + \hat{p}_2(\omega D_2 + Q_{2,loss})\left(1 + \frac{\hat{p}_2}{2\bar{K}_s}\right) \\ &= T_{loss}\omega + Q_{1,loss}\hat{p}_1\left(1 + \frac{\hat{p}_1}{2\bar{K}_s}\right) + Q_{2,loss}\hat{p}_2\left(1 + \frac{\hat{p}_2}{2\bar{K}_s}\right) \\ &= P_{loss,T} + P_{loss,Q_1} + P_{loss,Q_2} \end{aligned} \tag{13}$$

Where $P_{loss,T}$, P_{loss,Q_1},and $P_{loss,Q_2}$are the power losses associated with the torque loss, and flow rate losses from port 1 and 2.

3.5. Innas Hydraulic Transformer principle

The Innas Hydraulic Transformer (IHT) principle is different from the example above in that it provides an integrated combination of the motor and pump from **Figure 3** into a single rotation group [8]. **Figure 4** shows an example of the pV-diagram of a single piston in such a machine. Starting in the top left corner of **Figure 4a**, the piston is connected to p_1 and the chamber volume is minimal, V_{min}. Following the direction of the arrow, the chamber volume starts expanding, drawing in oil from the HPR. At some point, the valve plate disconnects the piston from p_1 and connects to p_0, after which the piston expands further to its maximum volume V_{max}. At this point, the direction changes, and the chamber volume starts to decrease while the piston is still connected to p_0. Shortly thereafter, the valve plate disconnects the piston from p_0 and connects to p_2, where the oil is displaced to the



(a) ideal IHT (b) area comparison

Figure 4: Ideal pV-diagram of a single piston in a transformer that uses the IHT principle.

work function. Finally, near the end of the discharge stroke, the piston is connected to p_1 again, after which the piston volume decreases further to its minimum volume again.

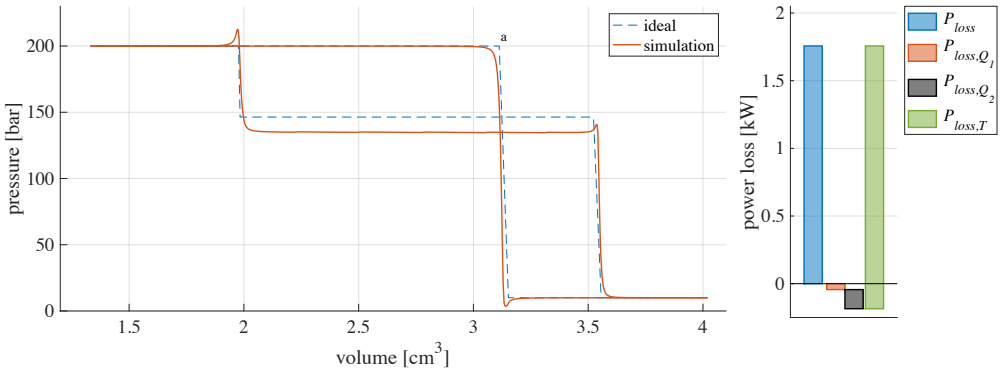
While the motor and pump action in an IHT do not occur in different machines, we can still define the displaced volume on ports 1 and 2, as is shown in **Figure 4a**. Therefore, equations (7) to (13) still hold. **Figure 4b** shows that the total area of the IHT pV-diagram can be divided into an equal sized motoring and pumping part, similar to the diagrams shown in **Figure 3**.

3.6. The problem with measuring displacement volume

Equations (10) and (12) define a flow rate loss and a torque loss, based on the theoretical displacement volume at ports 1 and 2 of the transformer. For pumps and motors, there is a well-known method for determining the displacement volume [18]. This method relies on measuring the flow rate at the high-pressure port of the unit at different operating speeds and different pressure levels. These measured flow rates are then extrapolated such that a displacement volume is derived for 0 rpm and 0 bar operating conditions. This method cannot be used for hydraulic transformers that are based on the IHT principle. The main reason for this, is that the pressure at port 2 cannot be chosen arbitrarily, since it is coupled to the ratio of ΔV_1 and ΔV_2. In other words, the displacement volume is different for each pressure level that is measured.

Furthermore, the pV-diagrams shown in **Figure 3** and **Figure 4** describe the ideal case, with perfect expansion and compression of the oil. **Figure 5a** shows the pV-diagram of a single piston, from a simulation model of the prototype transformer that is discussed in the next section. In this model, the transformer has some rotor friction, but has no leakage.

From **Figure 5a**, it is less clear how the losses are related to the geometric values of ΔV_1 and ΔV_2. For example, at corner a in **Figure 5a**, the pressure in the chamber starts to decrease before the chamber is at the ideal volume, since the flow area between port 1 and the piston chamber has to close at this point. Therefore, this chamber has effectively displaced less fluid from port 1. If the geometric value of ΔV_1 is used in (9) and (10), we find that the flow rate at port 1 is less than the ideal flow rate. This would result in a negative flow loss at port 1 ($Q_{1,loss} < 0$) for this machine. Something similar occurs for the displacement at port 2.



(a) pV-diagram (b) power loss

Figure 5: Ideal versus simulated pV-diagram for a piston in the FCT80 in (a) and division of power loss in (b), simulated at control angle 25°, and roughly 1500 rpm.

Figure 5b shows an example of the derived division of power loss for this simulated situation. The negative flow losses are countered by an exaggeration of the torque losses. Since the effective displacement at the ports is found to be less than the geometrical displacement volume, we calculate a higher torque loss in (12). The result is a torque loss that is larger than the overall power loss.

3.7. Conclusion

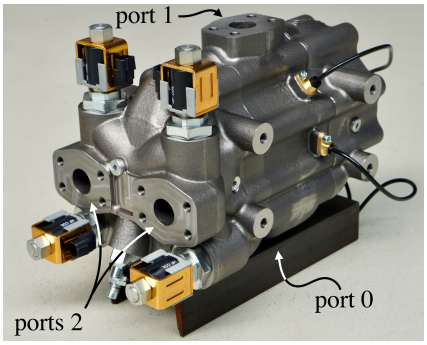
The theoretical division of losses for a hydraulic transformer can be made using pV-diagrams of ideal machines. This can be very useful in understanding the basic operating principle of hydraulic power transformation. Unfortunately, this division relies heavily on a good understanding of the displacement volume of the transformer at ports 1 and 2 at all possible operating conditions. To the authors knowledge, there is no known method to measure this displacement volume at the different operating conditions of an IHT.

4. FCT80

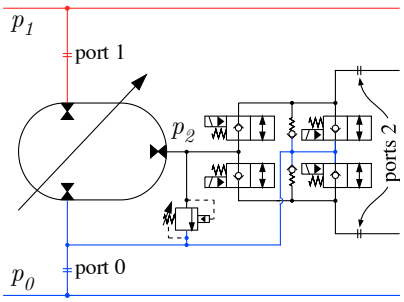
A new 80 cc/rev floating cup type hydraulic transformer (FCT80) has been designed by Innas [11]. **Figure 6a** shows a photo of the machined and assembled prototype. In the shown orientation, port 0 is not visible as it is on the bottom of the housing, and port 1 is found on the top. The front of the housing has two load ports 2. These two ports can for example be connected to either side of a hydraulic cylinder. There are four solenoid valves that can then be used to control which side of the cylinder is connected to p_2 and which side is connected to p_0, as is shown in hydraulic circuit in **Figure 6b**.

At high flow rates, the pressure difference over the control valves can become significant. To reduce the risk of cavitation at the port that is connected to p_0, there are two built-in check valves parallel to the solenoid valves. Furthermore, the circuit also shows an internal relief valve for safety reasons.

The FCT80 is controlled by an electric stepper motor, combined with a hydraulic servo motor [11]. In steady state conditions, this hydraulic servo motor will draw some high-pressure oil from the HPR, as it needs to deliver a certain continuous control torque while there is some internal leak flow. Since the motor is fed from within the transformer, this control power loss is automatically included in the overall efficiency measurements, in the form of a slightly higher flow rate at port 1. The power loss from the stepper motor will be negligible, as there is no continuous torque on the shaft of the stepper.



(a) photo of the FCT80 prototype



(b) hydraulic circuit of housing

Figure 6: The FCT80 prototype has two load ports, four control valves, two check valves, and a relief valve, all build into the housing.

5. TEST RESULTS

5.1. Test setup

The setup used to measure the performance of the FCT80 is shown in **Figure 7**. Several parameters are measured using the sensors that are listed in **Table 1**. The other components have been chosen such that the transformer can be operated in a steady state, i.e. constant pressure and flow rate on ports 0, 1, and 2.

The top left of **Figure 7** shows an electric motor that drives a 45 cc/rev pump. The output of this main pump is connected to port 1 of the FCT80. Between the pump and the transformer, there is a 5 L accumulator that is pre-charged at 157 bar. This part of the setup can be considered to be the HPR. The supply side of the pump is connected to port 0 of the FCT80. There is another 5 L accumulator between them, which is pre-charged to 4 bar. This part can be considered to be the LPR.

The speed of the main electric motor is controlled by the pressure measurement at port 1. The pressure in the LPR is maintained using a small charge pump in combination with a proportional valve. This is needed because for a different state of charge of the accumulators, there is a difference in total oil volume in the circuit. Additionally, the main pump has an external drain port (not shown in **Figure 7**) which also leaks some oil back to tank.

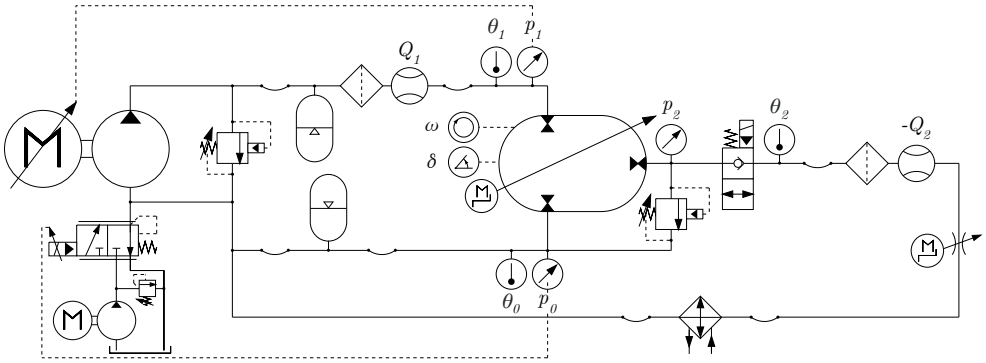


Figure 7: Hydraulic circuit used during the measurements.

Table 1: List of the sensors used in the transformer testbench.

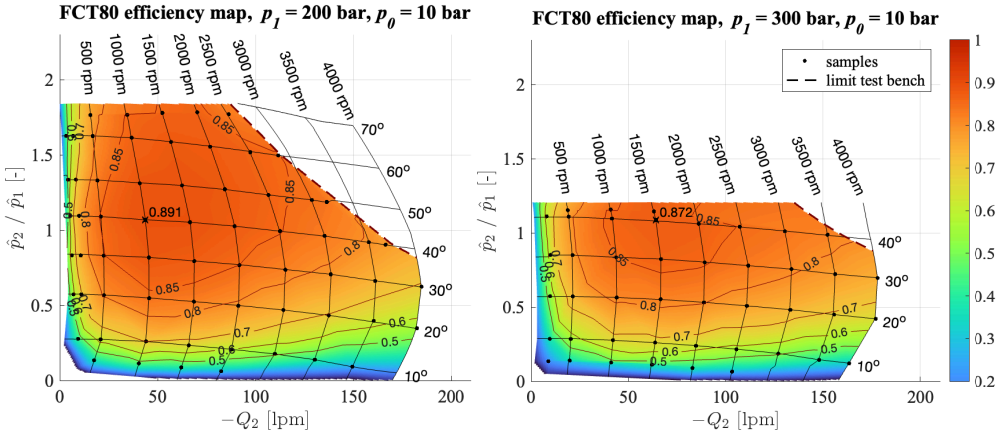
variable	symbol	sensor	range	accuracy
control angle	δ	Micro-Epsilon ES-U2	0.2 to 2.2 mm	0.04 µm
high pressure	p_1, p_2	Honeywell STJE 7500 psig	0 to 517.1 bar	±0.259 bar
low pressure	p_0	Omega PXM02MD0-040BARGV	0 to 40.0 bar	±0.02 bar
flow rate	Q_1, Q_2	VSE RS 400/32	1.0 to 400 l/min	0.5% MV*
speed	ω	Rheintacho SDN4 (in mod1)	0.1 Hz to 20 kHz	
temperature	$\theta_0, \theta_1, \theta_2$	Testo type 13 PT100 class B	-50 to 400°C	±0.3°C

*accuracy for this sensor is defined as a percentage of the measured value (MV)

As mentioned before, the control angle of the transformer is controlled by the electric stepper motor. The load flow from port 2 is throttled by means of a needle valve that is also controlled by a stepper motor. The flow from this load valve passes through a heat exchanger, before it is routed back to the rest of the LPR.

5.2. Steady-state efficiency map

In order to describe the steady state performance of the FCT80, it has been exposed to a grid of measurement points, or samples, at different operating conditions. First of all, the pressure in the HPR and LPR (p_1 and p_0) is kept constant for the whole grid. Each sample consists of a combination of a certain pressure p_2, controlled by the control angle, and flow rate Q_2, controlled by the load valve. Once these conditions are set, the procedure has been to wait until the oil temperature at port 1 is 50±1°C. This was achieved by controlling the cooling capacity of the heat exchanger. After some time, the temperature as well as all measured parameters are more or less stable, which means the transformer is operating in steady state. In this steady state, the sensor data is averaged over a period of 10 seconds. These average values per sample were used to calculate the efficiency maps shown in **Figure 8**. The black dots in these figures indicate the different sample points.



(a) $\hat{p}_1$ = 190 bar (b) $\hat{p}_1$ = 290 bar

Figure 8: Measured steady-state efficiency map of the FCT80 at p_0 = 10 bar, θ_1 = 50°C

In **Figure 8**, the output flow rate at port 2 ($-Q_2$) is shown on the x-axis, while the ratio between the output and input pressure ($\hat{p}_2/\hat{p}_1$) in shown on the y-axis. The maximum ratio of 1.84 and 1.21 correspond to a maximum output pressure $\hat{p}_2$ of 350 bar, which is the setting of the internal pressure relief valve. The somewhat horizontal lines are drawn at different control angles of the machine, while the vertical lines show different rotational speeds of the rotor. For this transformer, the maximum control angle is 70° and the maximum speed is 4000 rpm. The test bench is limited in terms of the maximum power that can be supplied to the transformer. The black dashed line shows the limit of the supply pump at the chosen HPR-pressure.

The face colour and other lines in **Figure 8** represent the measured steady state efficiency of the unit. The figures show that the efficiency of the transformer is very similar for the two different supply pressures. Overall, the efficiency is more than 0.7 for most of the center part of the field of operation. The maximum efficiency of 0.891 was measured at 40°, 1000 rpm, with a supply pressure of 200 bar. Looking at the edges of the efficiency map, we find that for low control angles and low rotor speeds, the efficiency decreases.

6. CONCLUSION

The introduction of CPR-systems and secondary control methods, have the potential to significantly reduce energy losses that are typically associated with current hydraulic systems. Hydraulic transformers can play a crucial role in the realisation of such systems. Since there are no commercially available hydraulic transformers, the FCT80 was developed as a new prototype transformer.

To describe the performance of the FCT80, a general definition for the power loss and efficiency of hydraulic transformers has been derived. The derived definitions describe the hydraulic input and output power of a transformer with respect to the minimum energy state of the CPR-system that the machine is part of. These definitions also take the compressibility of the oil into account.

For pumps and motors, it is common practice to divide the overall power loss into volumetric and mechanical losses based on the displacement volume of the tested unit. This approach was also used in an attempt to divide the overall transformer losses into volumetric and mechanical losses. However, since there is no known method to measure the actual displacement volume of an IHT for each possible control setting, the proposed equations for the division of losses are not applicable to the experimental measurements.

A specifically designed test setup was used to measure the steady state performance of the FCT80. The results are shown for two different supply pressures. For the measured samples, a peak efficiency of 0.891 was found. Furthermore, a large part of the field of operation was found to have an efficiency of 0.7 or higher, with decreasing efficiency for low control angles and low rotor speeds.

NOMENCLATURE

δ	Control angle	rad
η	Efficiency	
θ	Temperature	°C
ω	Rotational velocity	rad/s
$\overline{K}_s$	Average isentropic bulk modulus (1.76e9)	Pa
D	Displacement volume	m^3/rad
P	Power	W
p	Absolute pressure	Pa
$\hat{p}$	Pressure relative to p_0	Pa
Q	Flow rate	m^3/s
T	Torque	Nm
V	Volume	m^3
ΔV	Displacement volume per piston per rev.	m^3
z	Number of pistons	

subscripts

0	low pressure rail
1	high pressure rail
2	load pressure
i	port index (0,1,2)
in	input
$loss$	loss
max	maximum
min	minimum
out	output
Q	flow rate
T	torque
th	theoretical

REFERENCES

[1] Love, L. J., Lanke, E., and Alles, P. Estimating the impact (energy, emissions and economics) of the US fluid power industry. *Oak Ridge National Laboratory, Oak Ridge, TN,* 2012. doi:10.2172/1061537.

[2] Vael, G. E., Achten, P. A., and Fu, Z. The Innas Hydraulic Transformer: the key to the hydrostatic common pressure rail. Technical report, SAE Technical Paper, 2000. doi:10.4271/2000-01-2561.

[3] Linjama, M., Vihtanen, H.-P., Sipola, A., and Vilenius, M. Secondary controlled multi-chamber hydraulic cylinder. In *The 11th Scandinavian International Conference on Fluid Power, SICFP'09*, 2009.

[4] Achten, P., Vael, G., and Heybroek, K. Efficient hydraulic pumps, motors and transformers for hydraulic hybrid systems in mobile machinery. In 1. *VDI-Fachkonferenz Getriebe in Mobilen Arbeitsmaschinen, Friedrichshafen, DE, 7.-8. Jun, 2011*, pages 1-19. VDI-Wissens- forum, 2011. diva2:1051356.

[5] Heybroek, K., Vael, G. E. M., and Palmberg, J.-O. Towards resistance-free hydraulics in construction machinery. In *8th International Fluid Power Conference, Germany, Dresden,* volume 2, pages 123-138, 2012. diva2:1051345.

[6] Shen, W., Jiang, J., Su, X., and Karimi, H. R. Energy-saving analysis of hydraulic hybrid excavator based on common pressure rail. *The Scientific World Journal*, 2013. doi:10.1155/2013/560694.

[7] Siefert, J. and Li, P. Y. Optimal control and energy-saving analysis of common pressure rail architectures: HHEA and STEAM. In *BATH/ASME 2020 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers, 2020. doi:10.1115/FPMC2020-2799.

[8] Achten, P., Fu, Z., and Vael, G. Transforming futur hydraulics: a new design of a hydraulic transformer. In *The Fifth Scandinavion International Conference on Fluid Power, SICFP'97*, page 287, 1997.

[9] Li, X., Yuan, S., Hu, J., and Lv, J. Mathematical model for efficiency of the hydraulic transformer. In *2009 Asia-Pacific Power and Energy Engineering Conference*, pages 1-5. IEEE, 2009. doi:10.1109/APPEEC.2009.4918475.

[10] Zhou, J., Jing, C., and Wu, W. Energy efficiency modeling and validation of a novel swash plate-rotating type hydraulic transformer. *Energy*, 193, 2020. doi:10.1016/j.energy.2019.116652.

[11] Achten, P., Potma, J., Achten, S., Achten, J., and Mommers, R. The design of the FCT80 hydraulic transformer. In *The 18th Scandinavion International Conference on Fluid Power, SICFP'23*, 2023.

[12] Achten, P., Mommers, R., Nishiumi, T., Murrenhoff, H., Sepehri, N., Stelson, K., Palmberg, J.-O., and Schmitz, K. Measuring the losses of hydrostatic pumps and motors; a critical review of

ISO4409:2007. In *BATH/ASME 2019 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers, 2019. doi:10.1115/FPMC2019-1615.

[13] Williamson, C. and Manring, N. A more accurate definition of mechanical and volumetric efficiencies for Digital Displacement® pumps. In *BATH/ASME 2019 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers, 2019. doi:10.1115/FPMC2019-1668.

[14] Li, P. Y. and Barkei, J. Hydraulic effort and the efficiencies of pumps and motors with compressible fluid. In *BATH/ASME 2020 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers, 2020. doi:10.1115/FPMC2020-2801.

[15] Schänzle, C. and Pelz, P. F. Meaningful and physically consistent efficiency definition for positive displacement pumps - continuation of the critical review of ISO4391 and ISO4409. In *BATH/ASME 2021 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers, 2021. doi:10.1115/FPMC2021-68739.

[16] Werndin, R., Achten, P., Sannelius, M., and Palmberg, J.-O. Efficiency performance and control aspects of a hydraulic transformer. In *The 6th Scandinavion International Conference on Fluid Power, SICFP'99*, pages 395-407, 1999.

[17] Werndin, R. Efficiency performance and control aspects of a hydraulic transformer. Linköping University, Dept. of Mechanical Engineering, 1999. Master Thesis.

[18] Toet, G., Johnson, J., Montague, J., Torres, K., and Garcia-Bravo, J. The determination of the theoretical stroke volume of hydrostatic positive displacement pumps and motors from volumetric measurements. *Energies*, 12(3):415, 2019. doi:10.3390/en12030415.

Chapter 2

Fundamentals

REMAINING USEFUL LIFE ESTIMATION FOR RUBBER O-RING UNDER STORAGE CONDITIONS CONSIDERING DEPENDENT PERFORMANCE INDICATORS

peer reviewed

Rentong Chen[1], Shaoping Wang[1], Chao Zhang[1,2*], Boyu Shen[1], Zhouhe Xie[1]

[1]*School of Automation Science and Electrical Engineering, Beihang University, Beijing, 100191, P.R. China*
[2]*Research Institute of Frontier Science, Beihang University, Beijing, 100191, P.R. China*

* Corresponding author: Tel.: +86 10 82338365; E-mail address: cz@buaa.edu.cn

ABSTRACT

Rubber O-ring seals have been extensively used in various types of hydraulic actuators. If rubber O-ring seals are exposed to a heat environment during storage, it will result in material aging and mechanical properties change. Therefore, it is necessary to obtain an accurate reliability model for rubber O-ring under storage conditions. This paper develops a bivariate-dependent remaining useful life prediction model based on inverse Gaussian (IG) process and Frank Copula. The nonlinear explainable IG process considering unit-to-unit variability is utilized to describe the degradation process of two performance indicators, namely as compression set and compressive stress relaxation. The time scale function in IG process model is determined by material aging model. The Frank Copula is employed to capture the dependent relationship between these indicators. The two-stage parameter estimation method is adopted to estimate parameters in IG process and copula, separately. Bayesian and Expectation-Maximum algorithm are jointly utilized to update parameters in IG process degradation model under exponential family distribution framework. Maximum likelihood estimation is used to update parameter in Frank Copula. To validate the proposed method, aging degradation tests for O-ring seals are conducted. The results demonstrate that the proposed real-time parameter updating method also improves the accuracy of online RUL assessments.

Keywords: Remaining useful life prediction, Rubber O-Ring, Nonlinear inverse Gaussian process, Copula.

1. INSTRUCTIONS

Rubber O-rings play a significant role in preventing the leakage and facilitating the efficient transmission of fluid power into mechanical power. There are a number of stored seals in factories for immediate replacement. However, rubber O-rings are susceptible to various forms of degradation, including chemical reactions or temperature changing, which will lead to premature failure. Therefore, assessing the reliability of rubber O-rings under storage conditions is a critical aspect of ensuring the safe and effective operation of various complex hydraulic systems.

Mahankar et al. [1] provided a comprehensive review on mechanisms of hydraulic seal failure and the effect of medium to high temperatures on it. Morrell et al. [2] conducted the accelerated aging test for rubber O-ring, and it indicates that the predominant reaction contributing to compression set in these studies is oxidative degradation. Previous studies mainly focused on material mechanical property change under storage conditions and constructed reliability model based on ageing mechanism. Kömmling et al. [3] investigated the influence of compression during aging by conducting five years aging test. However, physics of failure model cannot reflect dynamic characteristics during degradation. In addition, it cannot incorporate multiple uncertainties, such as

time invariant and unit-to-unit variability [4].

In response to these limitations, data-driven methods have gained prominence. Stochastic process models have a great potential for capturing stochastic dynamic and they have been developed to conduct reliability assessment [5]. Sun et al. [6] used Gamma process model to describe the ageing process of rubber O-Ring. Arrhenius model is used to represent the shape parameter of Gamma degradation model. Sun et al. [7] used the Wiener process model with time scale transformation to describe the ageing degradation process. However, the degradation for mechanical components can be reflected by various dependent performance indicators Their dependence relationship may arise from shared failure modes or working operations. Wahab et al. [8] supported that there are two dependent performance indicators, namely as compression set and compressive stress relaxation, can reflect rubber O-Ring ageing degradation process. These two performance indicators are caused by the thermo-oxidative aging of the polymer. Therefore, it is necessary to consider the interdependence of the two PCs. Pan et al. [9] used bivariate time-varying copula to describe two dependent performance indicators of rubber O-Ring. However, to the best of our knowledge, few studies constructed a bivariate reliability model considering multiple uncertainties. Remaining useful life (RUL) model considering two dependent performance indicators are seldom developed. In addition, most of studies assume that the time scale function in stochastic process model to be linear or specific nonlinear function without physics of failure explanations [10].

The main contribution of this paper is constructing a bivariate reliability model based on nonlinear explainable IG process considering unit-to-unit variability and Frank copula. RUL prediction is conducted by updating the parameters in the reliability model utilizing Bayesian and Expectation-Maximum (EM) algorithm together. The rest of this paper is organized as follows. Section 2 develops bivariate reliability model based on nonlinear explainable IG process and Frank Copula. Section 3 shows and on-line RUL evaluation with new degradation data. Section 4 uses rubber O-ring ageing degradation data to show the accuracy and effectiveness of the proposed RUL prediction model. The whole paper concludes with Section 5.

2. MODEL DEVELOPMENT

2.1. Inverse Gaussian process model considering unit-to-unit variability

IG process is widely used to model monotonic degradation for mechanical component. IG process model is utilized to construct degradation model for performance indicators. The IG process $\{X(t), t \geq 0\}$ has the following properties:

- The IG process model with initial degradation value $X(0) \equiv 0$;
- The IG process model has independent degradation increments;
- The degradation increments for IG process model, $\Delta X(t) = X(t+\Delta t) - X(t)$ follows IG distribution:

$$\Delta X(t) \sim IG\left(\eta \Delta\Lambda(t;\boldsymbol{\gamma}), \lambda \Delta\Lambda^2(t;\boldsymbol{\gamma})\right) \tag{1}$$

where $\eta, \lambda > 0$. η denotes the degradation rate. $\Lambda(t)$ denotes the time scale function denoting the degradation trajectory. In this study, $\Lambda(t;\boldsymbol{\gamma})$ is determined by physical law function according to the failure mechanism, $\boldsymbol{\gamma}$ is the corresponding parameters in time scale function. IG distribution is one of the typical exponential family of distributions, whose standard PDF equation can be expressed as

$$p(x \mid \beta) = h(x) \exp\left(\beta^T \cdot T(x) - A(\beta)\right) \tag{2}$$

where x is the random variable, β is parameters in exponential family distribution, $T(x)$ denotes the sufficient statistic, $h(x)$ denotes the base measure, $A(\beta)$ is log partition function. Let $x_t = X(t)$, its PDF under the framework of exponential family distribution is expressed as:

$$f(x_t) = \frac{1}{\sqrt{2\pi x_t^3}} \exp\left\{ \begin{bmatrix} -\frac{\lambda}{2\eta^2} & \frac{\lambda\Lambda^2(t;\gamma)}{2} \end{bmatrix} \begin{bmatrix} x_t \\ \frac{1}{x_t} \end{bmatrix} - \left[-\frac{\lambda\Lambda(t)}{\eta} - \frac{1}{2}\log \lambda\Lambda^2(t;\gamma) \right] \right\} \tag{3}$$

where $\beta = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix}^T = \begin{bmatrix} -\frac{\lambda}{2\eta^2} & \frac{\lambda\Lambda^2(t;\gamma)}{2} \end{bmatrix}^T$, and $A_{IG}(\beta)$ can be expressed as $A_{IG}(\beta) = 2\sqrt{\beta_1\beta_2} - \frac{1}{2}\log(-2\beta_2)$. According to the properties of exponential family distribution, the expectation and variance of $X(t)$ can be obtained by

$$\mathrm{E}[X(t)] = \frac{\partial}{\partial\beta_1} A_{IG}(\beta) = \eta\Lambda(t;\gamma) \tag{4}$$

$$\mathrm{Var}[X(t)] = \frac{\partial^2}{\partial\beta_1} A_{IG}(\beta) = \frac{\eta^3}{\lambda}\Lambda(t;\gamma) \tag{5}$$

The failure time T represents the point in time at which the degradation path first reaches the predefined threshold d. The PDF and CDF of failure time T can be found in Eq. (6) and Eq.(7), respectively.

$$F_T(t) = \Phi\left(\sqrt{\frac{\lambda}{d}} \left(\Lambda(t;\gamma) - \frac{d}{\eta} \right) \right) - \exp\left(\frac{2\lambda}{\eta}\Lambda(t;\gamma) \right) \Phi\left(-\sqrt{\frac{\lambda}{d}} \left(\Lambda(t;\gamma) + \frac{d}{\eta} \right) \right) \tag{6}$$

$$\begin{aligned} f_T(t) = {} & \sqrt{\frac{\lambda}{d}} \Phi\left(\sqrt{\frac{\lambda}{d}} \left(\Lambda(t;\gamma) - \frac{d}{\eta} \right) \right) - \frac{2\lambda}{\eta} \exp\left(\frac{2\lambda}{\eta}\Lambda(t;\gamma) \right) \Phi\left(-\sqrt{\frac{\lambda}{d}} \left(\Lambda(t;\gamma) + \frac{d}{\eta} \right) \right) \\ & + \sqrt{\frac{\lambda}{d}} \exp\left(\frac{2\lambda}{\eta}\Lambda(t;\gamma) \right) \Phi\left(-\sqrt{\frac{\lambda}{d}} \left(\Lambda(t;\gamma) + \frac{d}{\eta} \right) \right) \end{aligned} \tag{7}$$

From the meaning of failure life time, the RUL L_j can be seen as the failure time of IG process $\{X'(l_j), l_j > 0\}$ with corresponding threshold $d - X(t_j)$, where $X'(l_j) = X(l_j + t_j) - X(t_j)$. Similarly, the corresponding CDF and PDF for RUL L_j can be also obtained.

Considering the design errors, processing techniques and different working conditions, unit-to-unit variability can be represented in degradation paths. It is more suitable to include unit-to-unit variability in the reliability model. The expectation of degradation data is used to represent unit-to-unit variability in the IG process model. We can find that the degradation rate η is showed as the form of $\frac{1}{\eta}$ or $\frac{1}{\eta^2}$ in Eq.(3). For the convenience of following modeling and computation. We assume that $\frac{1}{\eta}$ follows normal distribution as $\frac{1}{\eta} \sim N\left(\mu_\eta, \frac{1}{\sigma_\eta^2} \right)$, and there will be

$$g_\eta\left(\eta \mid \mu_\eta, \frac{1}{\sigma_\eta^2}\right)=\frac{\sigma_\eta}{\sqrt{2\pi}\cdot\eta^2}\exp\left\{-\frac{1}{2}\sigma_\eta^2\left(\frac{1}{\eta}-\mu_\eta\right)^2\right\} \tag{8}$$

The, the CDF and PDF for failure lifetime T can be expressed as Eq. (9) and Eq.(10), respectively.

$$\begin{aligned} F_T(t)=&\Phi\left(\sqrt{\frac{\lambda}{d}}\cdot\frac{\sigma_\eta\Lambda(t;\boldsymbol{\gamma})-\mu_\eta\sigma_\eta d}{\sqrt{\sigma_\eta^2+\lambda d}}\right)\\ &-\exp\left(2\lambda\mu_\eta\Lambda(t;\boldsymbol{\gamma})+\frac{2\lambda^2\Lambda^2(t;\boldsymbol{\gamma})}{\sigma_\eta^2}\right)\Phi\left(-\sqrt{\frac{\lambda}{d}}\cdot\frac{\left(\sigma_\eta^2+2\lambda d\right)\Lambda(t;\boldsymbol{\gamma})+\mu_\eta\sigma_\eta^2 d}{\sqrt{\sigma_\eta^4+\lambda d\sigma_\eta^2}}\right) \end{aligned} \tag{9}$$

$$\begin{aligned} f_T(t)=&\sqrt{\frac{\lambda}{d}}\frac{\sigma_\eta}{\sqrt{\sigma_\eta^2+\lambda d}}\Phi\left(\sqrt{\frac{\lambda}{d}}\cdot\frac{\sigma_\eta\Lambda(t;\boldsymbol{\gamma})-\mu_\eta\sigma_\eta d}{\sqrt{\sigma_\eta^2+\lambda d}}\right)\\ &-\exp\left(2\lambda\mu_\eta\Lambda(t)+\frac{2\lambda^2\Lambda^2(t;\boldsymbol{\gamma})}{\sigma_\eta^2}\right)\exp\left(2\lambda\mu_\eta+\frac{4\lambda^2\Lambda(t;\boldsymbol{\gamma})}{\sigma_\eta^2}\right)\Phi\left(-\sqrt{\frac{\lambda}{d}}\cdot\frac{\left(\sigma_\eta^2+2\lambda d\right)\Lambda(t;\boldsymbol{\gamma})+\mu_\eta\sigma_\eta^2 d}{\sqrt{\sigma_\eta^4+\lambda d\sigma_\eta^2}}\right)\\ &+\sqrt{\frac{\lambda}{d}}\cdot\frac{\sigma_\eta^2+2\lambda d}{\sqrt{\sigma_\eta^4+\lambda d\sigma_\eta^2}}\cdot\exp\left(2\lambda\mu_\eta\Lambda(t;\boldsymbol{\gamma})+\frac{2\lambda^2t^2}{\sigma_\eta^2}\right)\cdot\Phi\left(-\sqrt{\frac{\lambda}{d}}\cdot\frac{\left(\sigma_\eta^2+2\lambda d\right)\Lambda(t;\boldsymbol{\gamma})+\mu_\eta\sigma_\eta^2 d}{\sqrt{\sigma_\eta^4+\lambda d\sigma_\eta^2}}\right) \end{aligned} \tag{10}$$

2.2. Bivariate Reliability Model

The copula function is an effective tool in reliability engineering for describing the dependence between two degradation performance indicators. Frank copula, one of the Archimedean copulas, has symmetric structure. Frank copula can capture both positive and negative correlations between variables, so it has been widely used in bivariate reliability analysis The CDF and PDF of Frank Copula can be found in Chen et al. [11]. Considering two dependent performance indicators $X_1(t)$ and $X_2(t)$ with threshold d_1 and d_2, the reliability function $R(t)$ and failure lifetime T can be given by:

$$R(t)=\Pr\left\{X_1(t)<d_1, X_2(t)<d_2\right\} \tag{11}$$

$$T=\min\left\{T_1,T_2\right\} \tag{12}$$

According to Chen et al. [11], when the unit-to-unit variability is not taken into account, Eq. (11) can be furthered calculated by:

$$R(t)=C\left(R_1(t),R_2(t)\right) \tag{13}$$

In addition, RUL L_j at degradation time t_j is obtained by:

$$F_{L_j}\left(l_j\right)=C\left(F_{L_j,1}\left(l_j\right),F_{L_j,2}\left(l_j\right)\right) \tag{14}$$

Considering the unit-to-unit variability, the reliability and RUL function can be obtained by:

$$R^v(t)=\iint R\left(t\mid\eta_1,\eta_2\right)g_\eta\left(\eta_1\right)g_\eta\left(\eta_2\right)\mathrm{d}\eta_1\mathrm{d}\eta_2 \tag{15}$$

$$F_{L_j}^v\left(l_j\right)=\iint F_{L_j}\left(l_j\mid\eta_1,\eta_2\right)g_\eta\left(\eta_1\right)g_\eta\left(\eta_2\right)\mathrm{d}\eta_1\mathrm{d}\eta_2 \tag{16}$$

3. STATISTICAL INFERENCE

We assume that N samples with $M_i\left(i=1,2,\cdots,N\right)$ measurement times are conducted in the degradation test. Let $X_k\left(t_{ij}\right)$ denote the degradation data for k th performance indicator of the i th test sample at j th measurement $\left(k=1,2;i=1,2,\cdots,N;j=1,2,\cdots,M_i\right)$. Assuming that $\boldsymbol{\Theta}$ denotes the unknown parameters in the proposed model, and it will be $\boldsymbol{\Theta}=\left(\mu_{\eta 1},\sigma_{\eta 1}^2,\lambda_1,\boldsymbol{\gamma}_1,\mu_{\eta 2},\sigma_{\eta 2}^2,\lambda_2,\boldsymbol{\gamma}_2,\theta\right)$. Let $\boldsymbol{\Theta}_k^{\mathrm{IG}}$ denote the unknown parameters in IG process, that is $\boldsymbol{\Theta}_k^{\mathrm{IG}}=\left(\mu_{\eta k},\sigma_{\eta k}^2,\lambda_k,\boldsymbol{\gamma}_k\right)$. Assuming that $\boldsymbol{D}=\left(\boldsymbol{D}_1,\boldsymbol{D}_2\right)$ denotes the degradation data for two dependent performance indicators. We will discuss the unknown parameter estimation methods both in initial value determination stage (off-line stage) and on-line stage.

3.1. Initial Value Determination

Initial values for unknown parameters $\boldsymbol{\Theta}=\left(\mu_{\eta 1},\sigma_{\eta 1}^2,\lambda_1,\gamma_1,\mu_{\eta 2},\sigma_{\eta 2}^2,\lambda_2,\gamma_2,\theta\right)$ are determined by the off-line degradation data. The reliability model based on bivariate dependent performance indicators is given by:

$$\begin{aligned}
&\Delta X_k\left(t_{ij}\right)=X_k\left(t_{i,j+1}\right)-X_k\left(t_{ij}\right)\\
&\Delta X_k\left(t_{ij}\right)\sim IG\left(\eta_k\Delta t_{ij},\lambda_k\left(\Delta\Delta t_{ij}\right)^2\right),\frac{1}{\eta_k}\sim N\left(\mu_{\eta k},\frac{1}{\sigma_{\eta k}^2}\right)\\
&f\left(\Delta X_1\left(t_{ij}\right),\Delta X_2\left(t_{ij}\right)\right)=\prod_{k=1}^{2}\prod_{i=1}^{N}\prod_{j=1}^{M_i}f_k\left(\Delta X_k\left(t_{ij}\right);\boldsymbol{\Theta}_k^{\mathrm{IG}}\right)\prod_{i=1}^{N}\prod_{j=1}^{M_i}c\left(F_1\left(\Delta X_k\left(t_{ij}\right)\right),F_2\left(\Delta X_k\left(t_{ij}\right)\right);\theta\right)\\
&\left(k=1,2;i=1,2,\cdots,N;j=1,2,\cdots,M_i\right)
\end{aligned}\tag{17}$$

Then, the log-likelihood function based on complete data $\left(\boldsymbol{D},\eta_1,\eta_2\right)$ is given by:

$$\begin{aligned}
\ln L\left(\boldsymbol{\Theta}\mid\boldsymbol{D},\eta_1,\eta_2\right)=&\sum_{i=1}^{N}\sum_{j=1}^{M_i}\ln c\left(F_1\left(\Delta X_k\left(t_{ij}\right)\mid\eta_{1,i}\right),F_2\left(\Delta X_k\left(t_{ij}\right)\mid\eta_{2,i}\right);\theta\right)\\
&+\sum_{k=1}^{2}\sum_{i=1}^{N}\sum_{j=1}^{M_i}\ln\left[f_k\left(\Delta X_k\left(t_{ij}\right)\mid\eta_{k,i}\right);\boldsymbol{\Theta}_k^{\mathrm{IG}}\right]+\sum_{k=1}^{2}\sum_{i=1}^{N}\ln g_\eta\left(\eta_{k,i}\right)
\end{aligned}\tag{18}$$

Bayesian MCMC is employed to estimate unknown parameters:

$$\pi\left(\boldsymbol{\Theta}\mid\boldsymbol{D}\right)\propto\ln L\left(\boldsymbol{\Theta}\mid\boldsymbol{D},\eta_1,\eta_2\right)\cdot\pi\left(\boldsymbol{\Theta}\right)\tag{19}$$

where $\pi\left(\boldsymbol{\Theta}\right)$ is the prior distribution, $\pi\left(\boldsymbol{\Theta}\mid\boldsymbol{D},\boldsymbol{T}\right)$ is the posterior distribution. In our study, non-informative distribution is used to estimate $\boldsymbol{\Theta}$. The estimation results are chosen as the initial value in the on-line parameter update process. Then, the reliability function given the posterior distribution of $\boldsymbol{\Theta}$ can be expressed as:

$$R^v\left(t\mid\boldsymbol{D}\right)=\int_{\boldsymbol{\Theta}}R^v\left(t\mid\boldsymbol{\Theta}\right)\pi\left(\boldsymbol{\Theta}\mid\boldsymbol{D}\right)\mathrm{d}\boldsymbol{\Theta}=\mathrm{E}_{\pi(\boldsymbol{\Theta}\mid\boldsymbol{D},\boldsymbol{T})}\left[R^v\left(t\mid\boldsymbol{\Theta}\right)\right]\tag{20}$$

3.2. On-line Parameter estimation

In this section, the on-line parameter and RUL update algorithm based on on-line degradation data is

discussed. Note that the time scale parameter $\boldsymbol{\gamma}$ reflecting the degradation trajectory, there are not large fluctuations for $\boldsymbol{\gamma}$ during the whole service time. Therefore, $\boldsymbol{\gamma}$ will not be updated at the on-line stage, and it will be beneficial to reduce computational burden. As for other parameters in $\boldsymbol{\Theta}$, two-stage estimation method, inference function for the margins (IFM) is utilized to estimate the unknown parameters in IG process and dependence parameter in Frank copula, separately [12]. The two-stage estimation method has significant computational advantages in terms of computational efficiency.

Stage 1. Unknown Parameter Update in IG process

Assuming that there is degradation data $\boldsymbol{D}_{k,1:m}=\left[\Delta X_k\left(t_1\right),\Delta X_k\left(t_2\right),\cdots,\Delta X_k\left(t_m\right)\right]\left(k=1,2\right)$ for both performance indicators at time t_m. Using the IG distribution under exponential family framework, we can get the likelihood function:

$$p\left(\boldsymbol{D}_{k,1:m};\beta\right)=\prod_{j=1}^{m}h\left(\Delta X_k\left(t_j\right)\right)\exp\left(\beta^T\cdot T\left(\Delta X_k\left(t_j\right)\right)-A\left(\beta\right)\right) \tag{21}$$

Based on the Bayesian theory, there is:

$$p\left(\frac{1}{\eta_k}\mid\boldsymbol{D}_{k,1:m}\right)\propto p\left(\boldsymbol{D}_{k,1:m}\mid\frac{1}{\eta_k}\right)\cdot p\left(\frac{1}{\eta_k}\right) \tag{22}$$

where $p\left(\frac{1}{\eta_k}\right)$ is the prior information for $\frac{1}{\eta_k}$. It follows the normal distribution, which can be represented under exponential family framework. $p\left(\frac{1}{\eta_k}\mid\boldsymbol{D}_{k,1:m}\right)$ is the posterior distribution given the real time degradation $\boldsymbol{D}_{k,1:m}$, and it also follows normal distribution, that means $p\left(\frac{1}{\eta_k}\mid\boldsymbol{D}_{k,1:m}\right)\sim N\left(\mu_{\eta k,m},\frac{1}{\sigma_{\eta k,m}^2}\right)$. Regardless of the terms without $\frac{1}{\eta_k}$, Eq. (22) can be furthered expressed as

$$p\left(\frac{1}{\eta_k}\mid\boldsymbol{D}_{k,1:m}\right)\propto\left\{\prod_{j=1}^{m}\exp\left\{-\frac{\lambda}{2\eta^2}\Delta X_k\left(t_j\right)+\frac{\lambda\cdot\Delta\Lambda\left(t_j;\hat{\boldsymbol{\gamma}}\right)}{\eta}\right\}\right\}\cdot\exp\left\{-\mu_{\eta k}\sigma_{\eta k}^2\frac{1}{\eta_k}-\frac{1}{2}\sigma_{\eta k}^2\frac{1}{\eta_k^2}\right\} \tag{23}$$

Then, we can obtain the mean and variance of posterior distribution of $\frac{1}{\eta_k}$:

$$\begin{cases}\mu_{\eta k,m}^{\text{posterior}}=\dfrac{\mu_{\eta k}\sigma_{\eta k}^2+\lambda\Lambda\left(t_m;\hat{\boldsymbol{\gamma}}\right)}{\sigma_{\eta k}^2+\lambda X_k\left(t_m;\boldsymbol{\gamma}\right)}\\ \left(\sigma_{\eta k,m}^2\right)^{\text{posterior}}=\sigma_{\eta k}^2+\lambda X_k\left(t_m;\hat{\boldsymbol{\gamma}}\right)\end{cases} \tag{24}$$

Since $\frac{1}{\eta_k}$ follows normal distribution, it cannot be observed directly. Expectation-Maximum (EM) algorithm can be used to estimate unobserved latent variables. The log-likelihood function for complete data can be expressed as:

$$\log p\left(\boldsymbol{D}_{k,1:m},\frac{1}{\eta_k}|\boldsymbol{\Theta}\right)=\log p\left(\boldsymbol{D}_{k,1:m}|\frac{1}{\eta_k},\boldsymbol{\Theta}\right)+\log p\left(\frac{1}{\eta_k}|\boldsymbol{\Theta}\right) \tag{25}$$

Regardless of the base measure, Eq. (25) can be furthered calculated as:

$$\log p\left(\boldsymbol{D}_{k,1:m},\frac{1}{\eta_k}|\boldsymbol{\Theta}\right)\Leftrightarrow\sum_{j=1}^{m}\left\{\beta^T T\left(\Delta X_k\left(t_j\right)\right)-A_{IG}\left(\beta\right)\right\}+\left\{\alpha^T T\left(\frac{1}{\eta_k}\right)-A_N\left(\alpha\right)\right\} \tag{26}$$

Given the degradation data $\boldsymbol{D}_{k,1:m}$, the estimated value $\hat{\boldsymbol{\theta}}_{k,m}^{IG(g)}=\left[\mu_{\eta k}^{(g)},\sigma_{\eta k}^{2(g)},\lambda_k^{(g)}\right]$ at g th iteration step in the EM algorithm, we can get the posterior value $\hat{\boldsymbol{\theta}}_{k,m}^{IG\text{posterior}(g)}$ through Eq. (22). As for the EM algorithm under exponential family framework, $T\left(\frac{1}{\eta_k}\right)$ is replaced by $\mathrm{E}_{p\left(\frac{1}{\eta_k}|\boldsymbol{D}_{k,1:m},\hat{\boldsymbol{\theta}}_{k,m}^{IG\text{posterior}(g)}\right)}\left[T\left(\frac{1}{\eta_k}\right)\right]$ to complete E-Step. Note that $\frac{1}{\eta_k}$ is also involved in β, which also needs to be replaced. Let $q\left(\frac{1}{\eta_k}\right)=p\left(\frac{1}{\eta_k}|\boldsymbol{D}_{k,1:m},\hat{\boldsymbol{\theta}}_{k,m}^{IG\text{posterior}(g)}\right)$, and the Q function in E-Step is given by:

$$\begin{aligned}Q\left(\boldsymbol{\Theta}_k^{IG}|\boldsymbol{D}_{k,1:m},\hat{\boldsymbol{\theta}}_{k,m}^{IG\text{posterior}(g)}\right)=\sum_{j=1}^{m}\Bigg\{&\left\{\mathrm{E}_{q\left(\frac{1}{\eta_k}\right)}\left[\beta\right]\right\}^T T\left(\Delta X_k\left(t_j\right)\right)-A_{IG}\left(\mathrm{E}_{q\left(\frac{1}{\eta_k}\right)}\left[\beta\right]\right)\Bigg\}\\&+\left\{\alpha^T\cdot\mathrm{E}_{q\left(\frac{1}{\eta_k}\right)}\left[T\left(\frac{1}{\eta_k}\right)\right]-A_N\left(\alpha\right)\right\}\end{aligned} \tag{27}$$

Let $\frac{\partial}{\partial\boldsymbol{\Theta}_k^{IG}}Q\left(\boldsymbol{\Theta}_k^{IG}|\boldsymbol{D}_{k,1:m},\hat{\boldsymbol{\theta}}_{k,m}^{IG\text{posterior}(g)}\right)=0$, then we can get:

$$\mu_{\eta k,m}^{(g+1)}=\mu_{\eta k,m}^{\text{posterior}(g)} \tag{28}$$

$$\sigma_{\eta k,m}^{2(g+1)}=\left(\sigma_{\eta k,m}^{2}\right)^{\text{posterior}(g)} \tag{29}$$

$$\lambda_{k,m}^{(g+1)}=\frac{m}{\sum_{j=1}^{m}\left\{\left(\mu_{\eta k,m}^{2(g+1)}+\frac{1}{\sigma_{\eta k,m}^{2(g+1)}}\right)\Delta X_k\left(t_j\right)-2\mu_{\eta k,m}^{2(g+1)}\Delta\Lambda\left(t_j;\hat{\boldsymbol{\gamma}}\right)+\frac{\left(\Delta\Lambda\left(t_j;\hat{\boldsymbol{\gamma}}\right)\right)^2}{\Delta X_k\left(t_j\right)}\right\}} \tag{30}$$

That completes Stage 1.

Stage 2. Unknown Parameter Update in Frank Copula

The second stage is to update the unknown parameter θ in Frank copula. The estimated value in stage 1 can be used to calculated marginal distribution, and then we can get the θ by maximum likelihood estimation (MLE):

$$\hat{\theta}_m=\arg\max_{\theta}\ln L^c\left(\theta|\boldsymbol{D}_{k,1:m}\right) \tag{31}$$

That completes Stage 2. Then, we can also update RUL when the real-time degradation data comes. Algorithm 1 shows the whole procedure of on-line parameter update.

Algorithm 1: On-line parameter estimation procedure

Input: $\mu_{\eta,0}, \sigma^2_{\eta,0}, \lambda_{k,0}, \theta_0$ and $\hat{\gamma}$ (Obtained from the off-line reliability analysis), On-line degradation data $\boldsymbol{D}_{k,1:m}$, iteration times L, and EM algorithm iteration threshold ε

Output: $\mu_{\eta k,m}$, $\sigma^2_{\eta k,m}$, $\lambda_{k,m}$ and θ_m

Step 1: Initialize: $\left\{\mu_{\eta k}^{(0)}, \sigma_{\eta k}^{2(0)}, \lambda_k^{(0)}\right\}$

For $g = 1:L$ or $\left|\hat{\boldsymbol{\theta}}_k^{IG(g+1)} - \hat{\boldsymbol{\theta}}_k^{IG(g)}\right| \geq \varepsilon$

  Calculate $\left\{\mu_{\eta k,m}^{\text{posterior}(g)}, \left(\sigma^2_{\eta k,m}\right)^{\text{posterior}(g)}\right\}$ based on (24);

  Calculate $\left\{\mu_{\eta k}^{(g+1)}, \sigma_{\eta k}^{2(g+1)}, \lambda_k^{(g+1)}\right\}$ based on (28), (29) and (30);

End for

Step 2: Calculate $\hat{\theta}_m$ based on (31)

Step 3: Update RUL based on (14).

When getting the new degradation date D_{m+1}, repeat Step 1 ~ Step 3

4. CASE STUDY: RUBBER O-RING STORAGE DEGRADATION TEST

In this section, we conduct the rubber O-Ring storage degradation test to illustrate the effectiveness and advantages of the proposed RUL prediction method.

4.1. Rubber-O Ring Storage Degradation Test

The rubber O-ring is widely used both in static applications and dynamic applications. However, during the long-term storage, large number of spare parts will age due to its storage conditions. It will result in the reducing their sealing mechanical properties, eventually losing their use value or affecting the reliability of the mechanical system. The degradation performance indicator, compression set and compressive stress relaxation, can both describe the degradation process during the storage stage. The compression set reflects the sign of elasticity and deformation resistance of rubber materials. The compressive stress relaxation represents the change value of elastic pressure before and after aging. The smaller the compression permanent deformation, the better the resilience of the material and the stronger the deformation resistance. Compression set tends to increase with aging time. Wahab et al. [8] supported that these two performance indicators are statistically dependent with each other. Therefore, it is reasonable to assume that both two performance indicators could be described by IG process, and their dependence could be described by Frank Copula.

The degradation test rig is shown in **Figure 1**. Compression set ε (Performance Indicator 1, PI1) and compressive stress relaxation σ_t/σ_0 (Performance Indicator 2, PI2), were tested under storage environment. The specific test methods can be found in Pan et al. [9]. There are ten test samples in the degradation test. The degradation paths after linear transformation for compression set $X_1(t) = -\ln(1-\varepsilon)$ and compressive stress relaxation $X_2(t) = -\ln(\sigma_t/\sigma_0)$ are shown in **Figure 2**. Based on the engineering experience, the failure thresholds of PI1 and PI2 are set to be 0.4 and 0.7, respectively. Thus, the thresholds after linear transformation are $d_1 = 0.511$ and $d_2 = 0.3567$ .The aging model commonly used to describe the seal materials can be expressed as:

$$P = A\exp\left(-Kt^{\gamma}\right) \tag{32}$$

where P denotes the material properties for seal, t denotes the service time, K denotes the coefficient related to aging, A and γ are related constants. To determine the specific expression of $\Lambda(t)$ in IG process, Eq. (32) also needs to be linearized. Then, as for materials aging, $\Lambda(t)$ should be expressed as:

$$\Lambda(t) = t^{\gamma} \tag{33}$$

According to the principle of hold-out method, the degradation data from the first 8 samples are utilized to determine the initial values, while the remaining 2 samples are utilized for verifying the proposed model.



Figure 1: The degradation test rig.

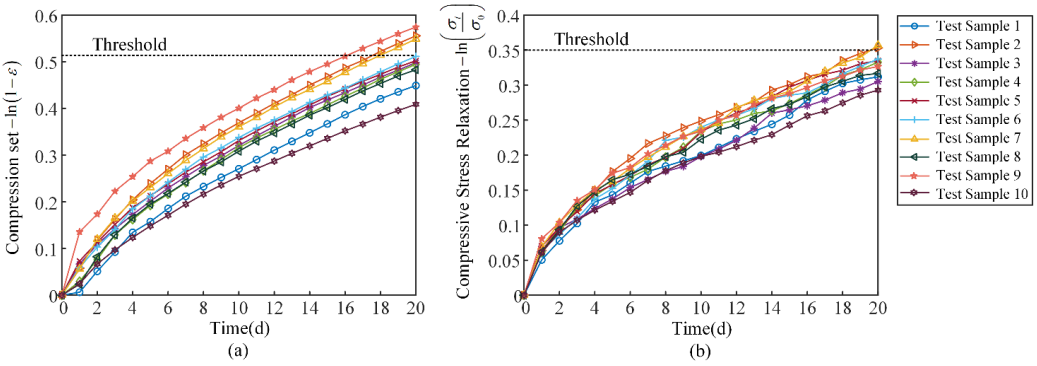


Figure 2: Linear transformation degradation paths (a) Performance Indicator 1, PI1: Compression set (b) Performance Indicator 2, PI2: Compressive stress relaxation

4.2. RUL prediction

The initial values are determined by Bayesian MCMC method proposed in Section 3.1. The estimation results are shown in **Table 1**. Note that the parameters in time scale functions γ_1 and γ_2 will not be updated at the online stage.

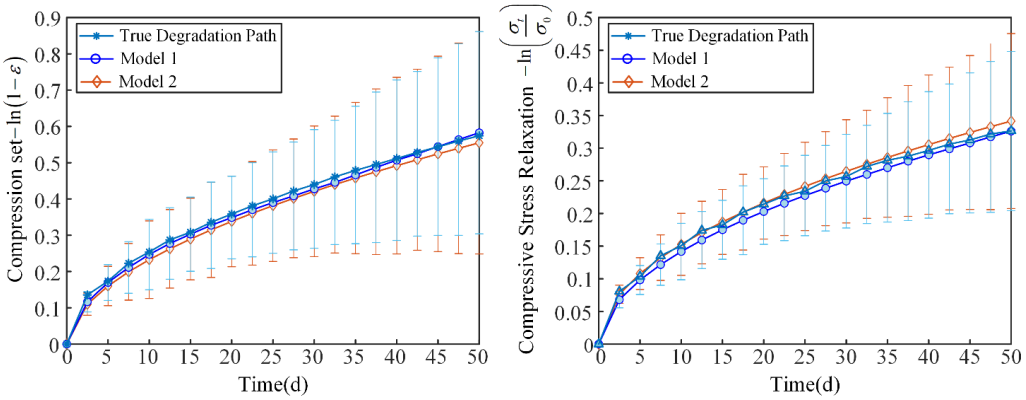
Table 1: Parameters Estimation Results

Unknow parameter	Estimation Results		
	Medium	2.5pc	97.5pc
$\mu_{\eta 1}$	0.9212	0.9008	0.9432
$\sigma_{\eta 1}^2$	79.83	77.62	81.93
λ_1	95.8	90.32	100.44
γ_1	0.5412	0.5311	0.5509
$\mu_{\eta 2}$	0.5312	0.5176	0.5542
$\sigma_{\eta 2}^2$	58.16	56.19	60.44
λ_2	149.6	145.63	154.37
γ_2	0.5321	0.5296	0.5398
θ	20.69	18.43	22.64

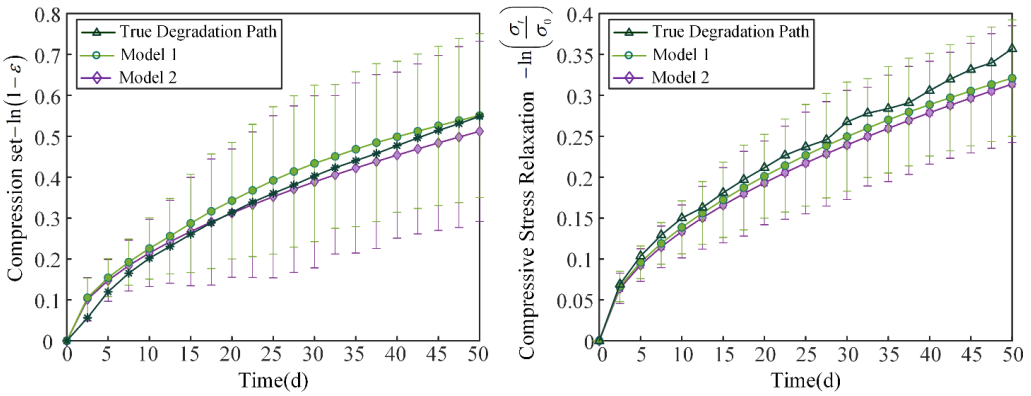
The 9th and 10th seal sample are used to conduct on-line parameters update and RUL prediction. The Bayesian-EM update algorithm is shown based on the method proposed in Section 3.2. We also compare two models in order to show the accuracy and effectiveness of the on-line reliability model, e.g.:

- Model 1 (The on-line RUL update algorithm proposed in our study): Considering PI1 and PI2, using Bayesian and EM algorithm to update unknown parameters in IG process
- Model 2: Considering PI1 and PI2, only using EM algorithm to update parameters in IG process

During the parameter update process, we can find that parameter estimation may lack accuracy during the early stages due to limited data. However, as more degradation data accumulates, the estimated values tend to stabilize and improve accuracy. **Figure 3** shows the estimated degradation paths and corresponding 95% confidence interval obtained based on the parameter update results. It can be found that predicted degradation values at different measurement time are all within the confidence intervals. As more degradation data accumulates, the predicted degradation paths become increasingly accurate. The estimated degradation values for 9th test sample are much more accurate than the estimated degradation values for 10th test sample. The estimated degradation values for 9th test sample almost consistent with the real degradation trajectory.



(a) Estimated degradation paths for 9th test sample



(b) Estimated degradation paths for 10th test sample

Figure 3: Estimated degradation paths

Figure 4 shows the reliability curves obtained by off-line parameter estimation results, 9th and 10th sample on-line parameter estimation results. We can find that differences between the off-line reliability evaluation and individual reliability evaluation due to the unit-to-unit variability. If the reliability of individual component is assessed, it is necessary to update the parameters in the reliability analysis model. Or else, there will be mistakes when using population parameters to conduct individual reliability.

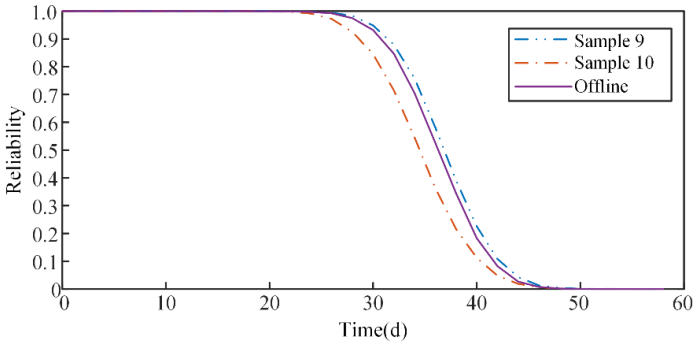


Figure 4: Estimated degradation paths

Figure 5 shows the PDF of estimated RUL. **Table 2** compares the mean square errors (MSE) between true RUL and estimated RUL. We can find that the estimated RUL by Model 1 is more accurate that the estimated RUL by Model 2. From **Figure 5**, it is noticeable that the PDF of RUL becomes progressively narrower as more degradation data accumulates. It indicates that the RUL prediction is getting more accurate. The PDF of RUL based on Model 1 is taller than the PDF of RUL based on Model 2. It means that Model 1 is more accurate than Model 2. In addition, the prediction accuracy of 9th test sample is higher than 10th test sample, which is also consistent with the prediction accuracy of degradation trajectories. Based on the above discussion, it can be concluded that employing Bayesian and EM algorithms together to update unknown parameters leads to more accurate RUL evaluation results.

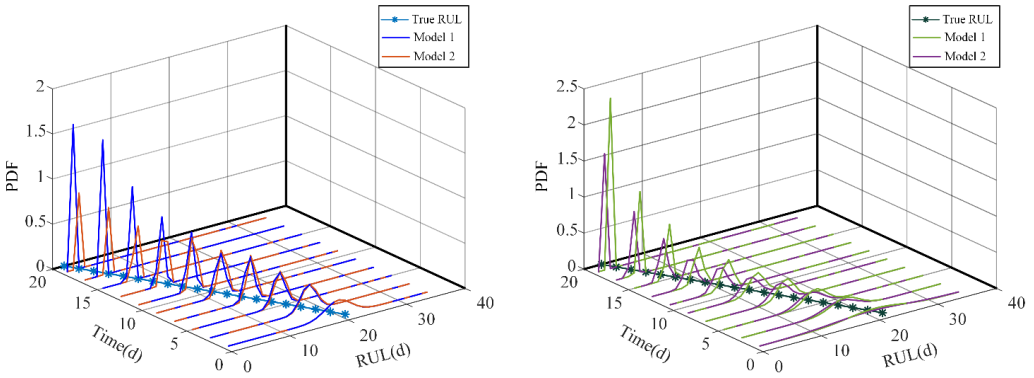


Figure 5: The PDF for RUL estimated by Model 3 (a) 9th test sample (b) 10th test sample

Table 2: MSE between true RUL and estimated RUL

Test Sample/MSE	Model 1	Model 2
9th Sample	0.3142	0.7962
10th Sample	0.9063	3.1562

5. CONCLUSION AND FUTURE WORK

This study utilizes nonlinear explainable IG process model considering the individual variability to describe two performance indicators. The time scale function is determined by failure mechanism model. The Frank Copula is employed to establish the reliability model, considering the correlation between the two performance degradation indicators. The initial values are determined by Bayesian MCMC method based on the degradation data. As for online RUL prediction, the two-stage parameter estimation approach is employed. The unknown parameter updates in the IG process model are achieved by combining the Bayesian and Expectation-Maximization algorithms, while the unknown parameter in the Frank Copula is estimated using maximum likelihood estimation. The effectiveness and accuracy of the proposed reliability assessment model are validated through an aging test of rubber O-ring seals conducted under the storage environment. The results indicate that the proposed model and algorithm significantly enhance the accuracy of remaining useful life prediction. Future work will concentrate on integrating the failure mechanism model and machine learning model to achieve more comprehensive and accurate reliability evaluation.

REFERENCES

[1] Mahankar PS, Dhoble AS (2021) Review of Hydraulic Seal Failures Due to Effect of Medium to High Temperature. Engineering Failure Analysis 127:105552

[2] Morrell PR, Patel M, Skinner AR (2003) Accelerated Thermal Ageing Studies on Nitrile Rubber O-rings. Polymer Testing 22:651-6

[3] Kömmling A, Jaunich M, Goral M, Wolff D (2020) Insights for lifetime predictions of O-ring seals from five-year long-term aging tests. Polymer Degradation and Stability 179:109278

[4] Liang B, Yang X, Wang Z, Su X, Liao B, Ren Y, Sun B (2019) Influence of Randomness in Rubber Materials Parameters on the Reliability of Rubber O-Ring Seal. Materials (Basel). 2019;12

[5] Zhang Z, Si X, Hu C, Lei Y (2018) Degradation Data Analysis and Remaining Useful Life Estimation: A Review on Wiener-process-based Methods. European Journal of Operational Research 271:775-96

[6] Sun B, Yan M, Feng Q, Li Y, Ren Y, Zhou K, Zhang W (2018) Gamma Degradation Process and Accelerated Model Combined Reliability Analysis Method for Rubber O-Rings. IEEE Access 6:10581-90

[7] Sun L, Gu X, Feng L, Di Y (2016) Reliability Analysis of Rubber O-rings Used in the Rockets. 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM): 1392-6

[8] Wahab A, Farid A (2011) Correlation between Compression-set and Compression Stress-relaxation of Epichlorohydrin Elastomers. Polymers and Polymer Composites 19:631-8

[9] Pan J, Bai G, Chen W (2018) Lifetime Estimation of Nitrile Butadiene Rubber O-rings under Storage Conditions Using Time-varying Copula. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 232:635-46

[10] Peng W, Li Y-F, Yang Y-J, Mi J, Huang H-Z (2017) Bayesian Degradation Analysis with Inverse Gaussian Process Models under Time-Varying Degradation Rates. IEEE Transactions on Reliability 66:84-96

[11] Chen R, Zhang C, Wang S, Qian Y (2021) Reliability Estimation of Mechanical Seals based on Bivariate Dependence Analysis and Considering Model Uncertainty. Chinese Journal of Aeronautics 34:554-72

[12] Joe H (1997) Multivariate Models and Multivariate Dependence Concepts. CRC press; 1997

DEVELOPMENT OF A HYDRAULIC ARTIFICIAL MUSCLE WITH HIGH FORCE DENSITY

peer reviewed

Mathias Niebergall*, Roman Strobel

[1]*Hydraulics Center of Excellence, University of Applied Science Ulm, Prittwitzstraße 10, 89075 Ulm*

* Corresponding author: Tel.: +49 731 96537-545; E-mail address: mathias.niebergall@thu.de

ABSTRACT

The actuation of mechanism like exoskeletons or devices for medical rehabilitation by means of fluid artificial muscles are convincing solutions due to their light weight, exceptional power capacity, remarkable resiliency, and low investment costs. The artificial muscle consist of an inner elastomeric hose surrounded by a textile, braided and reinforced by aramid fibers. The muscle is activated by fluid supply with a radial expansion of the inner pressurized hose accompanied by a corresponding axial contraction. Consequently circumferential stress in the textile reinforcement of the muscle converts into axial contraction force. The focus of this project is the development of high power hydraulic muscles that enable a significant higher pressure level as well as force density than known artificial muscles. Prototypes of new hydraulic artificial muscles have been developed and experimentally evaluated by means of a customized hydraulic test setup. Relating to the initial length of the muscle without fittings, a contraction of 32% has been measured. In this experiment the associated pressure level is 5 MPa. In a second experimental test the force depending on pressure has been measured and a high force density per mass of 60 kN/kg has been calculated.

Keywords: Hydraulic artificial muscle, McKibben muscle, High force density

1. INTRODUCTION

New innovative actuator systems for robotics, exoskeletons and devices for medical rehabilitation with hydraulic artificial muscles (McKippen principle) are remarkable powerful, compact, lightweight, robust, and less expensive solutions. The focus of the paper presented is a high power hydraulic muscle that enables a significant higher pressure level and force density than currently known artificial muscles. An overview of the known hydraulic muscle performances supplemented by results of the new developed hydraulic artificial muscle is shown in **Table 1** (compare section 6).

2. HYDRAULIC ARTIFICIAL MUSCLE

The actuator effect of a pressurized muscle towards a non-pressurized hydraulic artificial muscle is outlined in **Figure 1**. Fluid supply activates the initially non-pressurized muscle with the fiber angle α_0 (**Figure 1**, left). A radial expansion of the inner pressurized hose of the muscle leads to increasing circumferential stress in the surrounding textile sleeve and to a rising fiber angle α_1 (**Figure 1** right). The actuation effect is an axial contraction of the muscle that enables to lift a load F. The non-pressurized artificial muscle has an initial length of l_0 with a radius of r_0 and the pressurized muscle has a length of l_1 with a radius r_1.

The mechanical tensions of the hydraulic artificial muscle in axial direction (σ_a) and in circumferential direction (σ_c) have been determined according to strength calculations for cylindrical containers with

thin walls under internal pressure. The axial stress of the artificial muscle consist of two overlay causes, at first an internal hydraulic pressure and secondly an additional axial load *F*. Considering both causes, the axial stress can be calculated by the equation (1).

$$\sigma_a = \frac{p \cdot r}{2 \cdot t} + \frac{F}{2 \cdot r \cdot \pi \cdot t} \tag{1}$$

The circumferential stress of the artificial muscle, caused by an internal pressure, can be calculated by the equation (2).

$$\sigma_c = \frac{p \cdot r}{t} \tag{2}$$

The axial specific normal force can be calculated by the equation (3)

$$n_a = \sigma_a \cdot t \tag{3}$$

as well as the axial force F_a by the equation (4).

$$F_a = n_a \cdot 2 \cdot r \cdot \pi \tag{4}$$

The circumferential specific normal force can be calculated by the equation (5).

$$n_c = \sigma_c \cdot t \tag{5}$$

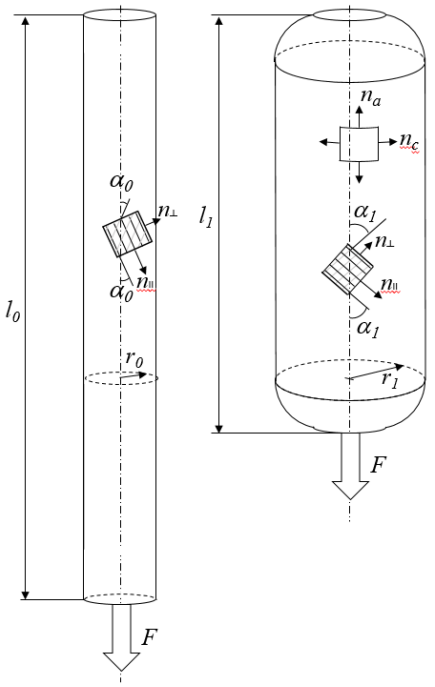


Figure 1: Actuation effect of a hydraulic artificial muscle (left: non-pressurized, right: pressurized)

The axial force F_a generated by the muscle enables the actuation effect, here the lift of the load F. The specific normal force in circumferential direction grows with increasing pressure and increasing radius as well as accordingly with increasing fiber angle. The forces in circumferential direction are always balanced.

3. DESIGN OF THE HYDRAULIC ARTIFICIAL MUSCLE AND OF THE TEST SETUP

The focus of the project is the development of high power hydraulic muscles that enable a higher pressure level, a higher force density than known artificial muscles and accordingly an improved compactness. Prototypes of the new hydraulic artificial muscles have been developed and manufactured. A sectional view of a prototype of an artificial muscle is shown in **Figure 2**. The muscle has an inner elastomeric hose surrounded by an aramid fiber textile sleeve connected by a fitting. The fitting consist of a swage nipple and a swage ferrule, as well as a screw coupling with an O-ring sealing element.

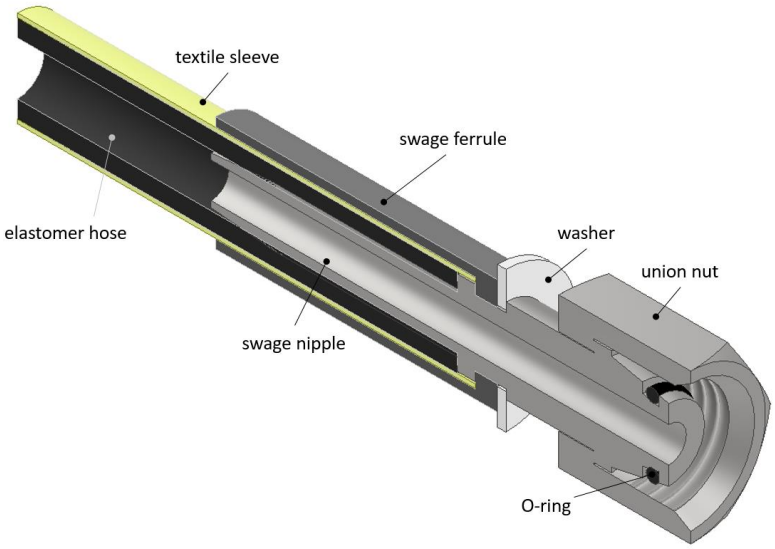


Figure 2: Sectional view of the hydraulic artificial muscle

The evaluation of the performance of hydraulic artificial muscles requires an appropriate hydraulic test setup, which has been designed and built for experimental investigations. The experimental setup is shown in **Figure 3**.

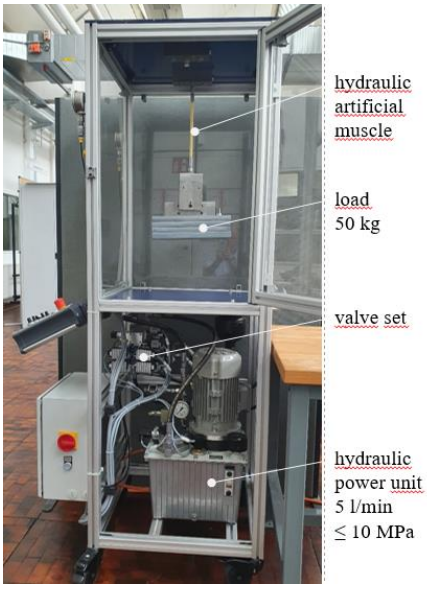


Figure 3: Hydraulic test setup for the investigation of the muscle contraction

4. EXPERIMENTAL DETERMINATION OF THE MUSCLE CONTRACTION

A non-pressurized hydraulic artificial muscle with an initial length of l_0 = 340 mm and a diameter of d_0 = 10 mm (radius r_0 = 5 mm) (compare **Figure 1**) has been configured. The mass of the muscle including fittings and filled with fluid is m = 0,145 kg. The artificial muscle has been activated by a volume flow with a corresponding increase in pressure up to 5 MPa. In **Figure 4** the non-pressurized muscle is shown on the left and the pressurized muscle on the right hand side. The actuation effect is a contraction of the artificial muscle with a remarkable action of force (depending on the small muscle diameter) to lift a load of 50 kg (F = 490 N). The pressurized artificial muscle is shortened to the length of l_1 = 232 mm, and has an expansion of the diameter to d_1 = 24 mm (radius r_1 = 12 mm). Relating to the initial length l_0 of the non-pressurized muscle, a contraction of ε = 32% has been calculated with equation (6). **Figure 5** shows the fiber angle of the non-pressurized muscle α_0 and of the pressurized muscle α_1. Initially the non-pressurized muscle has a fiber angle of α_0 = 25° (**Figure 5** left hand side). The volume flow into the muscle leads to an increasing pressure inside the muscle, to an increasing radial muscle volume, to an increasing circumferential stress in the textile sleeve, to an adaptive change of the fiber angle to α_1 = 50° (**Figure 5** right hand side) and finally to a muscle contraction with an intended actuation effect (**Figure 4** right hand side).

The contraction is calculated by the equation (6), without considering the length of the fittings.

$$\varepsilon = (l_0 - l_1) / l_0 \cdot 100 \qquad (6)$$

The stress in axial direction of the textile sleeve (thickness t = 0.45mm) has been calculated using equation (1) with the result of σ_a = 81.1 N/mm^2. The stress in circumferential direction calculated using equation (2) is σ_c = 133.3 N/mm^2. The internal hydraulic pressure dominates the stress in the textile. The stress of the artificial muscle under load is not critical for the aramid viber, due to a tensile strength of the aramid viber of σ_{AV} = 2800 N/mm^2. Improvements of the muscle robustness are carrying out at the connection of muscle (elastomer hose, textile sleeve) and swage ferrule/nibble.



Figure 4: Hydraulic artificial muscle non-pressurized (left) and pressurized (right)

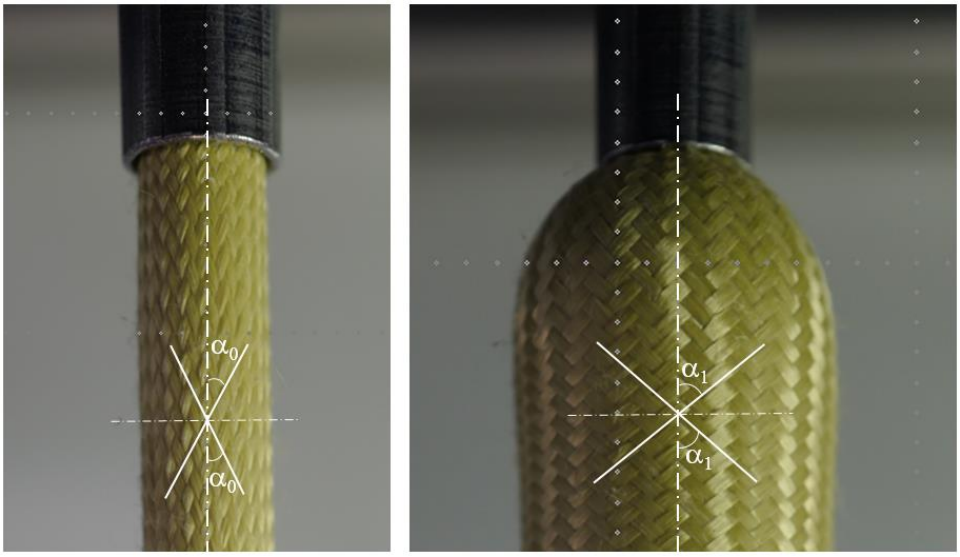


Figure 5: Fiber angle with a non-pressurized muscle α_0 = 25° and with a pressurized muscle α_1 = 50°

The conducted experimental test shows already a competitive contraction of ε = 32% with regard to the initial length of the muscle, without considering the fittings (compare **Table 1**). The relationship of the contraction and the applied pressure have been measured and is shown in **Figure 6**.

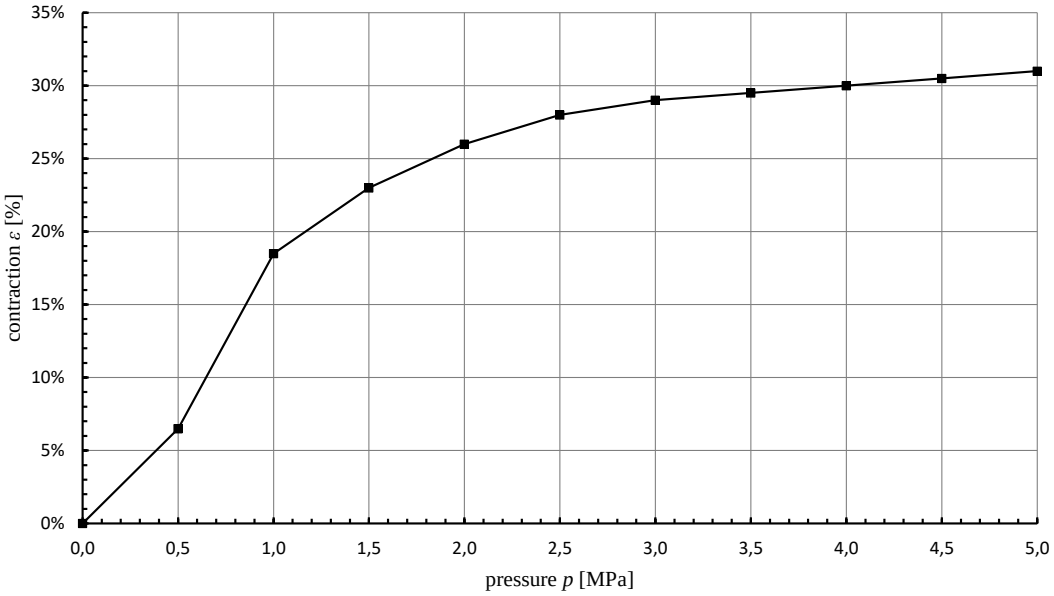


Figure 6: Measured contraction of the muscle with a load of 50 kg depending on the applied pressure

5. EXPERIMENTAL DETERMINATION OF THE FORCE DENSITY

To investigate the maximum force generated by the hydraulic artificial muscle, a measuring device has been developed (compare **Figure 7**). The measuring device has been equipped with a force sensor (20 kN) and a pressure sensor (10 MPa), and has been integrated in the chamber of the hydraulic test setup. The hydraulic artificial muscle has been clamped as shown in **Figure 7** to measure the blocked force and to prevent an axial contraction of the hydraulic artificial muscle. The muscle has an initial length of l_0 = 200 mm, a diameter d_0 = 10 mm, and a mass with fittings and filled with fluid of m = 0,120 kg. The measured force results of the blocked muscle depending on a continuous increased pressure within the muscle are shown in **Figure 8**. With a pressure level of 5.6 MPa a force of F_t = 7200 N has been measured and a force density per mass of δ_F = 60 kN/kg has been calculated with equation (7).

$$\delta_F = F_t / m \tag{7}$$

The energy performance of the hydraulic artificial muscle consists of muscle contraction and actuation force. The evaluation of the performance of the new muscle towards the state of the art is challenging, due to the variety of muscle specifications in terms of geometric dimensions, used materials, fiber angles, and pressure levels. To enable a comparability of the various muscle performances, the force

density per mass has been calculated or estimated for the considered muscles, if possible. **Table 1** show a summarized representation of the specifications and performances of selected relevant hydraulic artificial muscles, supplemented by the data of the new developed hydraulic muscle. The conducted experimental test shows already a very competitive result with a high force density per mass of δ_F = 60 kN/kg.

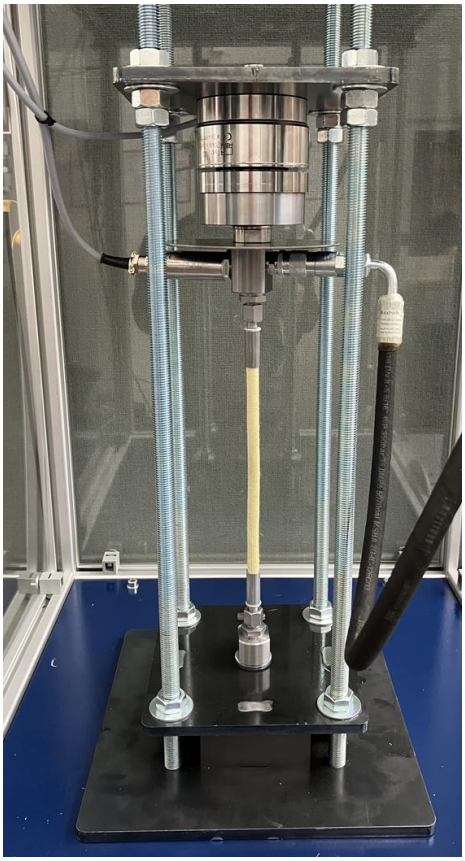


Figure 7: Measuring device to determine the maximum blocked force F_t of the clamped muscle

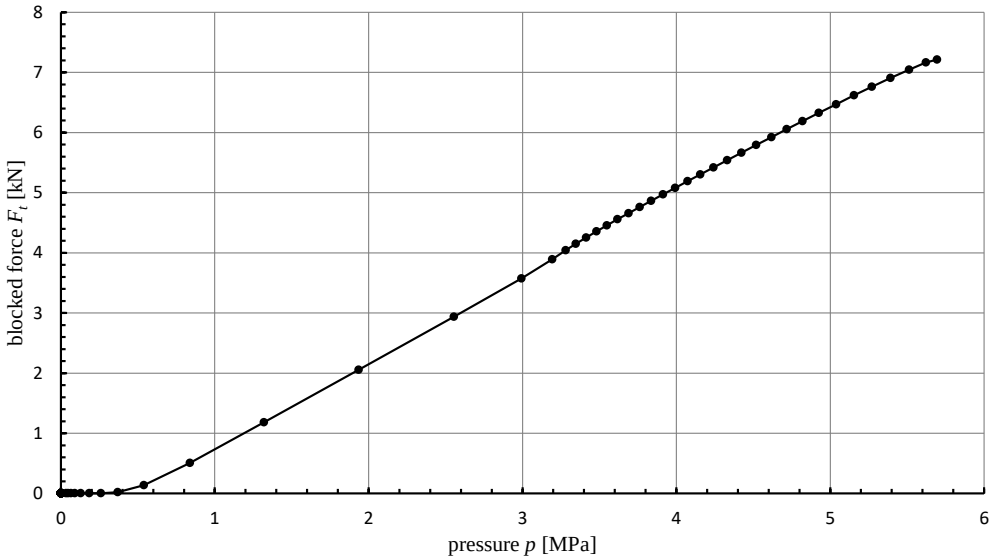


Figure 8: Measured force F_t depending on the pressure p within the hydraulic artificial muscle

Table 1 shows a selection of known publications of artificial muscle concepts with their specification data and the belonging force density per mass.

Table 1: Comparison of muscle specifications and performances

	present paper (2023)	Zhang et al. (2021) [1]	Zhang et al. (2017) [2]	Tiwari et al. (2012) [3]	Sangian et al. (2015) [4]	Iwata et al. (2012) [5]	Mori et al. (2009) [6]	unit
initial muscle length l_0	200	300	200	160	80	300	700	mm
initial outer diameter d_0	10	30	30	3.17	6	21	40	mm
initial fiber angle α_0	25	25	43		35	23.5	25	°
maximum pressure p	5.6	4	6	0.69	0.25	7	4	MPa
maximum contraction ε	**32**	23	20 [1]	19 [1]	28	23 [1]	25	%
maximum force F_t	7 200	23 500	12 000	40	26	8 000	28 000	N
mass of muscle m	0.12	0,6 [3]	0.5	0.002	0.0018	unknown	3	kg
force density per mass δ_F	**60**	40 [3]	24 [2]	20 [2]	14.4 [2]	not estimated	9.5	kN/kg

[1] calculated with equation (6)
[2] calculated from published measured forces and mass of the muscles
[3] estimated from published data in [1] and [2]

6. SUMMARY

The focus of the paper presented is the development of high power hydraulic artificial muscles that enable significant higher pressure levels and force densities. That is crucial for small and lightweight drives, for e. g. medical rehabilitation. Experimental results show a remarkable action of force depending on the small muscle diameter to lift heavy loads. The observed contraction of the artificial muscle is ε = 32% with regard to its initial length, without considering the fittings.
By means of a further experimental test the possible maximum blocked force F_t of a hydraulic artificial muscle (l_0 = 200 mm, diameter d_0 = 10 mm) has been measured. With a pressure level of 5.6 MPa a blocked force of F_t = 7200 N has been measured with a very competitive force density per mass of δ_F = 60 kN/kg (muscle mass m = 0,120 kg).

Those results show, that hydraulic artificial muscles are attractive future actuation solutions for e.g. robotic and exoskeleton applications or for various devices for medical rehabilitation.

Acknowledgement
The authors would like to thank the foundation council of the Hydraulics Center of Excellence of the Ulm University of Applied Sciences for the friendly support of the project.

NOMENCLATURE

d_0	diameter of the non-pressurized muscle	mm
d_1	diameter of the pressurized muscle	mm
F	force of the axial load	N
F_a	axial force generated by the artificial muscle	N
F_t	measured total force, caused by the internal pressure of the muscle	N
l_0	initial length of the textile sleeve	mm
l_1	length of the contracted textile sleeve	mm
m	mass of the contracted muscle including fittings and fluid	kg
n_a	specific normal force in axial direction (per unit of the circumference)	N/mm
n_c	specific normal force in circumferential direction (per unit of the textile mantle length)	N/mm
p	internal pressure	MPa
r	radius of the artificial muscle	mm
r_0	radius of the non-pressurized muscle	mm
r_1	radius of the pressurized muscle	mm
t	thickness of the textile mantle	mm
α_0	initial fiber angle of the non-pressurized muscle	°
α_1	fiber angle of the pressurized muscle	°
δ_F	force density per mass	kN/kg
ε	maximum contraction	%
σ_a	stress in axial direction	N/mm^2

σ_c	stress in circumferential direction	N/mm^2
σ_{AV}	tensile strength of aramid viber	N/mm^2

REFERENCES

[1] Zhang, Che, et al. (2021) A modified physical model of high strength water hydraulic artificial muscles considering the effects of geometry and material properties. Journal of Mechanical Engineering Science 236.5

[2] Zhang, Hou, et al. (2017) Modeling and experiments on the drive characteristics of high-strength water hydraulic artificial muscles. Smart Materials and Structures 26.5

[3] Tiwari, Rashi, et al. (2012) Hydraulic artificial muscles. Journal of Intelligent Material Systems and Structures 23.3

[4] Sangian, Naficy, et al. (2015) The effect of geometry and material properties on the performance of a small hydraulic McKibben muscle system. Sensors and Actuators, A: Physical

[5] Iwata, Suzumori, et al. (2012) A method of designing and fabricating McKibben muscles driven by 7 MPa hydraulics. International Journal of Automation Technology 6.4

[6] Mori M, Suzumori K et al. (2009) Development of very high force hydraulic McKibben artificial muscle and its application to shape-adaptable power hand. Proceedings of the 2009 IEEE International Conference on Robotics and Biomimetics

HYDRAULIC PILE HAMMER SURROGATE MODEL BASED ON PHYSICS-INFORMED NEURAL NETWORK

peer reviewed

Yajun Liu[1,*], Junfeng Zeng[1], Nengchang Guo[1], Guoda Pang[2], Judong Pang[2], Zhijian Wei[2], Yang Lu[2], Jianwei Liu[1]

[1] *South China University of Technology, School of Mechanical & Automotive Engineering, Guangzhou Guangdong 510700, China*
[2] *Guangdong Liyuan Hydraulic Machinery Co., Ltd, Technical department, Foshan Guangdong 528300, China*

* Corresponding author: Tel.: +86 13640338486; E-mail address: yajun@scut.edu.cn

ABSTRACT

The high fidelity digital model of hydraulic pile hammer is able to predict the energy conversion rate of hydraulic system under the specified controller parameters, and guide the matching process of controller parameters and construction conditions. Using neural network to fit the calculation results of simulation software can greatly reduce the calculation cost of prediction process. However, the unexplained feature of neural network output increases the application risk of this method. Based on the classical theory of physics-informed neural network (PINN), a PINN method based on inequalities is proposed in this paper. Based on this method, a physics-informed surrogate model network (PISMN) oriented to the simulation process of hydraulic pile hammer was constructed. It is proved that this method can constrain the output of the surrogate model, improve the stability of the training process, and the median prediction deviation of the prediction results was reduced by 53.0% in the validation set with perturbation.

Keywords: Hydraulic piling hammer, Proxy model, Neural networks

1. INSTRUCTIONS

Through the high fidelity digital simulation model of hydraulic hammer, the energy conversion rate of hydraulic system of the pile hammer can be calculated. With the help of the calculation results, the post-processing program can match the control parameters of the pile machine with the working conditions to maximize the energy conversion efficiency. However, the application of this method is limited by the tight computing power in pile hammer equipment. Therefore, it is necessary to generate a surrogate model, which should meet the following conditions: perform accurate calculations and replace simulation software, able to calculate the energy conversion rate of the hydraulic system in the construction process of the hydraulic pile hammer, and operate efficiently in the equipment with limited computing power.

There are lots of kinds of surrogate model, including polynomial response surface (PRS), Support vector regression (SVR), Radial basis functions (RBF), Kriging models, artificial neural networks (ANN), etc. By encoding the simulation database and storing it in the computing power constrained device, and using the traditional surrogate model, it is a method to approximate the prediction results of the simulation model. However, this method has the following shortcomings: it can not use the historical data of hydraulic piling construction process to correct the prediction results. Among those above methods, the surrogate model based on neural network can solve the above problems through reasonable structure design. Artificial neural network method has been widely used in the recent

research of surrogate model because of its advantages in computing complex problems [1]. Compared with other surrogate models, benefiting from its iterative learning mechanism, this method has the characteristics of high calculation accuracy and strong robustness. However, this method has the following shortcomings: lack of reasoning and causality expression ability, unable to explain the reasoning process of decision-making, etc. In order to solve the above problems, the researchers proposed to use the physics-informed method to generate the neural networks to solve the problem, and the method has been applied in many fields.

The physics-informed neural networks (PINN) was first proposed by Raissi et al., and has been introduced in several subsequent paper [2]. The research follows an earlier idea of using the laws of physics to constrain neural networks. In the classical PINN study, those previous research results [3] were combined with deep neural networks to overcome the limitations of the traditional partial differential equations (PDE) solver, so as to solve those more complex problems. In addition, the study of Arnold et al. [4] also proved that this method is able to solve the problem of poor generalization ability of neural network caused by small amount of training sample data or insufficient sampling.

However, it becomes difficult to apply PINN when the relevant parameters in the physical equation are time-varying or cannot be accurately measured for some reasons. At present, in the researches of controllers using physics-informed neural network, the physical background of the controlled object is certain, which can be described systematically by relative differential equations or partial differential equations. Therefore, the related research of PINN in complex systems or other systems with insufficient prior knowledge still needs further research.

2. ALGORITHM

2.1. Surrogate model overview

When the environmental factors of hydraulic pile hammer are determined, the efficiency of hydraulic system of pile hammer can be predicted by simulation software. In order to solve the problem caused by high computational power requirement, the surrogate model based on neural network is proposed to replace the simulation model. The surrogate model in this paper is mainly used to predict the time of reciprocating movement cycle of pile hammer. According to the calculation results, the parameters such as the number of hammer strikes per minute and the energy conversion rate of the pile hammer hydraulic system can be indirectly calculated. And the calculation results will also be used to optimize the controlling parameters of pile hammer during the construction process.

The following figure (**Figure 1**) shows the basic structure of the surrogate model, which is named physics-informed surrogate model network (PISMN) in this paper. The output of the model is the period of a reciprocating movement of the pile hammer, and the symbol is t. Input vector of the model is $\{s, p_a, r\}$, including the vector of controlling parameter during construction process(marked as s), measurable physical parameter vector of the pile machine(marked as p_a) and the vector of historical data during the pile hammer operation (marked as r). And the vector of controlling parameter during construction process include three indexes: the input power of the hydraulic pump (marked as P), the delay time(marked as t_d) before lifting the hammer, that is, the time that hammer stay on the pile anvil after the impact, and the lifting height of the hydraulic pile hammer (marked as h).

For the pile hammer hydraulic system, after the hammer impacts the pile anvil, the movement direction of the hydraulic medium flowing in the cylinder changes, and then the pressure and flow in the hydraulic system fluctuate. In addition, the reaction force may cause the hammer to bounce

upward and further aggravate the fluctuations of pressure and flow. Studies have shown that large amounts of energy are often lost when valves switch between the two states, and solutions such as soft switching have been proposed. Considering that a long delay time may slow down the process of construction, so it is necessary to choose a suitable delay time during the construction process.

The selection of each physical factor, which is included in the network input vector, is mainly based on the specifications of pile foundation construction and selection suggestion form hydraulic pile hammer designer. The vector of measurable pile machine physical parameters (P_a) is mainly used to distinguish different electro mechanical systems of hydraulic pile hammer. The parameters contained in vector P_a are the physical factors that have a significant impact on the simulation results in the simulation model and can be accurately obtained, including: pile hammer mass, hydraulic cylinder inner diameter and cylinder rod diameter, etc. The historical data of the construction process vector (r) mainly includes: the historical construction parameters and the its corresponding pile hammer movement period. In this model, we have $r = \{s_{k-1}, s_{k-2}, \cdots, s_{k-i}, t_{k-1}, t_{k-2}, \cdots, t_{k-i}\}$ and $i \in [1,5]$.

The PISMN proposed in this paper consists of two sub-networks: feature extraction network (marked as $\beta(\cdot)$) and prediction network (marked as $\mu(\cdot)$).The feature extraction network can be regarded as the encoder of historical data of pile machine construction. The input parameters of the network are: the measurable physical parameter vector and the historical data vector which is record under the same physical parameter condition. The output vector of the neural network represents the system feature $\hat{P}_u$, which is used to extract unmeasurable physical parameters that contained in historical data and used to reduce the dimension of input vector of prediction network. The prediction network is used to predict the period movement cycle of pile hammer, and the predicted value is denoted by $\hat{t}$. The prediction network input includes: the vector of controlling parameter during construction process (S), the measurable physical parameters of pile hammer (P_a) and system feature ($\hat{P}_u$) extracted by extraction network $\beta(\cdot)$.

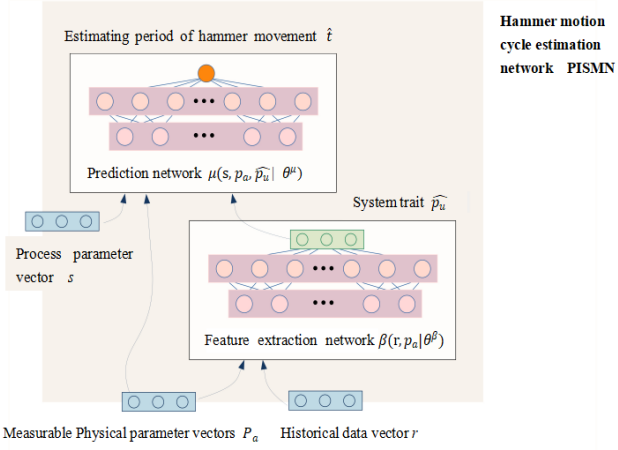


Figure 1: Structure of the PISMN algorithm

The number of hammer strikes per minute can be calculated directly through the reciprocating movement period of the pile hammer, which directly reflects the construction efficiency. In addition, the hydraulic efficiency of the pile hammer can be calculated by combining the pump outlet power and the lifting height of the pile hammer, which directly reflects the energy conversion efficiency of the pile hammer hydraulic system. The data transfer flow diagram of the surrogate model mentioned above is shown in the figure below (**Figure 2**).

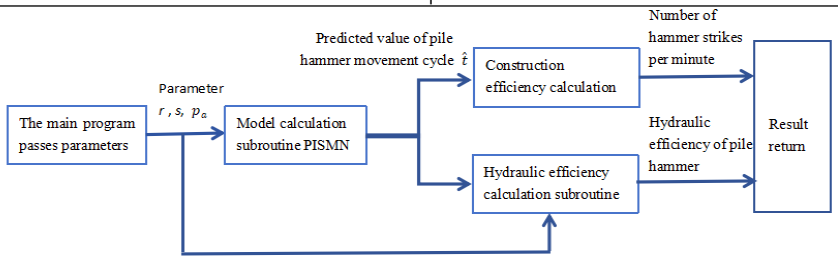


Figure 2: Flow chart of data in PISMN

2.2. Deployment of algorithms

The PISMN algorithm will fit the data set from the simulation platform mentioned in Chapter three. The simulation data generated in the platform include: the physical parameters $\{p_a, p_u\}$ of the hydraulic pile machine during simulation, where p_a are the parameters of the pile machine that can be measured , including the weight of the pile hammer (marked as m) ; the parameter p_u stands for pile machine parameters that are difficult to measure accurately , including overflow coefficient (marked as C_{im}, C_{em}), flow coefficient (marked as C_d) and so on. The above parameters are determined by the conclusion of Chapter two of this paper. Although the changes of these parameters have a great impact on the simulation results, the true values (marked as p_u) of these parameters will not be introduced into the PISMN training program. The PISMN network needs to extract the hidden features of physical parameter p_u which are contained in the historical data according to the historical parameters (marked as r) under the same physical parameters (marked as $\{p_a, p_u\}$) .

Table 1: The PISMN model training procedure pseudo-code

Algorithm 1 PISMN **Model training process**

Input : Simulation result data set R

Output : A trained predictive network model $\mu(\theta^{\mu})$;A trained feature extraction network model $\beta(\theta^{\beta})$; Optimization objective is:

$$\hat{\theta} = \arg\min \frac{1}{N} \sum_{i=1}^{N} L(t, \hat{t}, \gamma \mid \theta^{\mu}, \theta^{\beta})$$

1: function TRAIN-PISMN(R)
2: Initialize the simulation result R of the model analysis simulation platform, and establish the key-value pairs between parameter simulation parameter and simulation result.
3: while i < epoch do
4: Three groups of parameters { p_A, p_U, s } are randomly selected .
5: The construction parameters of the pile hammer and the corresponding movement period of the pile hammer under the other five groups of simulation parameters were randomly selected and recorded as r.
6: The predicted value of the movement period of the hammer is obtained by network forward calculation $\hat{t}$.
7: Obtain the partial derivative $\frac{\partial \hat{t}}{\partial h}$ of the network output with the drop weight height, and calculate the loss function L:

$$L = MSE(t, \hat{t}) + MSE_{penalty}(\gamma \cdot \max(-\frac{\partial \hat{t}}{\partial h}, 0), 0)$$

8: Update network model parameters:

$$\mu(\theta^{\mu}) \leftarrow \mu(\theta^{\mu(i+1)})$$

$$\beta(\theta^{\beta}) \leftarrow \beta(\theta^{\beta(i+1)})$$

9: end while
10: end function

The algorithm model training process is summarized as follows:

(1) The training process first loads the data set(R) of the simulation model, and establishes the

key-value pairs of the control parameters, physical parameters and the movement period of the pile hammer during the construction process of the pile machine, so as to realize the mapping between the input and output values during the training process. Then specify the training cycle, and enter the training cycle;

(2) Randomly select a set of physical parameters $\{p_a, p_u\}$; A set of construction parameters are selected based on the physical parameters, and the corresponding simulation software prediction results are recorded for the subsequent loss function calculation. Then, based on the physical parameters, another five groups of construction parameters and their corresponding pile hammer movement period are randomly selected as historical data (R).

(3) Parameters were passed into the PISMN feature extraction network to extract the features of the pile machine.

(4) Calculate the loss function, update the network parameters, and make it continuously optimized.

2.3. Experimental datasets

Based on the above, the training data set of the proxy model is generated by the simulation model. In the process of establishing the simulation platform, the degree of matching with the target model should be considered first. The high-fidelity digital model will improve the accuracy of the transfer learning algorithm in the process of predicting the data in the real machine domain. **Figure 3** shows the physical diagram and actual pile machine drawing of a certain type of hydraulic pile hammer.



Figure 3: Physical map and actual drawing of hydraulic pile hammer

For the hydraulic pile hammer simulation model, the relevant simulation test variables mainly focus on the physical quantities that have a certain impact on the hydraulic energy conversion rate of the model. The simulation database was established to record the following data contents: the values of all test factors in the simulation process, the simulation results (including the energy conversion efficiency of the hydraulic pile hammer in the process, the reciprocating movement cycle of the pile hammer, etc.). Among them, the test factor refers to a group of processed factors that are changed in the simulation process and are to be compared. Specific to this study are the physical quantities mentioned in before (overflow coefficient, flow coefficient, pile hammer mass, hydraulic cylinder internal geometry, etc.) and hydraulic pile hammer construction control parameters. Among them, the physical quantities that can be accurately measured will be read by the proposed algorithm in the subsequent research. The physical quantities that cannot be accurately measured are mainly used to record and distinguish different sets of simulation data. For the pneumatic test model, the simulation database is basically the same as that of the hydraulic pile hammer simulation model. Due to the

characteristics of the pneumatic platform, the input power P of the hydraulic pump in the construction control parameters of the pile hammer in the platform is changed to the output pressure of the reducing valve.

Generally, the simulation process of simulation software is divided into:

(1) Build a simulation sketch according to the actual schematic diagram;

(2) Select the sub-model of each component according to the actual situation;

(3) Set relevant parameters of the model;

(4) Set test factors and design simulation batch process based on relevant test factors;

(5) Conduct the simulation process and record the relevant simulation results;

(6) Post-processing of simulation data to facilitate data input of PISMN model, which can also be called data preprocessing of PISMN model. The following figure (**Figure 4**) is the simulation flow chart.

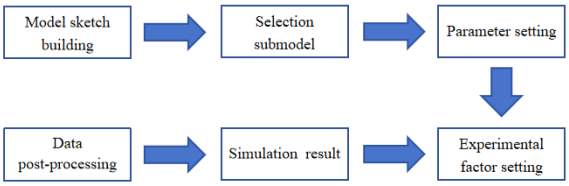


Figure 4: Simulation flow chart

3. EXPERIMENT AND RESULTS

In order to comprehensively evaluate the proposed agent model, this chapter will first introduce the training process of the neural network agent model, and then test the trained model with validation data set. In the above process, other neural network agent models will be used for horizontal comparative analysis. Finally, the influence of the physical knowledge embedding method on the agent model is further discussed.

3.1. Experimental results of model training

This section will first study the training process of the model, and compare the stability and convergence of different agent models in the training process. The following figure shows the change curves of the training process loss function of PISMN algorithm, BPNN algorithm and SGAN algorithm when the values are equal to 0, 0.5 and 1 respectively. At the same time, the variance of the return value of the training process loss function of each algorithm can be obtained through 10 repeated experiments. Then the curve and the area enclosed by it can be used to indicate that the return value of each method's loss function has a greater probability of falling in this region during the training process. When the parameter value of PISMN model described in the present paper is set to 0, the physical embedding method will no longer play an optimization role on the original network. Therefore, the following text will use the identification SMN to represent: the PISMN algorithm at that time. In addition, PISMN-G0.5 and PISMN-G1 methods shown in the figure respectively represent the PISMN algorithm when γ equals to 0.5 or 1.

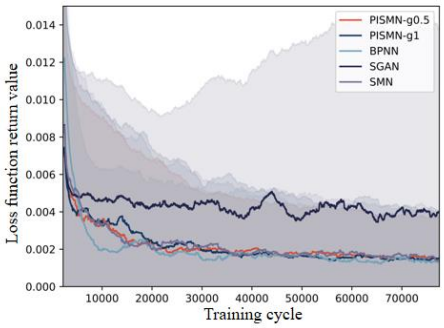


Figure 5: Change curve of loss function in training process of each algorithm

It can be observed from the experimental results in the following figure that all algorithms can converge effectively under this research background. However, the changes in the return value of the loss function of each algorithm are quite different. Among them, the convergence speed of the BPNN algorithm is faster, and the return value of the loss function at the end of training is also slightly lower than that of the PISMN algorithm. However, it can be seen from the research in the verification set in the following paper that the BPNN algorithm has overfitting phenomenon. Compared with other algorithms, there is a big difference between SGAN algorithm and other algorithms in the return value of loss function at the end of training process. In the relevant research of adversarial neural networks (Gans), problems such as non-convergence, mode collapse and slow convergence speed have become important constraints restricting the further application of GAN and its derivative methods. Although the SGAN algorithm is used to replace the traditional GAN algorithm, and the parameters of the SGAN algorithm are also optimized, the experimental results show that such a generation model is not suitable for the regression problem in the background of this study.

The following figures (**Figure 6** and **7**) shows the change curve of the ratio between the predicted value and the true value in the training process of each algorithm. In the figure, each algorithm uses two curves respectively to show this result. The two curves respectively represent the sum and difference of the mean and standard deviation of the ratio of predicted value and true value in 10 repeated experiments. The experimental results show that at the beginning of training, all the algorithms can quickly converge to within 10% error. However, at the end of the training stage, all the data within one standard deviation of the ratio of predicted value to true value fell within 5% deviation range in all samples of all algorithms except SGAN algorithm.

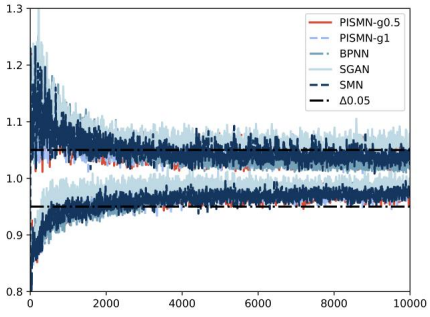


Figure 6: Local diagram at the beginning of training

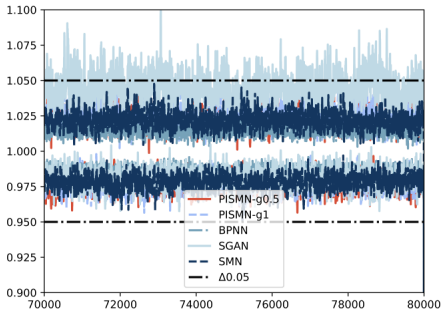


Figure 7: Local diagram at the end of training

3.2. Analysis of test set verification results

In order to evaluate each neural network agent model objectively, the following test sets are set to verify the comparison experiment. That is, after the training is completed, the data not used in the training process of the model (which does not affect the selection of ordinary parameters or hyperparameters) is used to perform an unbiased estimation of the prediction performance of the model.

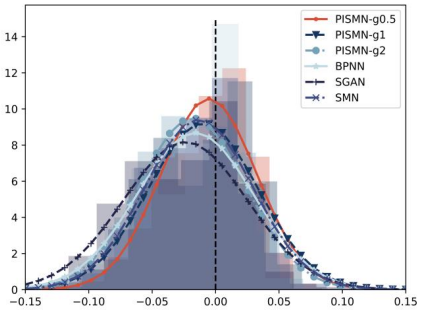


Figure 8: The absolute error distribution diagram of each algorithm in verification set 1

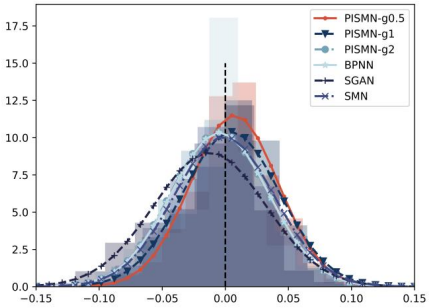


Figure 9: The absolute error distribution diagram of each algorithm in verification set 2

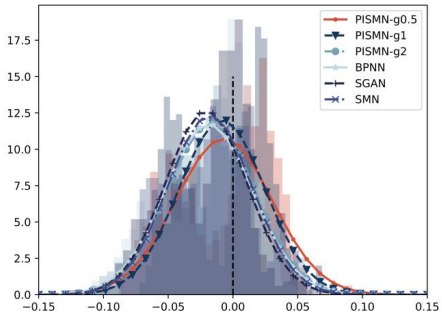


Figure 10: The absolute error distribution diagram of each algorithm in verification set 3

In the test experiments for each model, three validation sets with different uses were used. Each verification set is used separately to generate simulation parameters of the simulation model different from the training set. **Figure 8** to **Figure 10** respectively show the absolute error distribution diagram of the predicted values of each algorithm in the three verification sets.

According to the relevant data, in verification set 1, PISMN model with a value of 0.5 is superior to other methods in multiple indicators. In verification set 2, the BPNN method has a smaller deviation to the original simulation model, but the number of outliers is larger and the span of outliers is larger. In the relevant verification results of verification set 3, PISMN algorithm has better comprehensive performance, and the median and average of predicted value deviation have better performance in related algorithms. At the same time, it can be observed from Figure 3-18 that only two outliers appear in SGAN algorithm, and the variance of prediction bias of each algorithm has no significant change compared with the experimental results of other verification sets. The reason for this phenomenon may be that the degree of similarity between the data of verification set 3 and the training set is higher than that of verification set 1 and verification set 2. In addition, the prediction errors of each algorithm basically meet the normal distribution and meet the expectation. At the same time, it can be seen from Table 3-2 that when the value is 0.5, PISMN method can improve the prediction deviation median and mean accuracy by more than 54.4% compared with other algorithms in verification set I and verification Set III, and its comprehensive performance is significantly better than other comparison methods.

Further, from **Figure 8** to **Figure 10** above, it can be observed that the agent model optimized by the physical embedding method (i.e., the agent model with values of γ of 0.5 and 1 respectively) has a smaller difference between the upper and lower limits of the deviation of outliers, and the number of outliers distributed on both sides of the line γ=0 is more uniform than that of other agent models optimized by the physical embedding method. Therefore, it can be concluded that compared with the original model, the prediction accuracy and predicted value fluctuation of the model optimized by the physics knowledge embedding method are optimized to some extent. That is, by introducing the physical embedding method, the loss function term of the model is improved, which has a positive effect on the comprehensive performance of the agent model neural network.

From the above analysis, the conclusion of this section can be drawn: SGAN algorithm is not suitable for this research background, and BPNN method can converge faster than other algorithms. The PISMN method proposed in this paper and its derivative methods can converge stably. At the same time, the stable convergence of the algorithms also proves that it is feasible to explore the physical parameters of the pile machine by using the historical data (marked as r) of the hydraulic pile hammer. PISMN method has more stable performance and higher prediction accuracy than other algorithms in multiple verification sets. The comparison of SMN and PISMN ablation experiments also proved that

the physics knowledge embedding method can improve the comprehensive performance of the model.

4. DISCUSSION AND CONCLUSION

In this chapter, the agent model based on neural network is constructed, and the agent model of digital simulation of hydraulic pile hammer is constructed while extracting the relevant information of pile hammer by using the historical operation data of pile machine. The proxy model consists of two parts: the feature extraction network of historical data and the cycle prediction network of hammer movement. Among them, the function of the feature extraction network is to extract the physical parameters of the pile machine contained in the historical data, and pass them into the predictive neural network together with the measurable physical parameters and the construction control parameters of the pile machine, and finally achieve the prediction of the movement cycle of the pile hammer. Among them, the effectiveness of feature extraction network will be further discussed in the next chapter. By comparing PISMN network with other neural network-based proxy model algorithms, it can be concluded that PISMN has certain advantages in convergence speed and prediction accuracy.

REFERENCES

[1] Sun G, Wang S. A review of the artificial neural network surrogate modeling in aerodynamic design. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2019, 233(16): 5863-5872.

[2] Raissi M, Perdikaris P, Karniadakis G E. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. arXiv preprint arXiv:1711.10561, 2017.

[3] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational physics, 2019, 378: 686-707.

[4] Arnold F, King R. State–space modeling for control based on physics-informed neural networks. Engineering Applications of Artificial Intelligence, 2021, 101: 104195.

[5] Bhosekar A, Ierapetritou M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. Computers & Chemical Engineering, 2018, 108: 250-267.

[6] Sun G, Sun Y, Wang S. Artificial neural network based inverse design: Airfoils and wings. Aerospace Science and Technology, 2015, 42: 415-428.

[7] Zhao X, Huang B, Chen T, et al. Numerical simulations and surrogate-based optimization of cavitation performance for an aviation fuel pump. Journal of Mechanical Science and Technology, 2017, 31: 705-716.

[8] Lutter M, Ritter C, Peters J. Deep lagrangian networks: Using physics as model prior for deep learning. arXiv preprint arXiv:1907.04490, 2019.

[9] Antonelo E A, Camponogara E, Seman L O, et al. Physics-informed neural nets for control of dynamical systems. arXiv preprint arXiv:2104.02556, 2021.

[10] Rannow M B, Li P Y. Soft switching approach to reducing transition losses in an on/off hydraulic valve. Dynamic Systems and Control Conference. 2009, 48920: 613-620.

Chapter 3

Water Hydraulics

HOLISTIC EFFICIENCY MEASUREMENTS OF A MOBILE WORKING MACHINE: COMPARISON OF CONVENTIONAL MINERAL OILS AND A SUSTAINABLE WATER-BASED FLUID

peer reviewed

Sebastian Deuster[1]*, Katharina Schmitz[1]

[1]*RWTH Aachen University – Institute for Fluid Power Drives and Systems (ifas)*

* Corresponding author: Tel.: +49 241 80-47740; E-mail address: sebastian.deuster@ifas.rwth-aachen.de

ABSTRACT

The selection of pressure fluids plays an important role in hydraulic systems' efficiency. Depending on physical properties, especially the viscosity, different volumetric and hydraulic-mechanical losses occur in the system. This paper investigates the influence of hydraulic fluids with different physical properties on a crawler excavator's energy efficiency. For this purpose, the overall efficiency of the excavator is examined for different mineral oils and an alternative water-based fluid with deviating viscosity classes and indexes. The study results show that reducing the energy demand over a wide temperature range by lowering the viscosity grade or increasing the viscosity index is not generally feasible. To reduce losses by adapting the hydraulic fluid in a hydraulic system further physical and chemical properties such as pressure viscosity behaviour as well as the density must be considered.

Keywords: Efficiency measurement, Sustainable fluid, Excavator, Dig and dump, Viscosity, Viscosity Index, Water-based fluids, Viscosity improver

1. INTRODUCTION

Reducing greenhouse gas emissions is one of the main concerns regarding sustainable environmental and climate protection. Legally, the reduction is regulated by the Paris Climate Agreement [1]. In Germany, greenhouse gas emissions must be reduced by 95 % by 2050 compared to 1990. Globally, mobile machines are responsible for 1 % of total CO_2 emissions [2].

Optimizing the efficiency of hydraulic systems plays a decisive role. As shown in [3], major efficiency improvements can be achieved by adapting the hydraulic system´s architecture. This leads to a corresponding reduction in fuel consumption and consequently, to a reduction in greenhouse gas emissions. Further efficiency improvements can be reached by adaptation to the hydraulic systems components. The main attention concerning research activities is paid to improving hydraulic pump efficiency. Axial piston pumps are often part of the research work, because of their widespread usage [4, 5]. In addition to friction and volumetric losses in tribological contacts of hydraulic displacement units, power losses in hydraulic systems mainly occur due to pressure losses in pipes or components such as valves [6, 7]. An approach of today's research activities to reduce pressure losses in hydraulic components and flow paths is optimizing flow geometries [8–10]. Furthermore, research activities already exist to determine the influence of hydraulic fluids and their physical properties on the efficiency of hydraulic components [11] and elementary hydraulic systems [12].

In this paper, the influence of hydraulic fluids with different physical properties on the energy efficiency of a crawler excavator will be investigated. Therefore, the overall efficiency of the excavator is examined for different mineral oils and an alternative fluid with deviating viscosity classes and indexes. For this reason, hydraulic fluids are tested while executing an automated dig and dump cycle at different temperature stages. With this method, it is possible to compare and analyse

the efficiency of the excavator while using different hydraulic fluids. The research can be related to the investigations in [13]. The test excavator of the previous research was adapted to optimize the reproducibility as well as the accuracy of the measurements. The conventional drive unit, the diesel engine, was replaced by an external electronic drive unit. Furthermore, it is possible to temper the tank volume of the hydraulic unit as well as to record the drive power directly via torque measurements instead of fuel consumption measurements.
The paper is structured as follows: First, the state of the art is presented concerning the loss mechanism that occurs in hydraulic systems, divided into hydraulic mechanical losses and volumetric losses. After that, the losses are assigned to components that are used in the excavators' hydraulic system. Furthermore, the influence of the hydraulic fluid's physical properties on the loss mechanism is discussed. The state of the art is concluded with a presentation of the different viscosity behaviour of various base fluids used in this study concerning relevant operating parameters regarding temperature and pressure. The third chapter deals with the testing setup of the investigations. The excavator and its hydraulic system architecture are presented. Afterwards, the testing cycle is explained. Chapter three concludes with a presentation of the hydraulic fluids investigated. In the following chapter, the evaluation of the tests is discussed. For this purpose, different aspects are considered, such as energy consumption and temperatures when using the different fluids as well as the characteristics of different types of losses in the hydraulic system. The paper concludes with a summary of the results and an outlook on future research aspects.

2. STATE OF THE ART

The following chapter deals with the fundamentals of losses in hydraulic systems and their components. Therefore, first, the loss mechanisms in hydraulic systems are presented followed by the classification of these losses to specific components of hydraulic systems. The chapter is completed by the description of the viscosity behavior of different base oils, which have great significance on loss characteristics.

2.1. Loss mechanisms

According to [14], the power loss P_{loss} in an hydraulic system results from the input power P_{in} generated by the hydraulic pump and the total efficiency η_{total} of the hydraulic system (1).

$$P_{loss} = P_{in} \cdot (1 - \eta_{total}) \tag{1}$$

Energy losses in hydraulic systems result in the dissipation of heat energy to the hydraulic fluid. This leads to an increase in fluid temperature and corresponds to a decrease in viscosity. Based on the hydraulic system of mobile machines, energy losses in the system can be assigned to corresponding system components. Mainly, the losses occur in pumps/motors, valves and pipe/hose lines. Further losses occur, for example, in coolers or filters. The energy losses appear in superimposed form during the operation of the machine.

Hydraulic-mechanical losses

Hydraulic-mechanical losses occur in the form of friction losses in tribological contacts of hydraulic components. In these contacts, such as the piston/bushing contact of an axial piston pump, the separation of the surfaces of contact partners moving relative to each other is achieved by a hydrodynamic lubricating film. This lubricating film bears the load arising at the contact and separates the friction partners, in the best case, completely from each other. If the viscosity of the lubricants drops to such an extent that the separation of the surfaces can no longer be guaranteed, the load-bearing lubricating film can no longer be formed, and the contact is driven in boundary lubricating conditions. The influence of varying pressure-viscosity behavior of different base oils on the

formation of load-bearing lubricating films and the friction conditions in tribological contacts has been experimentally investigated in [15]. Consequently, friction in hydraulic-mechanical losses increase the lower the dependency of pressure on fluids viscosity.

Volumetric losses

Volumetric losses are defined in terms of volumetric loss flow that occurs functionally in hydraulic displacement units. Compression losses as well as internal and external leakage lead to volumetric losses. In hydraulic pumps and motors, internal leakage occurs because of a compensating volumetric flow from the high-pressure to a low-pressure side in function-related gaps. These volumetric losses occur for example, in the annular gap of a piston/bushing contact of an axial piston pump. The flow in this region is mostly laminar and can be described by Hagen-Poiseuille´s law [16]. Hagen-Poiseuille is an equation that connects the pressure drop and the volumetric flow rate with the hydraulic resistance. With an increase in viscosity, the volumetric losses decrease.

2.2. Hydraulic system losses

The architecture of a hydraulic system has a decisive influence on the efficiency of a mobile machine. In [7], the occurring losses during a 90° dig and dump cycle using a crawler excavator were determined experimentally. The losses of the individual hydraulic system components were considered based on the energy input of the fuel. 20 digging cycles were carried out by a test driver and the measurements were averaged. It was shown that about 7.4 % of the input energy could be converted into useful energy for the operation. Around a quarter of the losses can be attributed to the hydraulic system. In addition to system-related losses due to control tasks (such as load-sensing architecture), losses occur in flow guides like hoses, valves and pumps as well as motors. These losses are dependent on the viscosity of the fluid used. The influence of viscosity on different types of losses is possible due to the complexity and overlapping of losses by forming balance boundaries and will be examined in more detail in this paper. For this purpose, different aspects of the viscosity behavior of different base oils as well as different temperature-viscosity behavior will be considered.

2.3. Viscosity behaviour of different base oils

Viscosity generally describes the internal resistance against the shearing of a liquid. It is used as a measure of the internal friction of a liquid. This internal friction occurs between the molecules of a flowing medium [17]. Molecular chains slide on top of each other due to an initiated force.
The viscosity of lubricants generally depends on various parameters such as shear, temperature, and pressure. For mineral oils and synthetic oils with comparable molecular dimensions, Newtonian behaviour of the fluids can be assumed for relevant shear rates in hydraulic systems. Therefore, viscosity is shear-independent. The investigations done for this paper focus on the viscosity-temperature behaviour and viscosity-pressure behaviour.
Lubricants viscosity generally decreases with increasing temperature. The amount of the influence of temperature on viscosity is numerically described by the so-called viscosity index (VI). The higher the VI the lower the influence of temperature change on the lubricant’s viscosity. A temperature-stable viscosity favours the prediction of lubricant behaviour in systems concerning different operation temperatures. Furthermore, it lowers the fluid-dependent frictional losses for example while starting a system at low ambient temperature. The viscosity dependency of lubricant primary can be referred to the chemical structure of the base oil. Long carbon chains with only a few branches lie very near to each other and interfere with each other if the temperature rises and the molecules oscillate (e.g., synthetic oils and water-glycol). Otherwise, in mineral oil based lubricants the gaps between the chains are larger because of for example aromatic cyclic bounds. Thus, there are few contact points between the molecules since the hydrocarbon chains are spatially separated from each

other by cyclic structures. This gap increases as the temperature rises and the chains straighten. Therefore, the VI of synthetic oils and water-based fluids is higher than that of mineral oil based fluids.
By adding so-called viscosity improvers, the VI of mineral oils (HVLP [18]) can be brought to the same value as that of esters or polyalphaolefins. For this purpose, used polymers are formed as a coil at low temperatures and by rising temperatures they expand and fill the gaps between the hydrocarbon chains. In contrast to the base oils, the long chains are prone to shear stress. The chains can be cut, and as a result, they lose their viscosity-improving effect. Similar behaviour can be seen at HFC [19] fluids where polymers are used to achieve the needed viscosity. [20, 21]
The pressure-viscosity dependency is described by the pressure viscosity coefficient α. Pressure-viscosity dependency can be explained by taking a closer look at the chemical structure of base fluids. The more side branches fluids have the more the viscosity depends on pressure. If the (hydrocarbon) chains are pressed closer together the branches interlock with each other and build a further resistance against sliding between the chains. As already mentioned, mineral oils consist of a more branched structure like synthetic and water-based fluids. For this reason, their viscosity dependency on pressure is higher. The pressure dependency is important for the formation of load-bearing lubricant films in tribological contacts, for example in hydraulic pumps. Due to a high dependency of lubricant viscosity on pressure, the fluid can form more stable lubricating films. The pressure-dependent viscosity η_p can be explained by the exponential approach of Barus (2) [22]. The viscosity η_0 is obtained at a defined temperature under atmospheric pressure.

$$\eta_p = \eta_0 \cdot \exp(\alpha \cdot p) \tag{2}$$

3. TESTING SETUP

In the following chapter, the testing setup of the research is presented. Therefore, the crawler excavator and its hydraulic system are explained and followed by the integration of the hydraulic unit that drives the excavator. The sensor accuracies are shown to classify the reliability of the results. Next, the testing cycle is shown. After that, the investigated hydraulic fluids and their relevant properties for this study are explained.

3.1. Testing Excavator

The crawler excavator belongs to the class of compact excavators, and it is driven by tracks. In addition to linear actuators, which are required for digging, there also are actuators for further operations which are not regarded in the investigations. The test excavator is controlled by electro-hydraulic pilot valves instead of the usually implemented hydraulic joysticks. The pump of the excavator is pressure-controlled. The total weight of the machine is 1,790 kg. The gross power of the diesel engine is 12 kW. The nominal bucket filling level is 36 liters. The hydraulic tank consists of 15 liters and its complete hydraulic system capacity is 21 liters. The test excavator is equipped with appropriate measurement technology to record relevant data. The right side of **Figure 1** shows the main circuit of the hydraulic system with installed sensors. The main circuit consists of three linear actuators which are the boom (9), arm (8), and bucket (6) cylinders. The swivel motor (1) correspondingly drives the slewing gear for swiveling the upper carriage. The diesel engine drives the main hydraulic pump. Other actuators, such as the track motor, are not considered in this study because they are not activated during the executed test cycle.
Regarding the existing hydraulic drive system of the excavator (black box), the transferred power of the combustion engine to the shaft of the hydraulic pump can only be evaluated via fuel consumption. This measurement method only allows limited statements about the mechanical input power of the hydraulic pump. The reason for that is unknown parameters, such as the efficiency of the combustion

engine. Therefore, the conventional drive system of the excavator (black box) was replaced and an external hydraulic power supply (red box) was adapted to drive the excavator. The circuit diagram of the combined hydraulic system consisting of the hydraulic unit and excavator's hydraulic system (blue box) is shown in **Figure 1**. The pressure line of the power unit is connected to the hydraulic system of the excavator immediately after the pump of the excavator. The return line is discharged in front of the excavator´s internal cooler and filter. The throttle is needed to regulate the pressure of the pump. To calculate the input energy of the electric motor torque and rotation speed is measured. Furthermore, the pressure and temperature of the pressure line are identified as well as the leakage oil temperature and tank temperature. The tank temperature can be kept constant at the set value via a controlled air cooler. Heating is done while throttling the fluid by a pressure relief valve.

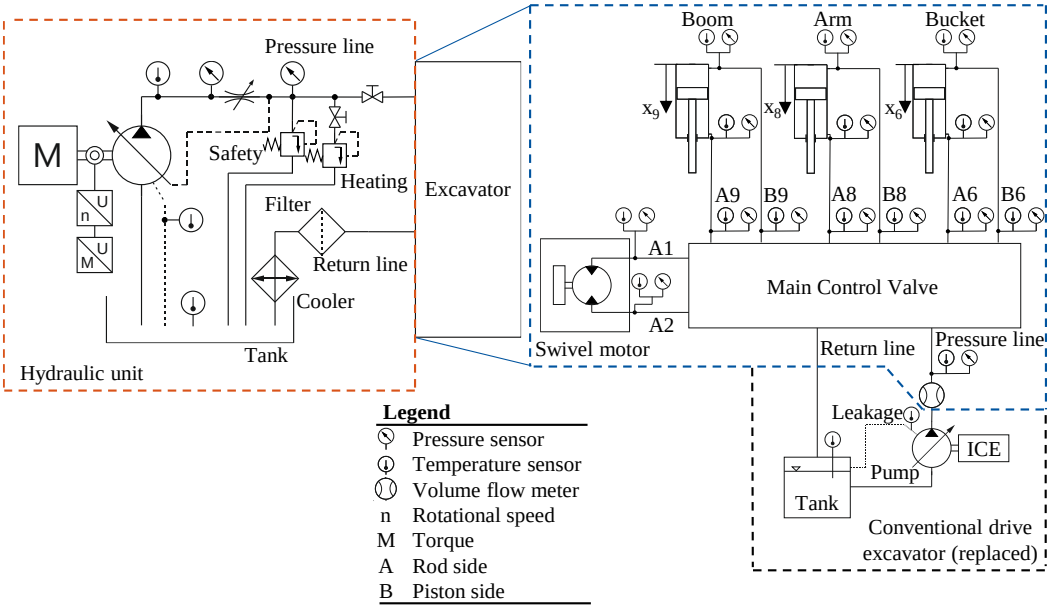


Figure 1: Hydraulic system excavator and hydraulic unit

3.2. Accuracy

In **Table 1** the measurement range and corresponding accuracy of the sensors used are shown. The accuracy refers to the measurement range of the sensors. The sensors' positioning in the excavator's hydraulic system is shown in **Figure 1**. For temperature measurement thermocouples of type K according to DIN EN 60584-1 [23] are used whose measuring range according to class 1 is -40 °C to 1,000 °C. Referring to the standard, the permissible limit deviation is 1 °C. All sensors are calibrated concerning their entire measurement chain regarding their measurement range during application.

Table 1: Sensor accuracy

Sensor type	Range	Accuracy
Pressure	0 … 250 bar	+/- 0.5 %
Volume flow	1 … 150 l/min	+/- 0.2 %
Torque	0 … 500 Nm	+/- 0.03 %
Rotational speed	0 … 500 rpm	+/- 0.04 %
Length	0 … 2,000 mm	+/- 0.001 %

3.3. Automated dig and dump cycle

The typical construction site activity of an excavator consists of the sum of different load cycles. The main task is the load cycle “digging” with 40 % [24]. The executed test cycle is based on test cycles defined by the Japan Construction Mechanization Association (JCMA) [25] for determining the energy consumption and energy losses of hydraulic excavators. A 90-degree dig and dump cycle is performed as a test cycle. The duration of the cycle is about 58 seconds.
The execution of the cycle is shown in **Figure 2**, and the time concerning the start of the cycle is also described. Instead of digging material, a weight (25 kg) is attached to the bucket and moved during the digging cycle. The weight does not leave the bucket during the cycle. Beginning from the start position, the excavator performs a digging movement after which it is guided back to the starting position (0-25 sec). The upper carriage is then swivelled by 90° (25-32 sec) and the excavator performs an emptying movement (32-44 sec). Finally, the arm is moved back to the starting position and the upper carriage is swivelled back (44-end). The whole movement is automated. An electro-hydraulic prototype control allows completely automated and reproducible digging cycles. For this purpose, a dig and dump cycle was run manually and the respective cylinder lengths as well as the upper carriage rotation were recorded. These data were stored in the control system in a look-up table, which contains the respective lengths of the actuators over time. After starting the implemented automatic dig and dump program, the lengths are fed into a controller, which actuates the slide position of the valves belonging to the actuators. With this automatization, the cylinder lengths and the rotational angle are controlled. For this, reason the output energy is comparable for different fluids. Thus, it is possible to perform a completely automated and reproducible 90° dig and dump cycle allowing to compare different fluids with various physical properties regarding their performance.

Section	Process
1	Digging movement 0 - 25 sec
2	Rotation upper carriage 25 - 32 sec
3	Unload movement 32 - 44 sec
4	Rotation to starting position 44 sec - end

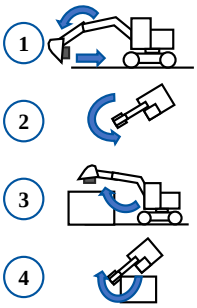


Figure 2: Dig and dump cycle

3.4. Hydraulic fluids

Three hydraulic fluids with different physical properties are part of the investigations. The most important properties regarding the investigations are shown in **Table 2**.
The first hydraulic fluid studied is a mineral oil based hydraulic oil with a viscosity class of 46. This is a frequently used hydraulic oil in various industries and mobile applications. Therefore, in this study, the HLP [26] is used as a reference fluid concerning the results. The second fluid is a mineral oil based fluid with the same paraffinic base oil type as the HLP, but it is concentrated with viscosity improvers (HVLP [18]). The lower viscosity class of 32 is chosen to specify if it is possible to reduce the energy consumption of the mobile machinery when reducing the overall viscosity of the hydraulic fluid. The higher VI is necessary to fulfil the minimum viscosity requirements of the pump at high temperatures. The third hydraulic fluid is based on a mixture of water and glycol (HFC [19]); its viscosity class is 46. Due to the high measurement effort related to the investigations in the study,

only three fluids could be investigated. The water content of the investigated HFC is about 40 %. The water-based hydraulic fluid is chosen as a representative of alternative and sustainable hydraulic fluids. In comparison to mineral oils, it reduces the impact of fossil-based carbon on the environment. Furthermore, it is over 90 % biodegradable within 28 days (OECD 301 [27]) and reduces the environmental impact of spilling hydraulic fluids. Accidents of mobile machinery play a decisive role concerning hydraulic oil spills [28]. Additionally, the HFC fluid has a high VI comparable to those of other sustainable biobased oils (e.g., ester).
For changing the fluids, a flushing process was done. The previously used fluid was drained, and the system was filled with the fluid to be tested. The actuators of the excavator were then moved to flush residual oil out of the system. This procedure was repeated until the hydraulic fluid represented the required purity demanded by ISO 15380 [29].

Table 2: Important properties tested hydraulic fluids [20, 21, 30, 31]

Hydraulic fluid	Viscosity class [32]	Viscosity index [33]	Density [g/cm³] (20 °C)	Specific heat capacity [kJ/(kg K)]	Thermal conductivity [W/(m K)]	Pressure Viscosity Coefficient [(GPa)$^{-1}$] (25 °C)
HLP	46	113	0.86	1.9–2.2	0.13–0.14	20.0
HVLP	32	195	0.85	1.9–2.2	0.13–0.14	20.0
HFC	46	206	1.08	3.3	0.3–0.43	2.0

Figure 3 shows the dependence of the hydraulic fluids' dynamic viscosity considered on temperature (a) and pressure (b). The different viscosity indexes of the fluids result in an intersection for HLP and HVLP, which is at about 100 °C outside the area relevant to this study. Between HFC and HLP, there is an intersection at 22 °C and 97 mPas. The viscosity curve of HFC ends at 60 °C because this is the limit of the application range of the water-based fluid. When considering the pressure dependence of the viscosity (b), the relatively low viscosity increase of HFC compared to the mineral oil can be identified, which is particularly important for high pressures in tribological contacts. At low pressures (>200 bar), the resulting deviations in viscosity are relatively small.

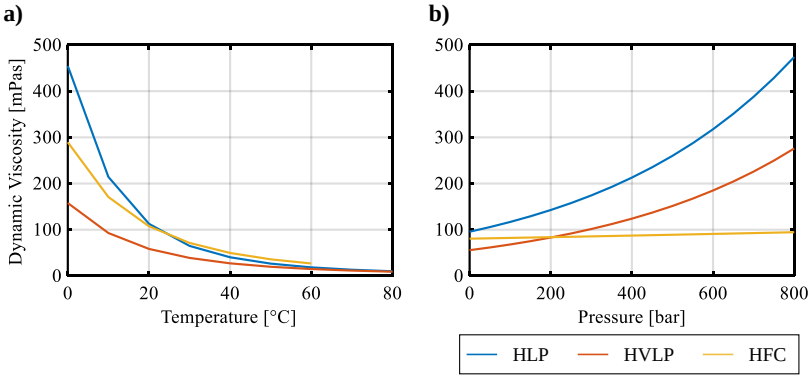


Figure 3: a) Temperature viscosity behaviour and b) Pressure viscosity behaviour investigated fluids

3.5. Testing temperatures

Tests were carried out at three different tank temperatures (**Table 3**). Due to different tank temperatures, it is possible to determine the influence of different viscosities during the operation on the efficiency of the excavator. The first stage is at a tank temperature of 25 °C ± 2 °C to simulate the losses that occur during a cold start of the machine or in operation at low ambient temperatures. The

second stage is at 40 °C ± 2 °C. The third is when the tank has been warmed up and a nearly stationary temperature is reached. Under normal loads, a steady-state tank temperature of 50 °C ± 2 °C is reached during operation at ambient temperatures of 20 °C ± 2 °C.
To achieve statistical accuracy, all measurement values were averaged over 50 cycles at each tank temperature stage for each oil.

Table 3: Testing temperatures

Temperature stage	Range
1	25 °C ± 2 °C
2	40 °C ± 2 °C
3	50 °C± 2 °C

4. MEASUREMENT RESULTS AND DISCUSSION

In this chapter, the measurement results of the excavator tests are shown. After that results of important measurements regarding the comparison of the different hydraulic fluids considered in the study are presented. First, the temperatures at important measurement locations of the hydraulic system are shown. To state the total energy consumption of the different hydraulic fluids the torque measurements are established as well as the presentation of energy consumption during the dig and dump cycle. All measurement data shown refer to the arithmetic average of 50 dig and dump cycles performed for each fluid at every temperature stage (Table 3). Overall, for each fluid, 150 cycles were performed.

4.1. Temperatures

The fluid temperatures during the operation of the excavator depend on the losses in the hydraulic system. The absorption and storage of heat energy by the lubricant are generally dependent on its thermal capacity.
In **Figure 4**, the tank temperature (a) and the leakage oil temperature of the main pump (b) are plotted over the cycle time. The plots are divided into three different temperature stages shown in **Table 3**.
The first row (a) shows the tank temperature over the cycle time. The maximum deviation of ± 2 °C from the respective reference temperature dependent on the stage can be observed. At stages 2 and 3, the tank temperatures are very similar (ΔT_{max}=0.2 °C). Whereas for temperature stage 1 the maximum deviation is 2 °C between HLP and HFC. At this temperature stage, cooling is done by adding an external heat exchanger to the system. The cooling of the coolant is done externally, and its temperature is not held at a constant value. Therefore, the coolant temperatures may differ between the measurements. Nevertheless, the deviations shown are within the previously defined tolerance range (**Table 3**).
The leakage oil temperature (b) indicates the heat development in the pump due to friction in tribological contacts and leakage volume flow. According to the dig and dump cycle progression (**Figure 2**), it can be observed that the maximum temperature increase takes place during the digging process. Especially during the relatively slow movements of digging and the retraction of the bucket, considering the whole cycle, a significant increase in temperature and a temperature peak occurs (t=21 s). This peak is more significant at stages 2 and 3 than at stage 1. This can be described by the increasing leakage flows with increasing temperature and decreasing fluid viscosity.
When considering the temperature difference between the three different hydraulic fluids, it can be noticed that the water-based HFC shows a lower leakage oil temperature than the mineral oil based fluids (HLP, HVLP). This trend can be seen over the entire cycle time and for all temperature stages.

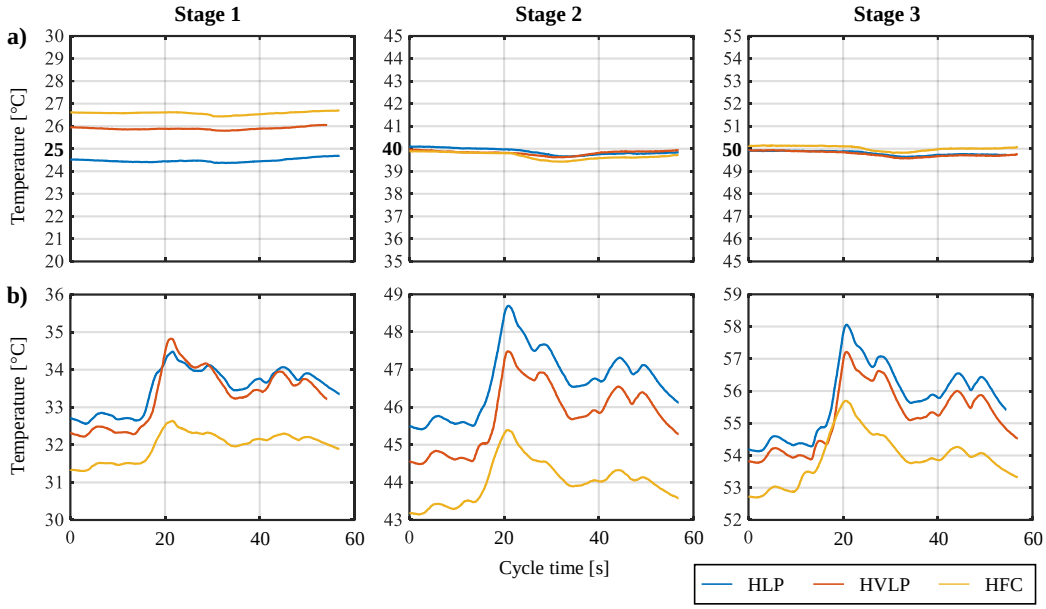


Figure 4: Temperature behaviour over cycle time. a) Tank temperature b) Leakage oil temperature

4.2. Torque and energy consumption

The torque is measured at the shaft between the electric motor and the hydraulic pump. Due to the constant rotational speed (1,500 rpm), it is proportional to the input power. Therefore, the torque can be used to determine, via the cycle time at which operation of the dig and dump cycle the corresponding amount of power is needed. **Figure 5** shows the torque curve over the cycle time for the different temperature stages and hydraulic fluids. In general, it can be concluded that the torque decreases for all fluids as the input temperature increases and the viscosity consequently decreases. Furthermore, the torque when using HFC is higher than when using the mineral oil based fluids at all temperature stages over nearly the entire dig and dump cycle. This is even though the viscosity at stage 1 is lower than that of the HLP (**Table 2**). One exception is the section when the excavator is performing a digging process (t=18-20 s). The low cylinder speeds in this process result in a lower torque when using HFC than with the mineral oil based fluids. An explanation for this behaviour is part of current investigations.

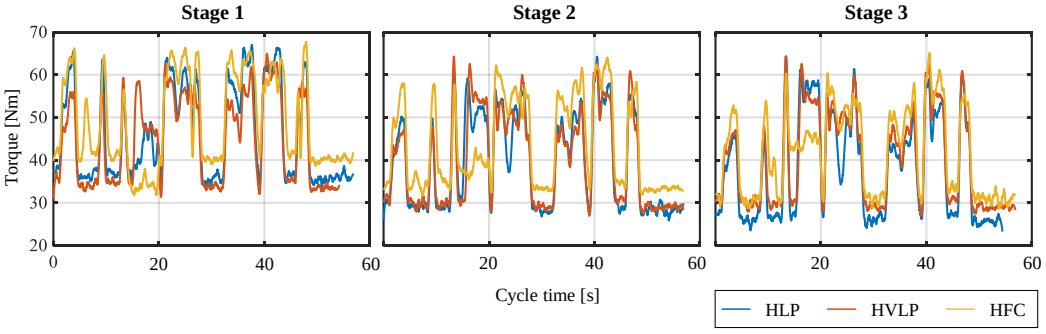


Figure 5: Input torque of different temperature stages

By integrating the power from the multiplication of torque and speed over the cycle time, the energy required to perform the dig and dump cycle can be determined. **Figure 6** shows the energy consumption per average of 50 cycles in kilojoules. The standard deviations for energy and temperature are shown for 50 cycles. The data points between the measured temperature stages are identified by linear interpolation.

The plotted measurement data show that the energy requirement to fulfill the dig and dump cycle is higher when using HFC than when using mineral oil based hydraulic fluids. Furthermore, an intersection between HLP and HVLP can be seen, which can be explained by the different viscosity-temperature behaviour of the oils. In addition, the energy drop over temperature is lower for the two fluids with high VI than for HLP. When considering the standard deviations for energy, these are higher for HLP, than for the other fluids considered. Since the standard deviations for temperature are comparable for all three, it is assumed that this can be linked to the lower VI since the deviations in viscosity are higher in this respect at slightly fluctuating temperatures. At the first temperature stage, the standard deviations for temperature are smaller than for the other temperature stages, this is due to the additional constant cooling power of the additional heat exchanger (see 4.1).

The higher energy requirement when using HFC can be described on the one hand by higher flow losses caused by its higher dynamic viscosity compared to the mineral oil according to the Hagen-Poiseuille law (see chapter 2.1). On the other hand, due to the lower dependence of viscosity on pressure, higher friction losses may occur in tribological contacts, thus increasing the hydraulic mechanical losses. To qualify this assumption, further qualitative investigations are necessary regarding the losses occurring in the pump.

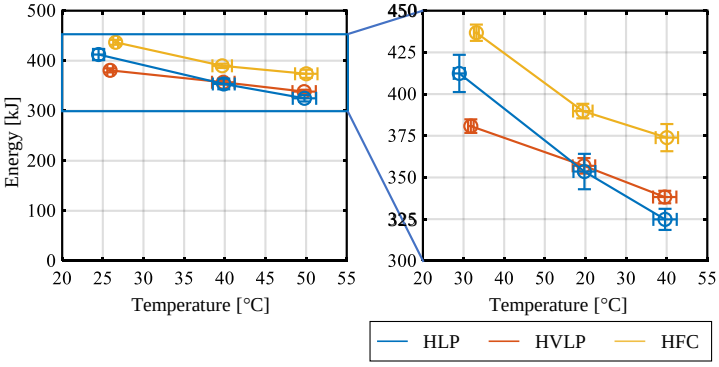


Figure 6: Energy consumption dig and dump cycle

5. CONCLUSION AND OUTLOOK

The research results of this study show different effects in terms of efficiency as well as temperature distribution when using different commercial hydraulic fluids in the excavator. The water-based HFC shows lower temperatures in the leakage oil port of the pump. Explanations can be referred to different thermal properties or a lower leakage oil flow concerning reduced volumetric losses. Further investigations to prove these assumptions are part of current investigations. Regarding energy consumption, the viscosity dependence between HLP 46 and HVLP 32 can be identified. In the low-temperature range (25-40 °C), the energy requirement is below that of HLP, with an intersection at around 40 °C, whereby in the range from 40 °C up to 50 °C, HLP has the lowest energy requirement. The intersection of the kinematic viscosities of the two oils, however, is at a significantly higher temperature of around 100 °C. The energy requirement when using HFC is significantly higher than when using HLP (Stage 1: 5.5 %/Stage 2: 9.3 %/Stage 3: 13.1 %) and HVLP

(Stage 1: 12.8 %/Stage 2: 8.4 %/Stage 3: 9.5 %) at all temperature stages.
In summary, the simplified use of low viscous fluids in hydraulic systems does not necessarily lead to a reduction in energy consumption. This requires a more detailed understanding of the system as well as the identification of losses and the temperature balance of the system. The three commercially available hydraulic fluids show different behaviours regarding these aspects. Especially for fluids with alternative base fluids that are not mineral oil based (e.g., water, ester), the effects on different types of losses must be known here due to their different physical and chemical properties.
Future research work will deal with the subject of dividing losses occurring in the excavator's hydraulic system. Therefore, the segments of the excavator's hydraulic system will be divided into balance sheet boundaries and losses in these segments will be calculated. With this method, it will be possible to evaluate in which area of the system which number of losses occur when using different fluids at different temperature stages. Furthermore, biobased hydraulic fluids (esters) will be included in the considerations.

NOMENCLATURE

P_{loss}	Dissipation power	W
P_{in}	Input power	W
η_{total}	Total efficiency	-
η	Dynamic viscosity	mPas
p	Pressure	bar
α	Pressure-viscosity-coefficient	GPa^{-1}

REFERENCES

[1] *Paris Agreement*, **2015**.
[2] Huang, L.; Krigsvoll, G.; Johansen, F.; Liu, Y.; Zhang, X. Carbon emission of global construction sector. *Renewable and Sustainable Energy Reviews*, **2018**, *81*, 1906–1916.
[3] Vukovic, M. Hydraulic hybrid systems for excavators. Dissertation, RWTH Aachen; Shaker Verlag GmbH, **2017**.
[4] Bräckelmann, U.; Breuer, D.; Stoll, S.; Piepenstock, U., Eds. *Optimized piston guidance design for swash plate pumps*, **2010**.
[5] Shorbagy, A.; Ivantysyn, R.; Weber, J. Holistic Analysis of the Tribological Interfaces of an Axial Piston Pump – Focusing on the Pump Efficiency. *Chem Eng & Technol*, **2023**, *46*, 5–13.
[6] Ligocki, A.; Greye, B.; Meyer, F., Eds. *Ein System zur Detektierung von hydraulischen Verlustleistungen in mobilen Arbeitsmaschinen*, **2014**.
[7] Sturm, C. *Bewertung der Energieeffizienz von Antriebssystemen mobiler Arbeitsmaschinen am Beispiel Bagger;* KIT Scientific Publ: Karlsruhe, Baden, **2015**.
[8] Idelčik, I.E.; Ginevskiĭ, A.S., Eds. *Handbook of hydraulic resistance;* Begell House: Redding, Conn., **2007**.
[9] Bügener, N. *Analyse und Verbesserung des Ansaugverhaltens von Axialkolbenpumpen in Schrägscheibenbauweise;* Shaker: Aachen, **2014**.
[10] Küppers, U. *Entwicklung von energie(druckverlust)optimierten Bogen-Elementen mit nicht kreisförmigen Querschnitt nach dem Mäander-Prinzip:* Bremen, **09/2006**.
[11] Otto, N. *Erprobung nachhaltiger Hydraulikfluide auf Ester- und Wasserbasis*, 1st ed.; Shaker: Herzogenrath, **2019**.
[12] Chiavola, O.; Frattini, E.; Palmieri, F.; Fioravanti, A.; Marani, P. On the Efficiency of Mobile Hydraulic Power Packs Operating with New and Aged Eco-Friendly Fluids. *Energies*, **2023**, *16*, 5681.
[13] Deuster, S.; Schmitz, K. Bio-Based Hydraulic Fluids and the Influence of Hydraulic Oil Viscosity on the Efficiency of Mobile Machinery. *Sustainability*, **2021**, *13*, 7570.

[14] Hlawitschka, E. Zum Temperaturverhalten von Hydraulikanlagen. *Zeitschrift Agrartechnik*, **1979**, *29*, 86–88.
[15] Deuster, S.; Holzer, A.; Schmitz, K. *Influence of the Pressure-Viscosity Behavior of Different Base Oils on the Formation of Lubricating Films in Tribological Contacts (reviewed conference article);* IEEE Global Fluid Power Society PhD Symposium: Neapel, **2022**.
[16] Sutera, S.P.; Skalak, R. The History of Poiseuille's Law. *Annu. Rev. Fluid Mech.*, **1993**, *25*, 1–20.
[17] NEWTON, I. *PHILOSOPHIAE NATURALIS PRINCIPIA MATHEMATICA (LATIN,1687);* BENEDICTION CLASSICS: [S.l.], **2015**.
[18] DIN Deutsches Institut für Normung e.V. *Hydrauliköle: Teil 3: Hydrauliköle HVLP, Mindestanforderungen;* Beuth Verlag GmbH: Berlin, **2017**.
[19] DIN Deutsches Institut für Normung e.V. *Familie H (Hydraulische Systeme): Anforderungen an Druckflüssigkeiten in den Kategorien HFAE, HFAS, HFB, HFC, HFDR und HFDU;* Beuth Verlag GmbH: Berlin, **2020**.
[20] Mang, T.; Dresel, W., Eds. *Lubricants and lubrication;* Wiley-VCH: Weinheim, Germany, **2017**.
[21] Totten, G.E.; Shah, R.J.; Forester, D.R., Eds. *Fuels and Lubricants Handbook: Technology, Properties, Performance, and Testing, 2nd Edition;* ASTM International: 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, **2019**.
[22] Barus, C. Isothermals, isopiestics and isometrics relative to viscosity. *American Journal of Science*, **1893**, *s3-45*, 87–96.
[23] DIN Deutsches Institut für Normung e.V. *Thermoelemente: Teil 1: Thermospannungen und Grenzabweichungen;* Beuth Verlag GmbH: Berlin, **2014 (Accessed** September 25, 2023).
[24] Helmus, M.; Fecke, M. *Standardisierung definierter Lastzyklen und Messmethoden zur Energieverbrauchsermittlung von Baumaschinen: Schlussbericht zum Forschungsvorhaben*, **2015**.
[25] JCMA - Japan Construction Mechanization Association Standards Committee. JCMAS H 020:2007 - (Japan Construction Mechanization Association Standard): Earth-moving machinery – Fuel consumption on hydraulic excavator – Test procedures. http://www.cema.or.jp/general/news/pdf/17th/17thJapanese (Accessed December 18, 2020).
[26] DIN Deutsches Institut für Normung e.V. *Hydrauliköle: Teil 2: Hydrauliköle HLP, Mindestanforderungen;* Beuth Verlag GmbH: Berlin, **2017**.
[27] OECD Organisation for Economic Co-operation and Development. Test No. 301: Ready Biodegradability.
[28] Guerin, T. Root causes of fluid spills from earthmoving plant and equipment: Implications for reducing environmental and safety impacts. *Engineering Failure Analysis*, **2014**, *45*, 128–141.
[29] DIN Deutsches Institut für Normung e.V. *Schmierstoffe, Industrieöle und verwandte Produkte (Klasse L) – Familie H (Hydraulische Systeme) – Anforderungen für die Kategorien HETG, HEPG, HEES und HEPR (ISO 15380:2011);* Beuth Verlag: Berlin, **03.2012**.
[30] Cusatis, P.; Sherman, J.; Fasano, P.; Bishop, R. Property and Performance Evaluation of Water Glycol Hydraulic Fluids. In: *Fire resistant fluids*. Sherman, J., Ed.; ASTM International: Conshohocken, PA, **2014**; pp. 109–125.
[31] Wan, G.T.Y.; Spikes, H.A. The Elastohydrodynamic Lubricating Properties of Water-Polyglycol Fire-Resistant Fluids. *A S L E Transactions*, **1984**, *27*, 366–372.
[32] DIN Deutsches Institut für Normung e.V. *Flüssige Industrie-Schmierstoffe – ISO-Viskositätsklassifikation (ISO 3448:1992);* Beuth Verlag GmbH: Berlin, **2010**.
[33] DIN Deutsches Institut für Normung e.V. *Mineralölerzeugnisse Berechnung des Viskositätsindex aus der kinematischen Viskosität;* Beuth Verlag GmbH: Berlin, **2004 (Accessed** May 15, 2021).

TRIBOLOGICAL PROPERTIES OF HYDRAULIC CYLINDER PISTON SEALINGS IN WATER AND OIL HYDRAULICS

Franc Majdič[1]*, Nejc Novak[1], Jan Pustavrh[1], Blaž Laznik[1], Ana Trajkovski[1]

[1]*University of Ljubljana / Faculty of Mechanical Engineering, Aškerčeva 6, SI-1000 Ljubljana, Slovenia, EU*

* Corresponding author: Tel.: +386 1 4771 411; E-mail address: franc.majdic@fs.uni-lj.si

ABSTRACT

Hydraulics is indispensable in everyday life because of its ability to work with large forces. Hydraulic cylinders are an important hydraulic element and contain different types of seals. This work presents an investigation of the influence of three different shapes and materials of piston seals in a hydraulic cylinder on the friction force. The friction was tested in two hydraulic fluids, water and mineral oil, at six different piston rod travel speeds on a test rig prepared for this purpose. It was found that the proportion of the friction force in mineral oil is between 60 % and 85 % of the friction force generated by the same piston seal in water. However, a long-term test carried out on a test rig provided data on the leakage of the piston seal before the start of the long-term test, after 50 km of travel, and after 100 km of travel of the piston rod at full load.

Keywords: hydraulic, cylinder, piston seals, friction, mineral oil, water, long-term test

1. INTRODUCTION

Among the most important components of the hydraulic cylinder, which make motion of parts of machines, are the seals, whose task is to prevent the undesirable flow of fluids between the chambers and to prevent the penetration of hard particles. Thus, the main task of the seals is to ensure a sufficiently large sealing force, while the frictional force resulting from the sealing force is also important. There are two types of seals, static and dynamic seals. Static or stationary seals seal the gap between two parts facing each other, while dynamic seals seal the gap between two parts moving toward each other. In this study, dynamic seals in hydraulic cylinders are discussed. In seals friction causes heat generation, leakage and wear. Seals are required to withstand high pressures and long life at competitive price [1].

Yanada et. al. [2] experimentally investigated the behaviour of hydraulic actuators controlled by servo valves with respect to different sealing materials, different external load amplitudes, and different cylinder dimensions. Friction tests were conducted to determine a more accurate friction model. They developed a more accurate dynamic friction model that includes fluid friction dynamics. They demonstrated that hydraulic actuators with HNBR seal material produce a higher frictional force than those with FKM and NBR material. Moreover, they proved that the frictional force exhibits the characteristic Stribeck effect.In their research, Márton et. al. [3] presented a practical method for measuring the friction of seals in hydraulic cylinders. The method is based on measuring the chamber pressures in cylinder and measuring the velocity of piston rings. The friction force is measured at a constant speed of travel of the piston rod. If we have a system with no external load and the piston moves at a constant speed, the friction force can only be calculated from the pressure difference.
In the work of Ottestad and colleagues [4], a method for reducing static friction ("stick-slip" effect) between the piston seals and the cylinder tube is presented. Static friction is reduced by introducing

relative rotation of the piston to the hydraulic cylinder tube when the piston is at rest. The relative rotation of the piston is achieved by additional cylinders, which through the mechanism ensure the rotation of the piston rod and thus the piston. The additional cylinders are controlled by a servo valve that ensures extremely high switching speeds. The results show that the static friction is reduced in a small interval around zero speed. Thus, outside this interval, the usual Coulomb friction is present.

In the work of G. K. Nikas [5] and G. K. Nikas et al. [6], an investigation was carried out (experimental setup and mathematical model) of elastomer seals with rectangular piston rod cross-section was carried out. The operating pressure range of the oil is between 3.4 and 34.5 MPa, and the temperature range is between -54 °C and +135 °C. They found that the frictional force was 27% to 53% higher when the piston rod was pulled out than when it was pulled into the cylinder tube. This was true not only for piston seals with rectangular cross-sections, but also for piston rod U-seals. They also found that lower pressures and lower speeds lead to minimal leakage and increased friction.

In 2021, Pan et. al. [7] experimentally investigated the frictional force generated by a single seal as a function of velocity, external load, seal diameter and seal shape. The friction force was determined using the principle of measuring chamber pressures in cylinder. The resulting friction force shows a characteristic Stribeck effect when the piston rod is pulled out and retracted into the cylinder tube. Initially, the frictional force decreases with increasing speed, then the frictional force increases almost linearly with increasing speed. In the following, it was demonstrated that for the same material and shape of the seal, the frictional force increases as the diameter of the seal increase. A comparison of different seal types showed that the KGD seal type achieves the highest frictional force. In addition, the KGD seal is the least sensitive to changes in external load or pressure compared to the other two piston seal types (YB and KD). braking, which is true both when the piston is pulled out and when it is pulled into the cylinder tube.

Wang et al. [8] experimentally investigated the influence of the surface roughness of the piston rod in cylinder when sealing with an O-seal under different operating conditions. The results showed that the roughness of the piston rod surface has a great influence on the behaviour of the seal and therefore should not be ignored. The pull-out and pull-in frictional forces and leakage are increased when the roughness of the piston rod surface is greater. The resulting frictional force generates heat, which increases the temperature in cylinder.

In 2017, Ma, Wang, and Gu [9] conducted an experimental study of the friction of piston seals in cylinder as a function of different operating conditions. They found that under high-frequency operating conditions, the seal assembly T is the best. Under steady-state operating conditions, the M seal assembly is best.

In 2021, Mahankar and Dhoble [10] studied how medium and high temperatures affect the failure of hydraulic seals, which are mostly made of elastomers. They argue that it is important to have a good understanding of the failure mechanisms in order to improve seal life and reliability. All of the seal wear mechanisms studied are the result of the temperature of the seals and the temperature of the area in which the seal operates. Seal decay and performance degradation occur at moderate to high temperatures but become more noticeable at higher temperatures. Oil viscosity and elasticity of seals are strongly dependent on the operating temperature. On the process of thermal aging of seals, high pressures, improper installation of seals, chemical and thermal influences have a highly degrading and abrasive effect. The process of thermal aging affects the material, mechanical and chemical properties of seals. In many cases, two or more failure mechanisms can occur simultaneously.

In their work, Chen and colleagues [11] presented a method for determining the failure or reduction in functionality of seals in hydraulic cylinder during operation. Acoustic emission signals are caused by an internal leak in the hydraulic cylinder and technique is suitable for measuring internal leakage that is less than 1 l/min. Shanbhag et al. [12] improved the acoustic emission technique. Based on the acoustic emissions, they can distinguish between perfectly worn seals, moderately worn seals, and unworn seals.

In the work of Cristescu et al. [13], an experimental investigation of piston friction force was carried out seals in hydraulic cylinder. They tested a type of seal, the U-seal. They found that the frictional

force increases with increasing pressure. However, at high pressures, the frictional force increases more slowly.
Literature suggest multiple studies of performance of the seals and the influencing parameters such as frictional force, hydraulic fluid type, speed of the moving piston rod, temperature generated form friction and surrounding area. Motivation for this study to evaluate the performance of different type and materials of seals in water and oil, to choose the most appropriate seal for the application. New test rig was developed to investigate comparison of different seal types (KGD, YB, KD) on frictional force. A dynamic test was performed to determine the dependence of friction forces on speed and acceleration. Two different hydraulic fluid were used to assess the performance of seals and their suitably application in water and oil hydraulics.

2. EXPERIMENTAL

In the present study, we tested three different cylinder piston seals at two test rigs. On the first one we performed measurements of the frictional force of all three different piston seals, they were lubricated once with tap water and the second time with mineral hydraulic oil ISO VG 46. On the second oil-hydraulic test rig, we then conducted an endurance test for the selected piston seal made of PTFE with bronze additive.

2.1. Tested cylinder piston sealings

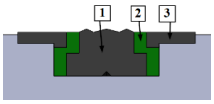
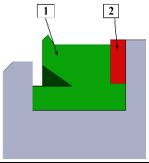
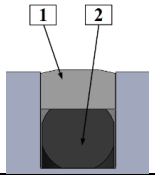
Three different types of piston seals with different geometries and made of different materials were tested, requiring different operating parameters for use (Table 1). Each seal has a different installation geometry, so three different hydraulic cylinder pistons were produced. The following materials were used for the seals: nitrile rubber (NBR), thermoplastic polyurethane (TPU) and polytetrafluoroethylene (PTFE) with bronze. The "YB50 39 4.2" seal made of PTFE with bronze is designed for maximum pressure (600 bar) and maximum sliding speed (15 m/s).

2.2. Friction measurement of piston seals lubricated once with oil and once with water

Frictional force measurements were performed on a single-axis pneumatic-hydraulic test rig (Figure 1). Using a pneumatic cylinder and a pneumatic system (Figure 2), we performed an externally controlled movement of a hydraulic cylinder piston with three various seals. We have a system with no external load and the piston moves at a constant speed, the friction force can only be calculated from the pressure difference.

Table 1: Three tested hydraulic cylinder piston seals

Type of the seal	KGD 50 34	KDA 50 39 9	YB 50 39 4.2
Material	NBR (1), TPE-E (2), POM (3)	TPU (1), POM (2)	PTFE + bronze (1), NBR (2)
Max. pressure	400 bar	500 bar	600 bar
Temp. range	-30 °C…+110 °C	-40 °C…+100 °C	-30 °C…+100 °C
Max. sliding vel.	0.5 m/s	0.5 m/s	15 m/s

Shape of the seal	(S1)	(S2)	(S3)
			

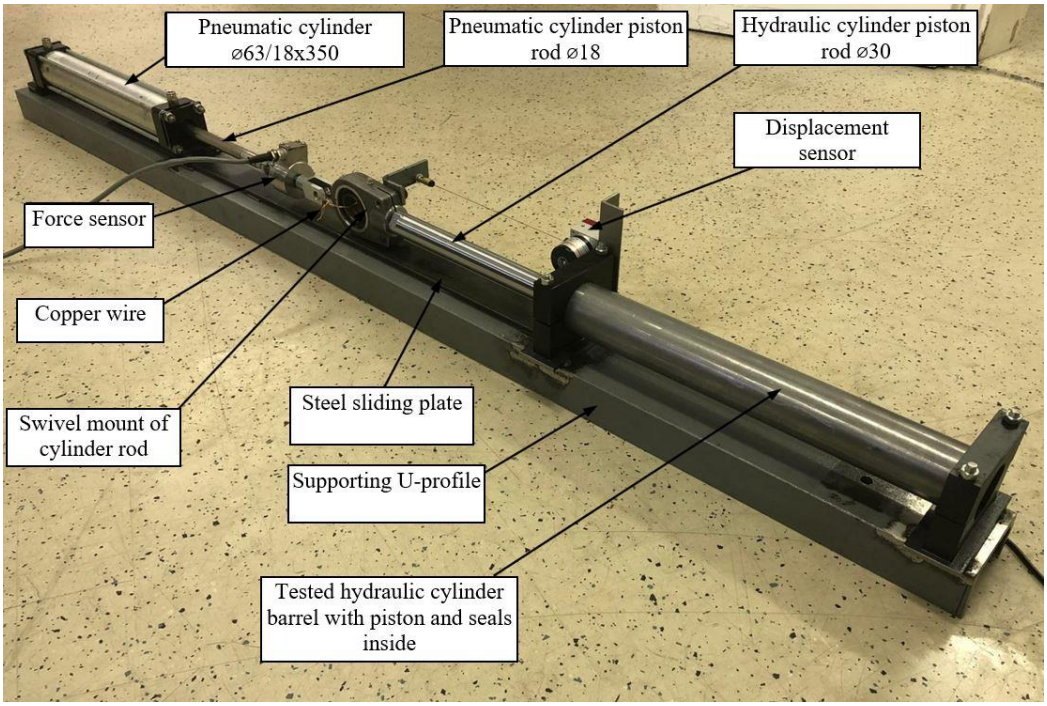


Figure 1: The first, pneumatic-hydraulic test rig - friction force measurements of piston seals (120 mm x 150 mm x 1730 mm)

The pneumatic part of the test rig (Figure 2) consists of a conditioning unit (OZ1), a 5/3-way pneumatic valve (1V1), which controls the direction of the compressed air flow passing through adjustable non-return flow-control valves (1V2 and 1V3), which indirectly control the speed and movements of the piston rod of the pneumatic cylinder (1A). The behaviour characteristics of the seals in water and in oil were measured with a force sensor (U2A) and a wire-type displacement sensor (Figure 1).

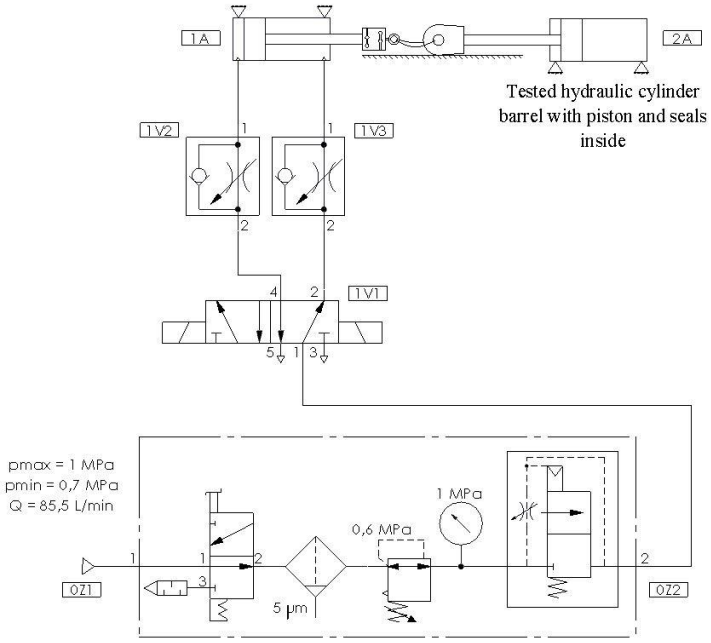


Figure 2: Functional diagram of the pneumatic test rig for measuring the frictional force of the hydraulic-cylinder piston seal

The procedure for measuring the frictional force of piston seals on the pneumatic-hydraulic test rig (Figure 1 and Figure 2 consisted of cleaning and degreasing all of components of the hydraulic test part, mounting the seal, setting the initial conditions and start measuring with Lab View.

2.3. Durability test of piston seals in oil hydraulics

The oil-hydraulic test station (Figure 3 and Figure 4) for piston seal sustainability testing consists of a hydraulic power unit with a 20 kW drive electric motor and a variable displacement axial piston pump with a displacement of 20 cm^3/rev. The pump pushes the hydraulic oil through the pressure line past the pressure relief valve through the 4/3 directional control valve alternately to the tested hydraulic cylinder ⌀50/30x380. The load on the tested cylinder is generated by the brake cylinder ⌀50/30x502. The pressure in the brake cylinder was provided alternately by two pressure valves set at 320 bar. The return line of the brake cylinder was filled with an additional hydraulic pump with a maximum filling pressure of 6 bar. The test rig also has a return line and by-pass filter as well as an oil-air cooler.

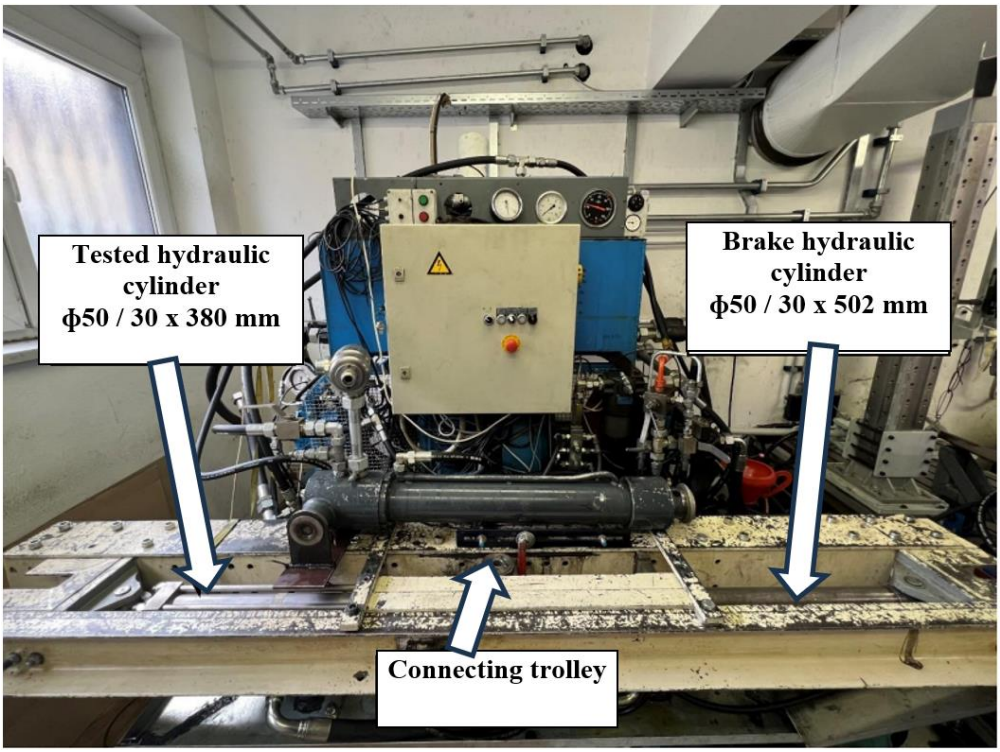


Figure 3: The second hydraulic test rig for long-term tests of cylinder piston seals (185 mm x 160 mm x 2000 mm)

The durability test, for the best piston seal from the first test rig, was performed in such a way that the Beckhoff industrial controller switched the 4/3 directional control valve, which directed the hydraulic oil alternately to one side and the other of the tested hydraulic cylinder. The load was generated by the piston rod of the tested cylinder, which was connected to the brake hydraulic cylinder. Both connections were made through two relief valves with bypass valves to the low-pressure pump, cooler and bypass filter.

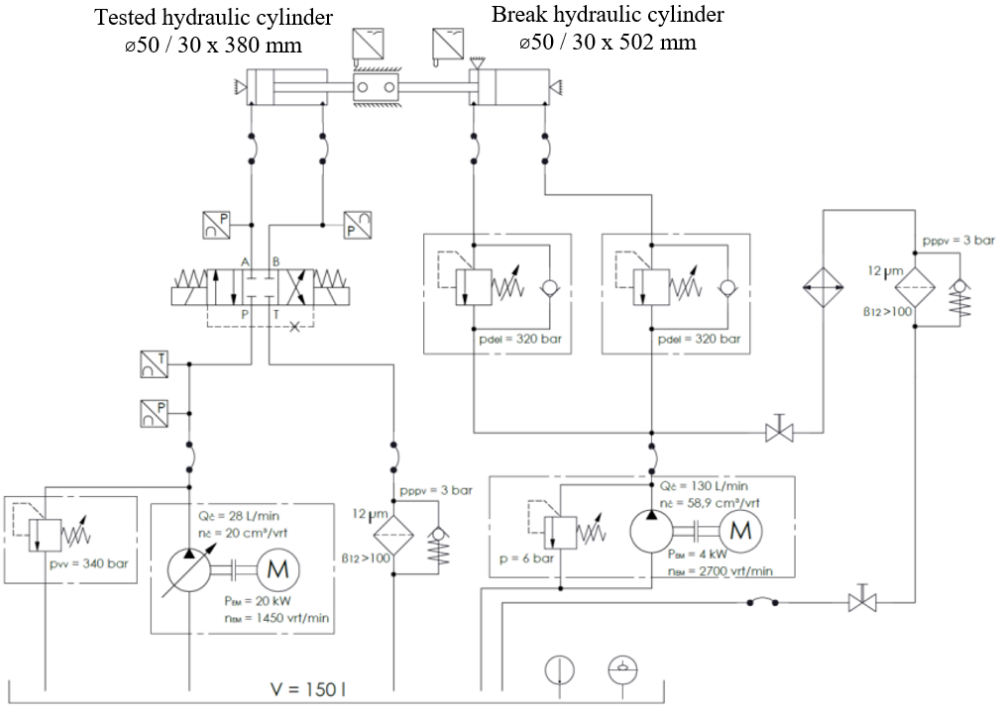


Figure 4: Functional diagram of the hydraulic test rig for long-term test of the hydraulic-cylinder piston seal

To measure the working pressure of the system, which was set to 320 bar with pressure relief valves, we used a measuring device The Parker Service Master Plus and pressure sensors (Figure 5). Then the pressure sensors are connected to both lines of the tested HV, as shown in the hydraulic diagram in Figure 4.

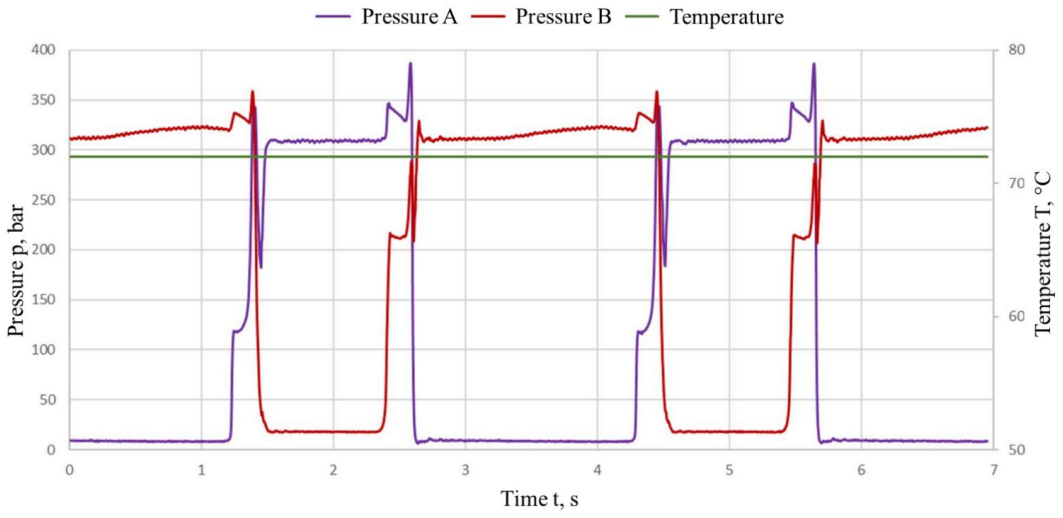


Figure 5: The measured duty cycle of the long-term test - the course of acting pressures and temperatures

3. RESULTS

3.1. Seal friction force in oil and water

The measured frictional force of the first piston seal "KGD 50 34" in mineral oil is shown in Figure 6. The analysis indicates that the highest average frictional force was observed during start-up at a speed of 270 mm/, while the lowest average frictional force occurred at a constant speed of 70 mm/s.
Transitioning to the static starting frictional force of the same piston seal in water, as depicted in Figure 7 the highest average frictional force was noted at a start-up speed of 125 mm/s, whereas the lowest average frictional force was recorded at a speed of 275 mm/s, but a constant sliding speed was not attained.
Figure 8 shows the measured frictional force of the second piston seal "KDA 50 39 9" in mineral oil. The highest average frictional force during start-up was observed at a speed of 125 mm/s, while the lowest average frictional force was measured at a constant speed of 100 mm/s.
The starting friction force and frictional force at constant speed for the second piston seal "KDA 50 39 9" in water is detailed in Figure 9. The highest average frictional force was measured at a speed of 35 mm/s when the piston seal was started up, and the lowest average frictional force occurred at a constant speed of 175 mm/s.

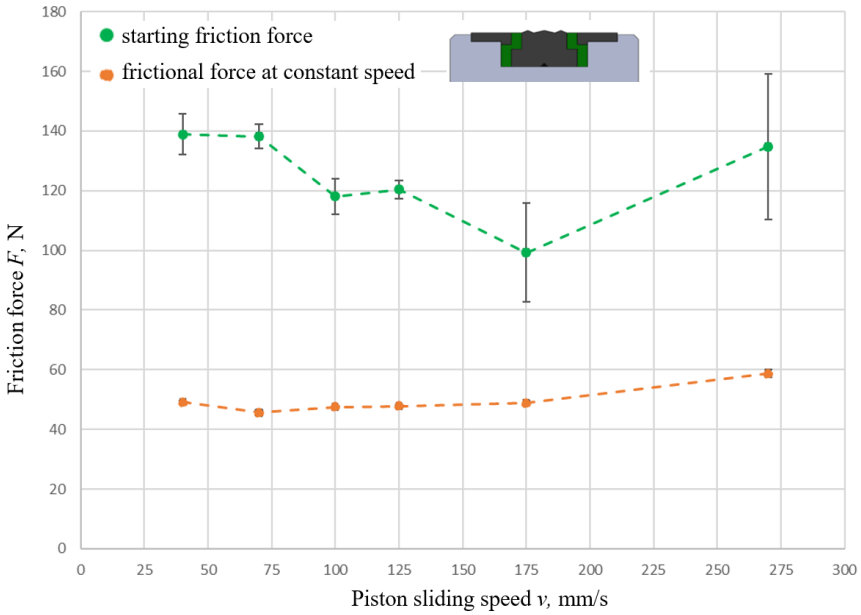


Figure 6: Measured friction force of piston seal "KGD 50 34" in mineral oil

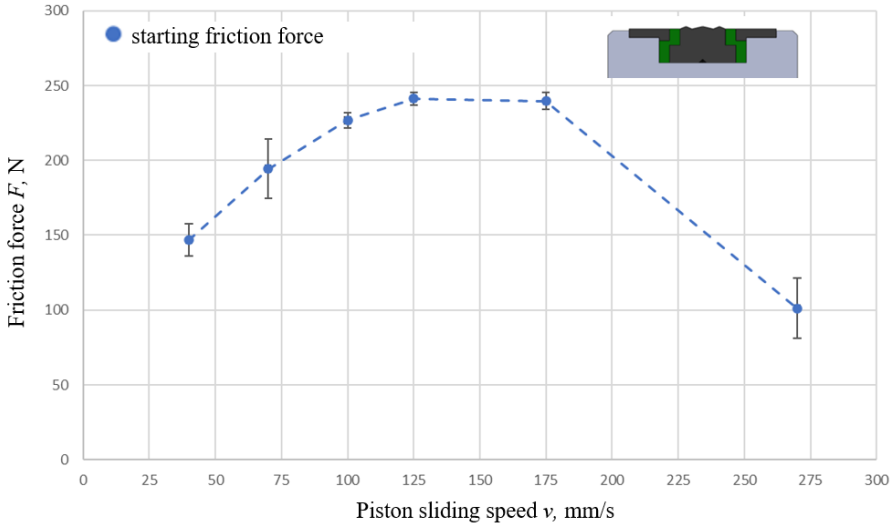


Figure 7: Measured friction force of piston seal "KGD 50 34" in water

The starting friction force and frictional force at constant speed, measured with the third piston seal "YB 50 39 4.2" in mineral oil is shown in Figure 10. The highest average frictional force was measured at a speed of 270 mm/s when the piston seal was started up, and the lowest average frictional force was measured at a constant speed of 35 mm/s.
Figure 11 shows the measured frictional force of the third piston seal " YB 50 39 4.2" in water. The highest average frictional force was measured at a speed of 270 mm/s when the piston seal was started up, and the lowest average frictional force was measured at a constant speed of 35 mm/s.

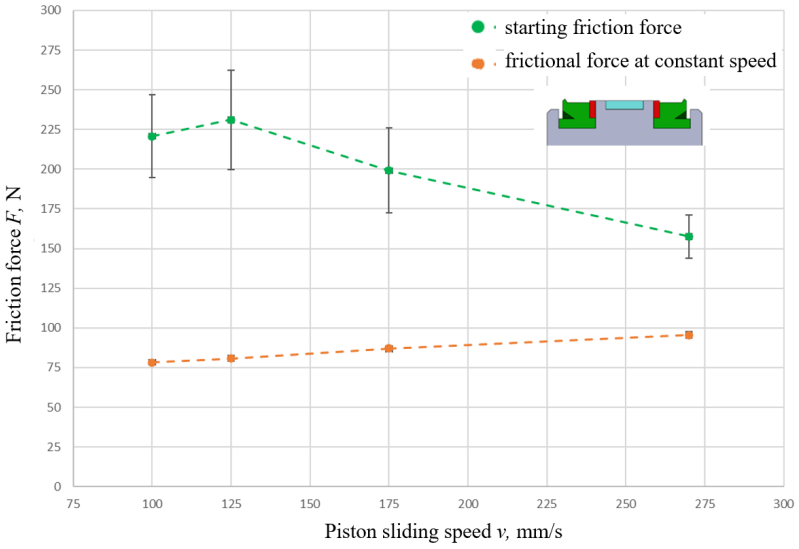


Figure 8: Measured friction force of piston seal "KDA 50 39 9" in mineral oil

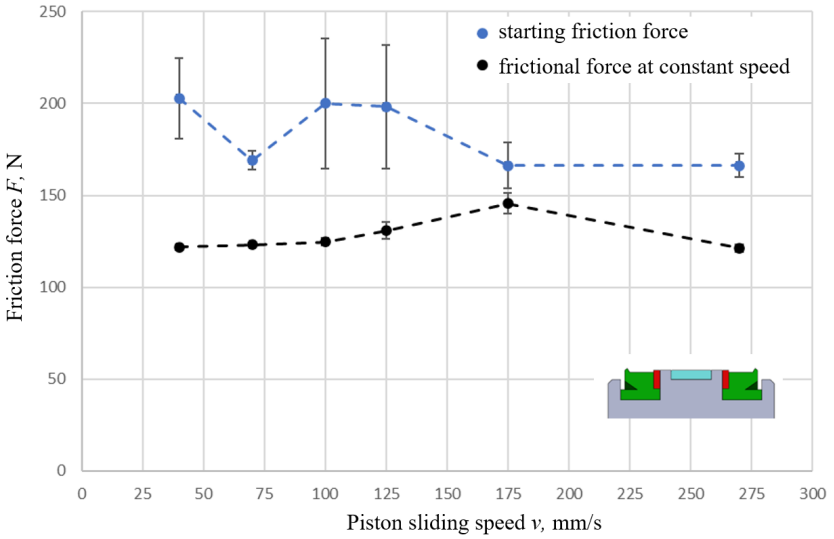


Figure 9: Measured friction force of piston seal "KDA 50 39 9" in water

With seal KDA 50 39 9, the frictional force in mineral oil is 60% of the force generated by the same type of seal in water. The seal type YB 50 39 4.2 generates in mineral oil 80% of the force generated by the same type of seal in water. However, with the KGD 50 34 water seal, the frictional force does not stabilize at a lower value because the speed does not reach a constant value. The reason for this is a greater frictional force and thus greater damping, which means that the piston sealing unit does not reach a constant speed until the piston rod is pulled out of the tube. Therefore, we have not calculated the friction ratio in mineral oil versus water for the KGD 50 34 seal type, since we have only specified the maximum value of the force. It is clear that the holding force in mineral oil is often greater than in water (seal type KPD 50 39 4.2) or the holding forces in both hydraulic fluids are the same (seal type YB 50 39 4.2). At speeds above 175 mm/s, the adhesive forces either reach equal values (KPD 50 39 4.2) or the adhesive force in water exceeds the force in mineral oil (YB 50 39 4.2). Extremal values of starting friction force and friction force at constant speed in mineral oil and water are represented in Table 2 and Table 3.

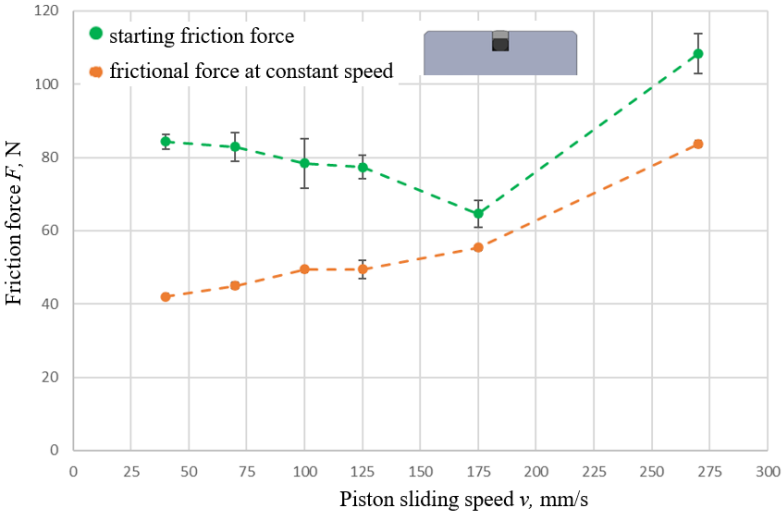


Figure 10: Measured friction force of piston seal "YB 50 39 4.2" in mineral oil

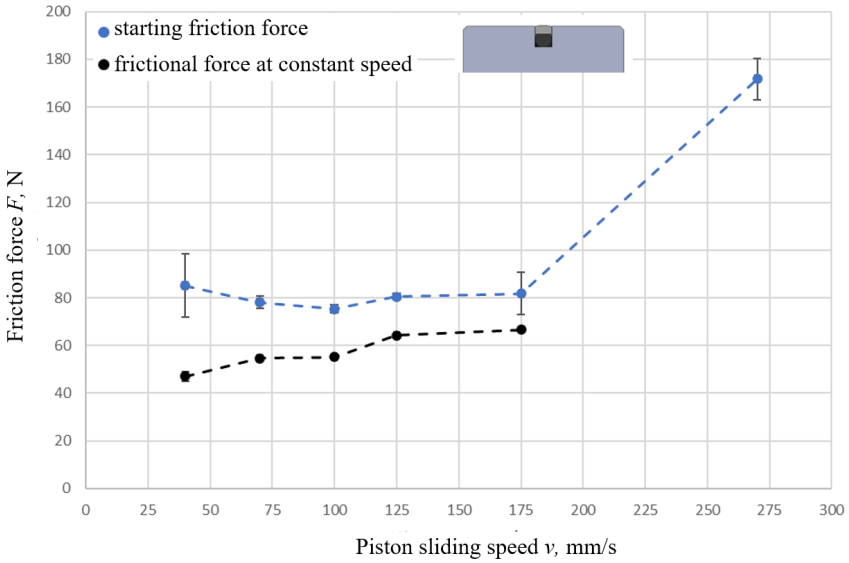


Figure 11: Measured friction force of piston seal "YB 50 39 4.2" in water

Table 2: Measured starting friction force in mineral oil and in water

Seal type	Mineral hydraulic oil				Water			
	F_{min}	v (F_{min})	F_{max}	v (F_{max})	F_{min}	v (F_{min})	F_{max}	v (F_{max})
	N	mm/s	N	mm/s	N	mm/s	N	mm/s
KGD 50 34	100	175	137	270	100	270	246	125
KDA 50 39 9	160	270	230	125	165	270	205	35
YB 50 39 4.2	67	175	110	270	75	100	172	270

Table 3: Measured friction force at constant speed in mineral oil and in water

Seal type	Mineral hydraulic oil				Water			
	F_{min}	v (F_{min})	F_{max}	v (F_{max})	F_{min}	v (F_{min})	F_{max}	v (F_{max})
	N	mm/s	N	mm/s	N	mm/s	N	mm/s
KGD 50 34	45	70	58	270	-	-	-	-
KDA 50 39 9	77	100	95	270	145	175	145	175
YB 50 39 4.2	42	35	84	270	47	35	63	125

For sealings there are a few new alternatives due to impact on health and the environment for PTFE and other PFAS materials such as he hydrolysis-resistant high-performance polyurethane [14]. Mentioned new polyurethane material will also be included in our next researching activities.

3.2. Piston seal durability test in mineral hydraulic oil

The best tested seal from the first test (YB 50 39 4.2) was then subjected to sustained testing on the oil hydraulic test rig. Figure 12 shows the average increase in leakage of the "YB 50 39 4.2" piston seal over the duration of the endurance test. The internal leakage increased from an initial 0.0011 ml/s to an average of 0.009 ml/s after the seal had been run for 100 km. After 100 km, the worn seal is still inside of acceptable limits, according to visual inspection, as shown in Figure 13. At 30x surface magnification, a slight wear of the seal edge can be seen, but it is still below the critical limit.

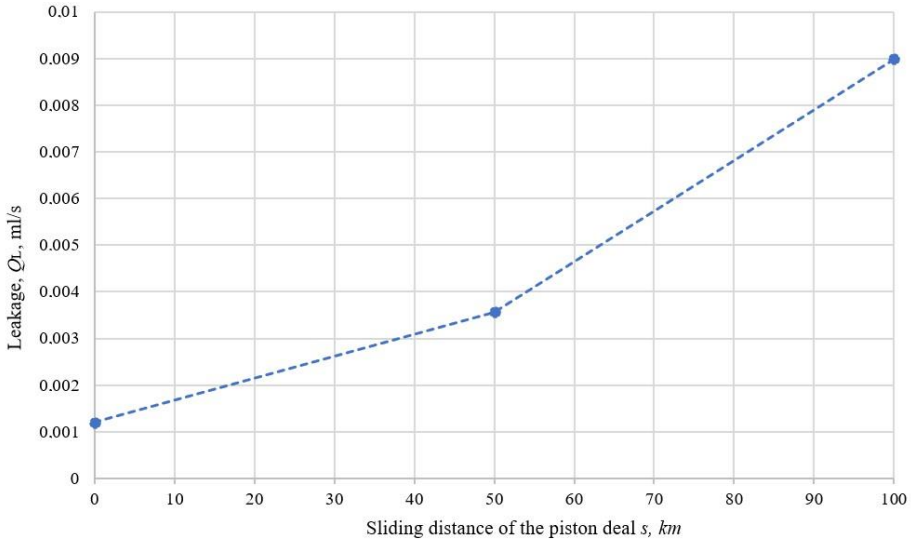
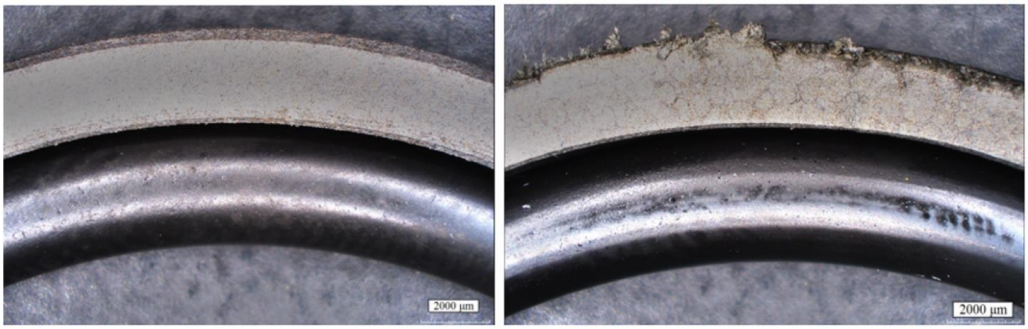


Figure 12: Leakage of the seal "YB 50 39 4.2" on the durability test with mineral hydraulic oil



a b

Figure 13: Surface of tested seal "YB 50 39 4.2"; a) before and b) after 100 km of durability test

4. CONCLUSION AND OUTLOOK

Frictional force was measured at six different speeds for pulling the piston rod out of the cylinder tube in mineral oil and water at the first test rig. The measurements were performed with three different piston seals. An endurance test was performed for the "YB 50 39 4.2" seal type and the leakage of the piston seal assembly was measured in three different piston positions of the tested cylinder. It was found that the seal leakage increased during the durability test. After 100 km of travel, the leakage increased by 7.5 times at the working pressure of 320 bar. Despite the multiple increase, the seal is still usable for most applications with an average final leakage of 0.009 ml per second. It would be advisable to make further research on the influence of working pressure and temperature on the wea rate of the seal and consequently, leakage. We have shown that the amplitude of the frictional force is larger in tap water than in mineral oil. The fraction of frictional force in mineral oil is 60% to 85% of the frictional force generated by the same piston seal unit in water. Lower friction of seals in mineral oil compared to water is mainly due to the significantly thinner lubricating film of water compared to the thickness of the lubricating film of mineral hydraulic oil. In the present study, we have investigated the influence of two hydraulic fluids (mineral oil and water) on the frictional force. We found that the frictional force of piston seals is greater in water than in mineral oil. Considering the ratio of the frictional force in mineral oil and water, we chose the "YB 50 39 4.2" PTFE seal for use in water and oil hydraulics, since the frictional force values in water and mineral oil differ by only 15%. In our future study promising materials such as PFAS (new polyurethane material) will be investigated along with well-established NBR and TPU materials.

ACKNOWLEDGEMENT

The authors would like to thank the company MAPRO, which donated the tested seals and hydraulic cylinder.

NOMENCLATURE

F	Friction force	N
FKM	Fluoroelastomers or fluoro rubber due to American standard (ASTM)	
HNBR	Hydrogenated Nitrile Butadiene Rubber	
ISO VG 46	Mineral hydraulic oil	

KD	A sealing material that assures a good reaction against shock pressure peaks and low friction in the low pressure range (Aston Seals)	
KDA	A sealing material mainly used with high pressure and the backup ring offsets large gaps or structural deflections (Aston Seals).	
KGD	A sealing rubber element with low permanent deformation which assures good sealing performance (Aston Seals).	
NBR	Nitrile Butadiene Rubber	
$n_{\check{c}}$	Rotational speed of the pump shaft	Rpm
p	Pressure	Pa, bar
p_{del}	Working pressure	Pa, bar
p_{EM}	Power of driven electromotor	Pa, bar
p_{pdv}	Pressure of filter by-pass valve	Pa, bar
p_{vv}	Pressure setting of the main pressure relief valve	Pa, bar
PTFE	Polytetrafluoroethylene, a thermoplastic polymer	
PTU	Thermoplastic polyurethane	
v	Sliding velocity	m/s, mm/s
V	Volume	m^3, L
YB	A dynamic seal element which assures exceptional low friction and high speed performance (Aston Seals).	

REFERENCES

[1] Laznik B (2023) Comparison of piston seals friction in water and oil hydraulics, diploma thesis, University of Ljubljana, Faculty of Mechanical Engineering, Ljubljana, Slovenia

[2] Tran X B, Hafizah N, Yanada H (2012) Modeling of dynamic friction behaviors of hydraulic cylinders. Mechatronics, Volume 22, pg. 65-75.

[3] Márton L, Fodor S, Sephiri N (2011) A practical method for friction identification in hydraulic actuators. Mechatronics, Volume 21, pg. 350–356.

[4] Ottestad M, Nilsen N, Hansen M R (2012) Reducing the Static Friction in Hydraulic Cylinders by Maintaining Relative Velocity Between Piston and Cylinder. 12th International Conference on Control, Automation and Systems, Jeju Island, Korea

[5] Nikas G K (2018) Fast performance-analysis of rectangular-rounded hydraulic reciprocating seals: Mathematical model and experimental validation at temperatures between −54 and +135°C. Tribology International, Volume 21, pg. 34-51.

[6] Nikas G K, Almond R V, Burridge G (2014) Experimental Study of Leakage and Friction of Rectangular, Elastomeric Hydraulic Seals for Reciprocating Motion from−54 to+135°C and Pressures from 3.4 to 34.5 MPa. Tribology International, Volume 57, pg. 846-865.

[7] Pan Q, Zeng Y,Li Y, Jiang X, Huang M (2021) Experimental investigation of friction behaviors for double-acting hydraulic actuators with different reciprocating seals. Tribology International, Volume 153.

[8] Wang B, Meng X, Peng X, Chen Y (2021) Experimental investigations on the effect of rod surface roughness on lubrication characteristics of a hydraulic O-ring seal. Tribology International, Volume 156.

[9] Ma K, Wang J, Gu L (2017) Experimental Study on Friction of Hydraulic Cylinder in Different Sealing Systems. The 4th International Conference on Mechatronics and Mechanical Engineering. Havana, Kuba.

[10] Mahankar S, Dhoble A S (2021) Review of hydraulic seal failures due to effect of medium to high temperature. Engineering Failure Analysis, Volume 127.
[11] Chen P, Chua P S K, Lim G H (2007) A study of hydraulic seal integrity. Mechanical Systems and Signal Processing, Volume 21, pg. 1115-1126.
[12] Shanbhag V V, Meyer T J J, Caspers L W, Schlanbusch R (2020) Condition monitoring of hydraulic cylinder seals using acoustic emissions. The International Journal of Advanced Manufacturing Technology, Volume 109, pg. 1727-1739.
[13] Critescu C, Dumitrescu C, Radoi R, Dumitrescu L (2014) Experimental Research for Measuring Friction Forces from Piston Sealing at the Hydraulic Cylinders. Tribology in Industry, Volume 36, No.4, pg. 456-474.
[14] Freudenberg (2023) High-Performance Material Offers Alternative to PTFE: Better Properties At Lower Costs, Available on: https://www.fst.com/news-stories/press-releases/2023/high-performance-material-offers-alternative-to-ptfe/ , last visited: 9th December 2023.

peer reviewed

NUMERICAL MODEL OF PISTON/CYLINDER INTERFACE WITH CONSIDERATION OF TURBULENCE EFFECT FOR WATER HYDRAULICS

Swarnava Mukherjee[1], Haotian Han[1], Lizhi Shang[1*], Georg Herborg[2], Stig Kildegaard Andersen[2]

[1]*Maha Fluid Power Reseach Center, Purdue University, West Lafayette, Indiana 47906*
[2]*Danfoss High Pressure Pumps, Nordborg, Denmark*

* Corresponding author: Tel.: +1 765 496 3174; E-mail address: shangl@purdue.edu

ABSTRACT

Axial piston machines find their use over a wide range of the power spectrum owing to their superior reliability, efficiency, and power density. They are also a key component in applications like reverse osmosis and firefighting wherein the working fluid is water. Utilizing low viscous fluid, such as water, as a working fluid poses challenges in designing the critical lubricating interfaces of the piston pumps. Specifically, low viscosity makes it difficult for the lubricating interfaces to provide sufficient bearing and sealing functions in challenging operating conditions. In order to maintain the lubricating interface performance in water hydraulic piston pumps, costly materials, and tight manufacturing tolerances are often utilized. To improve the efficiency and cost-effectiveness of these pumps, accurate numerical simulation tools that consider the fluid and structure interaction are needed to provide valuable insights into these lubricating interfaces. Although the Reynolds equation is a reliable method for determining the fluid pressure distribution in an oil-based piston pump, it assumes a laminar flow which may not be applicable to water piston machines. For example, in an inclined piston/cylinder interface of a water hydraulic pump, there may be regions in the film wherein the large gap height combined with the low viscosity of water induce turbulence effects. If the traditional Reynolds equation is used in such a scenario, it is likely to overestimate the leakage flow through the interface as it does not account for turbulence. Therefore, it is important to incorporate the effect of turbulence in the diffusive terms of the Reynolds equation to accurately describe the Poiseuille flow with high Reynolds numbers. The challenge is further compounded by the micromotion and deformation of the solid body, resulting in the unevenness of the gap height in the lubricating film. Therefore, the consideration of turbulence can only be applied regionally in such cases. The current study proposes a fluid-structure interaction model with the consideration of the localized turbulent effects. This modeling approach is applied to the piston/cylinder interface of an axial piston machine that uses water as the working fluid. The approach stems from the modification of the Poiseuille term to incorporate a function of the Reynolds number. The fluid dynamics considering the turbulence effect was validated against the solution of the Navier Stokes equation using commercial CFD software. The modified Reynold equation was implemented in an axial piston pump EHL model coupled with the multi-body dynamics. The simulation results from the novel pump model were compared to a measurement and the accuracy of the proposed model was found largely improved from the traditional laminar solution. The calculated flow rate was found to be 54.6% lower with the additional consideration of the turbulence effect in the studied case.

Keywords: Turbulent lubrication, Axial piston machine, Tribology, Experimental validation, Water hydraulics

1. NOMENCLATURE

Symbol	Name	Units
p	Pressure	$[Pa]$
K	Bulk modulus	$[Pa]$
V	Volume	$[m^3]$
ρ	Density	$[kg/m^3]$
$\dot{m}_{in}$	Mass inflow rate	$[kg/s]$
m	Mass	$[kg]$
$\overrightarrow{v_{IF}}$	Inertial frame body velocity	$[m/s]$
$\overrightarrow{F_{i,IF}}$	Inertial frame body force	$[N]$
I_{BF}	Body frame inertia tensor	$[kg \cdot m^2]$
$\overrightarrow{\omega_{BF}}$	Body frame angular velocity	$[rad/s]$
$\overrightarrow{M_{i,BF}}$	Body frame moment	$[N \cdot m]$
ϕ_p	Pressure flow factor	$[-]$
μ	Viscosity	$[Pa \cdot s]$
h	Film thickness	$[m]$
$\overrightarrow{v_m}$	Mean surface velocity	$[m/s]$
ϕ_R	Roughness flow factor	$[-]$
R_q	RMS surface roughness	$[m]$
ϕ_C	Contact flow factor	$[-]$
ϕ_S	Shear flow factor	$[-]$
$\overrightarrow{v_t}$	Bushing velocity	$[m/s]$
$\overrightarrow{v_b}$	Piston velocity	$[m/s]$
$\vec{v}$	Fluid velocity	$[m/s]$
Re	Reynolds number	$[-]$
L_{pist}	Piston length	$[m]$
d_{pist}	Piston diameter	$[m]$
r_{clear}	Nominal radial clearance	$[m]$
L_{bush}	Bushing length	$[m]$
d_{bush}	Bushing diameter	$[m]$

2. INTRODUCTION

Swashplate type axial piston machines are widely used across various industries like agriculture, aerospace, farming and mining equipment, reverse osmosis plants, and firefighting equipment. Their primary advantage is their compactness and efficiency even in demanding operating conditions. Typically, these machines operate with hydraulic oil as the working fluid. However, over the past few years, several researchers have demonstrated the use of water as the working fluid for these machines[1]–[3]. Using water as a working fluid has several advantages: it is sustainable, environmentally friendly, non-flammable, non-toxic, and acts as an effective coolant for the machine. However, water demonstrates certain challenges when used as the working fluid in such machines: its lower viscosity leads to higher leakage flow rates, reduces the load-carrying ability of the lubricating interfaces of the machine, and tends to induce wear in the tribological contacts of the machine. The design of lubricating interfaces of these machines is a critical aspect of designing efficient axial piston machines[4]. Among the three lubricating interfaces of an axial piston machine, the piston/cylinder interface is by far the most critical interface in determining the efficiency and operating limits of the machine[2] owing to its completely hydrodynamic nature. Especially when using water as the working fluid, the design of these interfaces becomes more challenging owing to the physical properties of water. It is therefore crucial to have simulation tools capable of accurately predicting the behavior of the lubricating interfaces to design efficient axial piston machines.

Several modeling approaches for the piston/cylinder lubricating interface of axial piston machines exist in the literature. Wieczorek et al[5] proposed a simulation model for the critical lubricating interfaces of an axial piston machine. Huang et al[6] introduced a fluid-structure interface model for the lubricating interfaces to capture the elasto-hydrodynamic effects. Pelosi et al[7] introduced a fluid structure thermal interaction model to capture thermo-elasto-hydrodynamic effects in the piston/cylinder interface. Ransegnola et al[8] introduced a universal mixed Reynolds equation to capture the cavitated regions of the interface. Mukherjee et al[9] introduced a mutual interaction model between the piston/cylinder gaps to predict distributive fluid behavior in lubrication interfaces. A key limitation in the above-stated simulation models for these lubricating interfaces is that there is no consideration made for localized regions in the interface where the flow might not remain laminar. This is especially important in water hydraulics applications owing to the lower viscosity of water and the possibility of a larger magnitude of wear. Ng and Pan[10] proposed a linearized turbulence theory to account for turbulent regions in lubricating interfaces. This approach was further utilized by Lv et al[11] to model a misaligned journal bearing with mixed lubrication considerations. The current work aims to utilize the linearized turbulence theory to formulate a universal mixed turbulent Reynolds equation tailored for simulating the piston/cylinder interface in axial piston machines.

This paper first describes the modeling approach employed for the piston/cylinder interface operating with water as a working fluid. The turbulent Reynolds equation is introduced to account for localized turbulence effects in the lubricating interface. This is followed by a short comparison of the proposed model with commercial CFD software Simerics MP+ for a simplified geometry. The proposed model is then validated with experimental measurements performed on a reference axial piston machine. Finally, a case study is shown that demonstrates the clearance thresholds at which consideration of turbulence effects becomes important.

3. MODELING APPROACH

The proposed modeling approach for the piston/cylinder interface was developed using a multi-

physics simulation suite, Multics[12]. The base piston/cylinder interface Multics model comprises of several sub-models as shown in **Figure 1**.

The pressure within the displacement chambers, employing a lumped parameter method, serves to provide the boundary conditions for the interface model as well as the forces and moments on the piston. A module dedicated to the dynamics of the piston is employed to solve the equations of motion encompassing all six degrees of freedom for the piston. The resultant piston positions and velocities are subsequently applied to assess the influence of the squeeze and wedge phenomena on the lubricating interface. In this section, a brief description of each of the models is presented.

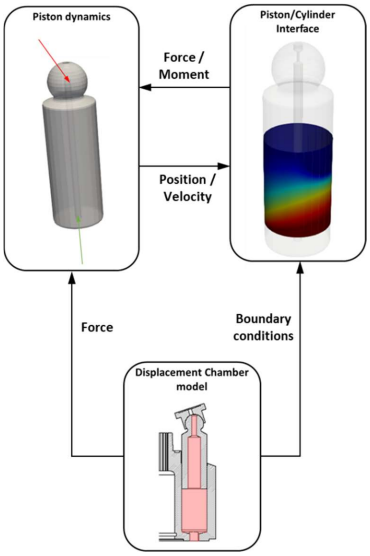


Figure 1: Model Overview

The displacement chamber model solves for the pressure inside a given displacement chamber using the pressure build-up equation as shown in Eq. (1). The volume and volume derivative in Eq. (1) are evaluated using the solution to the piston dynamics equations. $\dot{m}$ in represents the net mass flux into a given displacement chamber. The connection between the displacement chambers and the inlet and outlet ports is assumed to involve orifices with varying areas. These orifice areas represent the momentary flow constriction from a specific displacement chamber to either the outlet or inlet ports. These areas are evaluated using the CFD tool developed by Huang et al[13]. The CFD analysis assesses the smallest instantaneous cross-sectional area along the flow path from the displacement chamber to either of the ports by utilizing 3D representations of the unit. It computes an inviscid flow equation along the flow path and subsequently traces the perpendicular cross-sectional area along the inviscid streamlines to record the minimum cross-sectional area. This assessment is conducted at incremental shaft angles to document the aforementioned variations in the orifice area throughout a complete shaft rotation.

$$\frac{\partial p}{\partial t} = \frac{K}{V}\left(\frac{1}{\rho}\dot{m}_{in} - \frac{\partial V}{\partial t}\right) \tag{1}$$

The forces and moments arising from the pressurized fluid in the displacement chamber and the lubricating interface are used to solve the conservation of linear and angular momentum, as shown in Eq. (2) and (3) respectively, on the piston body. Further details on the force balance of the piston in an axial piston machine can be found in[14].

$$m\frac{d\overrightarrow{v_{IF}}}{dt} = \sum\overrightarrow{F_{\iota,IF}} \tag{2}$$

$$[I_{BF}]\frac{d\overrightarrow{\omega_{BF}}}{dt} = \sum\overrightarrow{M_{\iota,BF}} \tag{3}$$

The Reynolds equation dictates the pressure distribution in a lubricating interface. This pressure distribution is critical to assess the power loss as well as volumetric leakage through the interface. The traditional form of the Reynolds equation is written in terms of the pressure in the lubricating interface as shown by Hamrock et al[15]. This formulation, however, requires the application of the Reynolds boundary condition in the cavitated regions of the interface. This leads to an inaccurate estimation of the flow rate through the interface. In order to alleviate this, Ransegnola[16] proposed a density based formulation of the Reynolds equation. Ransegnola[12] also utilized statistical mixed lubrication relations[17]–[19] to formulate the universal mixed Reynolds equation as shown in Eq. (4).

$$\nabla\cdot\left(\phi_p\left(\frac{Kh^3}{12\mu}\vec{\nabla}\rho\right)\right) = \nabla\cdot\left(\rho\overrightarrow{v_m}\left(\phi_R R_q + \phi_C h\right)\right) + \vec{\nabla}\cdot\left(\rho\frac{\phi_S}{2}R_q(\overrightarrow{v_t} - \overrightarrow{v_b})\right) + \frac{\partial\rho\left(\phi_R R_q + \phi_C h\right)}{\partial t} \tag{4}$$

It is however noted that Eq. (4) is based on the assumption of laminar flow in a lubricating interface. The use of water in axial piston machines combined with the material choices to work with such a low viscous lubricant may lead to wear of the piston and/or bore running surfaces leading to a large gap between the piston and bore. The large gap resulting from the wear coupled with the low viscosity of water leads to the possibility of turbulent flow in the interface geometry, breaking the assumption of laminar flow. Ng and Pan[10] proposed a linearized theory for turbulent lubrication to account for localized turbulent effects in a lubricating interface. This model was also successfully adopted by other researchers[11], [20]. The fundamental idea behind the formulation is the modification of the Poiseuille term in the Reynolds equation as shown in Eq. (5) where $\boldsymbol{Re}$ corresponds to the local Reynolds number in the interface. The combined local Couette and Poiseuille mean velocity in the interface is represented in Eq. (6). As a result, the local Reynolds number in the interface can be expressed as shown in Eq. (7).

$$\nabla\cdot\left(\phi_p\left(\frac{Kh^3}{\boldsymbol{f(Re)}\mu}\vec{\nabla}\rho\right)\right) = \nabla\cdot\left(\rho\overrightarrow{v_m}\left(\phi_R R_q + \phi_C h\right)\right) + \vec{\nabla}\cdot\left(\rho\frac{\phi_S}{2}R_q(\overrightarrow{v_t} - \overrightarrow{v_b})\right) + \frac{\partial\rho\left(\phi_R R_q + \phi_C h\right)}{\partial t} \tag{5}$$

$$\vec{v} = \phi_C \overrightarrow{v_m} - \phi_p \frac{h^2}{12\mu} \vec{\nabla} p \tag{6}$$

$$Re = \frac{\rho |\vec{v}| h}{\mu} \tag{7}$$

The function in the Reynolds equation proposed by Ng and Pan[10] is shown in Eq. (8). This variation is graphically represented in **Figure 2**. It is observed that this function has a value close to 12 at low Reynolds numbers recovering the original formulation of the laminar Reynolds equation (4). However, the value of the function increases with the Reynolds number to account for the turbulent effects in the lubricating interface.

$$f(Re) = 12 + 0.00113 Re^{0.9} \tag{8}$$

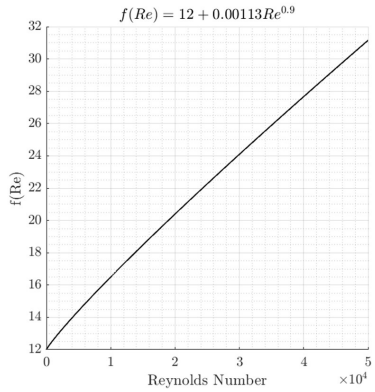


Figure 2: Turbulence Function

In order to assess the validity of the proposed formulation, a simplified thin gap geometry is constructed as shown in **Figure 3** which is representative of a worn piston/cylinder interface of a water hydraulic pump that was studied by the authors' research group. The dimensions in of the simplified geometry are normalized to L_{pist}, d_{pist} and r_{clear}, which correspond to the piston running surface length, piston diameter and nominal radial clearance of the reference unit. The geometry represents a piston tilted inside the bore. The gap height of the simplified geometry is a few times higher than the nominal designed clearance to represent the wear. In order to isolate the pressure driven flow component, a total pressure of 80 bar is specified at the bottom of the geometry and a total pressure of 5 bar is specified at the top of the geometry. No piston motion, body deformation, and thermal effect are considered.

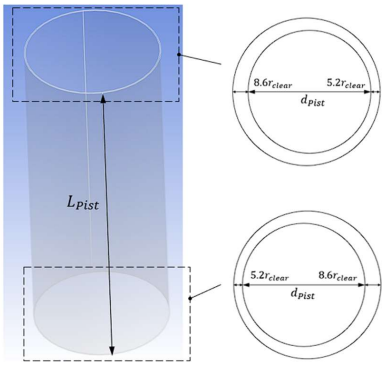


Figure 3: Simple Geometry

Simerics MP+ was used to solve the above-described simplified case with and without a standard $k = \epsilon$ turbulence model. The boundary conditions were specified to be a total pressure of 80 bar on one axial end of the geometry and a total pressure of 5 bar was specified on the other end. The same problem is recreated using the proposed model with Eq. (5) and without Eq. (4) the turbulence factor as well. The boundary conditions for the proposed model were specified pressures of 80 bar and 5 bar at the 2 different axial ends of the geometry. It is also noted that for the model in Simerics MP+, the 5 bar inlet was specified with a fully developed velocity profile. A comparison of the pressure distribution in the fluid geometry is shown in **Figure 4**. It is observed that the pressure distribution prediction demonstrates a significant difference between the laminar and turbulent variants. The distribution predicted by the proposed model closely matches the ones predicted by Simerics MP+.

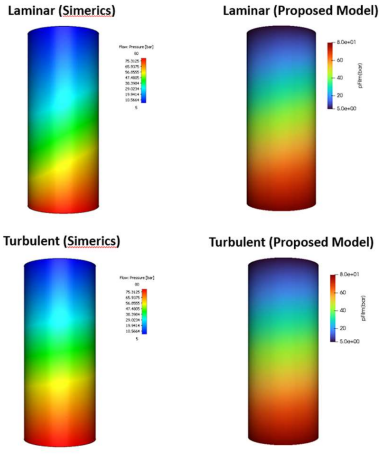


Figure 4: Pressure Distribution Comparison

In order to observe the differences quantitatively, the mass flow rate at the top boundary of the geometry for all the cases is shown in **Figure 5**. Based on the Reynolds number (7) calculation implemented in the proposed model, the turbulence factor multiplier is increased from 0.00113 to 0.00286. It is observed that both the laminar and turbulent mass flow rates predicted by the proposed model and Simerics MP+ match quite well. This match provides confidence in the implementation of the proposed turbulence formulation; however, for a given set of geometry and operating conditions, the turbulence factor may need to be studied further. **Figure 5** presents the prediction of the flow rate with and without the consideration of the turbulence models in both Simerics MP+ and the proposed model by activating and deactivating the turbulence models respectively.

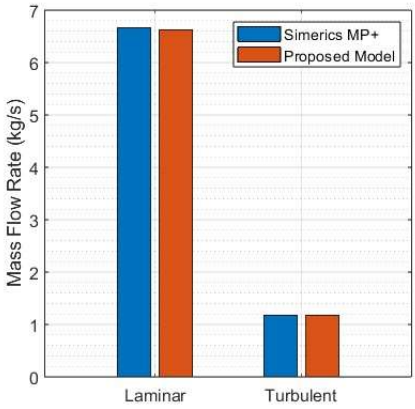


Figure 5: Mass Flow Rate Comparison on Simplified Geometry

4. MODEL VALIDATION

The developed model was validated with experimental measurements performed on the reference fixed displacement swashplate-type water hydraulic pump. The specifications of the reference pump are shown in **Table 1**. Steady-state measurements were performed on the unit according to ISO4409 standards. Steady-state inlet flow rate, outlet flow rate, and a volumetric loss flow rate were measured for the purposes of validation of the proposed model. The highest speed and highest outlet pressure operating conditions were chosen to demonstrate the fidelity of the proposed model. In order to provide a comprehensive insight into the model predictions, steady-state measurements were performed with a set of new unworn bores and a set of severe worn-out bores.

Table 1: Parameters of the reference unit

Parameter	Value [Units]
Displacement	444 [cc/rev]
Maximum speed	1500 [rpm]
Minimum speed	100 [rpm]
Maximum outlet pressure	85 [bar]

Minimum outlet pressure	10 [bar]
Maximum inlet pressure	5 [bar]
Minimum inlet pressure	2 [bar]
Number of pistons	9

In order to incorporate the worn-out bores into the simulation model, the worn-out bores were measured using a contact probe instrument. It is noted here for the reference unit under consideration, insignificant wear was observed on the piston compared to the bore. Thus, for the simulations discussed in the study, the piston wear was neglected. **Figure 6** shows the point cloud that was captured using the contact probe instrument on a worn bore.

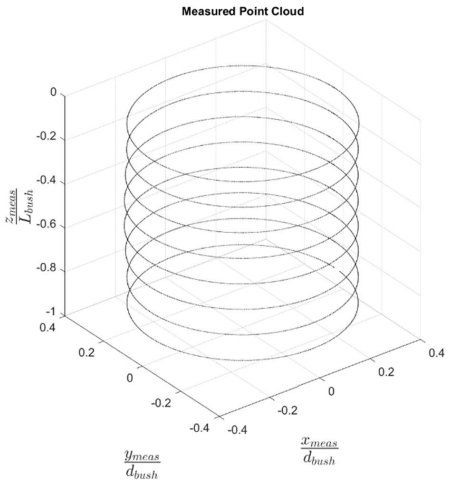


Figure 6: Point Cloud Measurement

The simulation model requires a tabulated input of the wear profile that dictates the variation of the actual profile from the nominal bore diameter. In order to achieve this, a method of the fitted cylinder is utilized to derive the deviation of the measured profile from the nominal bore radius. **Figure 7** shows a representation of the evaluation made to derive the wear profile of a given bore. The geometric centers of each axial layer from **Figure 7** are first evaluated by fitting a circle through these points. A straight three-dimensional line is then fit through these derived centers. A cylinder with a radius of the nominal bore is then constructed about the fitted line. The deviation of each measured point is then evaluated from the resulting fitted cylinder.

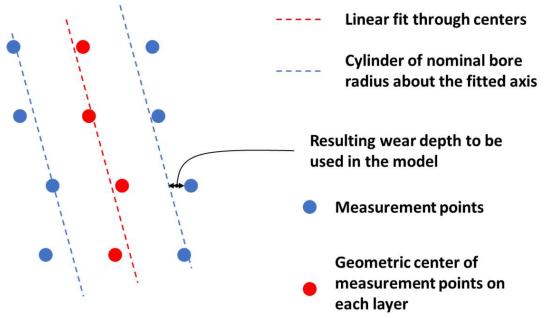


Figure 7: Fitted Cylinder Method to Evaluate Wear Depth

The wear depth distribution on the bore normalized with the nominal radial clearance is shown in **Figure 8**. It is observed that the worn profile demonstrates about 8 times the nominal clearance in some regions. This represents the worst-case wear for such large units with low viscosity working fluids like water. Unlike traditional units with small clearances and high viscosity working fluids, the lubrication flow in these regions of high clearance need not be completely laminar.

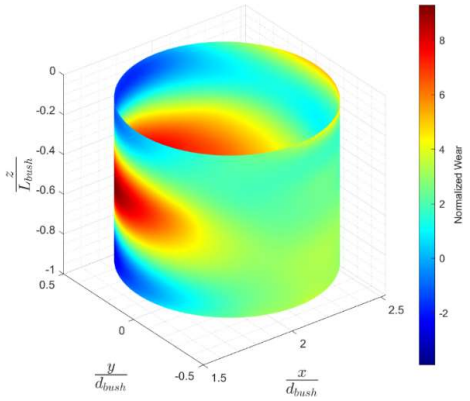


Figure 8: Normalized Bore Wear Profile

The operating condition of interest was chosen to be a differential pressure of 80 bar across the pump at 1500 rpm. Leakages were measured with a set of worn-out bores as well as new unworn bores. The volumetric leakage flow normalized with the theoretical flow rate of the unit is shown in **Table 2**. It is noted here that the experimental setup measured the cumulative leakages from all lubricating interfaces through the drain port on the unit. However, for the purposes of comparison of the experimental measurements with the simulation model of the piston/cylinder interface, the majority of the external leakages were assumed to be rising from the piston/cylinder interface. The simulated leakages discussed in this study are evaluated as the average leakage across all 9 piston/cylinder interfaces of the unit.

Table 2: Measured Leakages

Case	Normalized volumetric leakage flow rate
Unworn bore	0.027
Worn bore	0.106

The simulated variation of the normalized leakage flow through the piston/cylinder interface for the two sets of bores is shown in **Figure 9**. It is observed that for the unworn bore, there is no significant difference between the laminar and turbulent Reynolds equation since the flow is essentially laminar everywhere in the interface. However, for the worn bore, there is a significant difference between the laminar and turbulent Reynolds equations. This is primarily due to the large local clearances in the interface that trigger turbulent flow in the interface. In order to compare the simulated leakages with the experimentally measured ones, an average of the simulated leakages from **Figure 9** over an entire shaft revolution across all 9 piston/cylinder interfaces is performed to provide an estimate of the steady state. In order to evaluate the piston/cylinder interface leakage from the measured volumetric flow loss, it was assumed that the piston/cylinder interface contributes to the majority of the volumetric flow loss. This is because the slipper/swashplate and cylinder block/valveplate interface components are new and therefore provide adequate sealing in their respective interfaces. A comparison between the experimentally evaluated leakages and simulated leakages are shown in **Figure 10**. For the unworn bore, both the models predict the leakage very close to each other because of the low Reynolds number in the unworn gap. The simulation results are expected to be lower than experimentally obtained values since other leakage sources such as slippers and the valveplate are ignored. For the worn bore, it is observed that the turbulent Reynolds equation predicts a leakage much closer to the experimentally measured ones compared to the laminar Reynolds equation. There is however some degree of over-prediction. This is attributed to the coefficients from Eq. (8). Based on Ng and Pan[9], the multiplier and exponent in the turbulence function are specific to geometry and operating conditions.

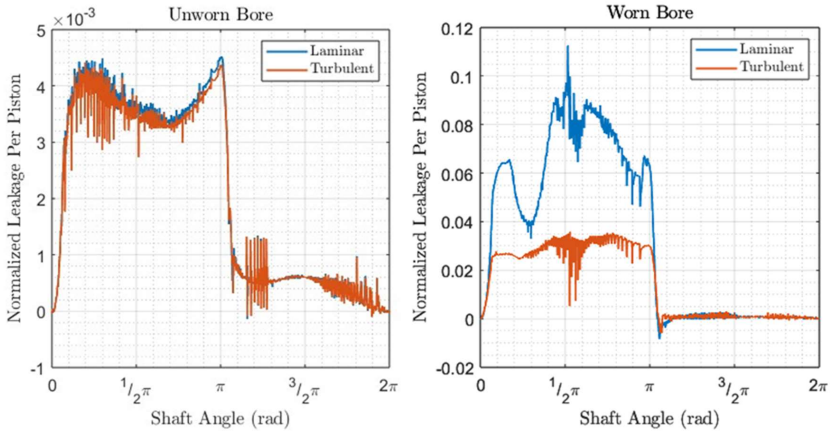


Figure 9: Variation of Leakage from the Piston/Cylinder Interface

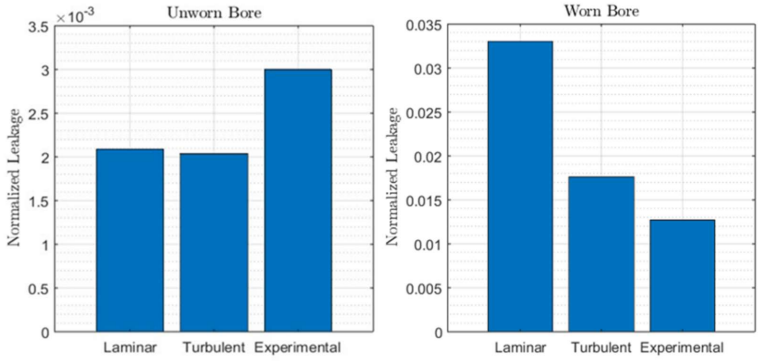


Figure 10: Simulated vs. Experimental Leakage

5. RESULTS AND DISCUSSION

The importance of turbulence in large clearances has been established in the previous section. In order to provide more insight into the applicability of the turbulent Reynolds equation, several cases with partially worn bores have been performed both with the laminar and turbulent Reynolds equations. Table 3 shows the different simulated cases. The 0% wear case corresponds to the nominal bore diameter and the 100% wear case corresponds to the full wear profile from **Figure 8**. The percentages in between correspond to a linearly increasing wear profile between the nominal geometry and fully worn geometry. This study provides insight into a threshold amount of wear after which turbulent effects become non-negligible.

Table 3: Simulation Case Study with Increasing Wear Intensities

Case	% Wear
1	0
2	25
3	50
4	75
5	100

Figure 11 shows the variation of mean leakage between the different cases predicted using the laminar and turbulent Reynolds equations. It is evident that at about 50% wear the prediction between the laminar and turbulent equations deviates indicating that turbulence effects become significant. The larger local clearances combined with the low viscosity of water invalidate the laminar flow assumption locally in the interface.

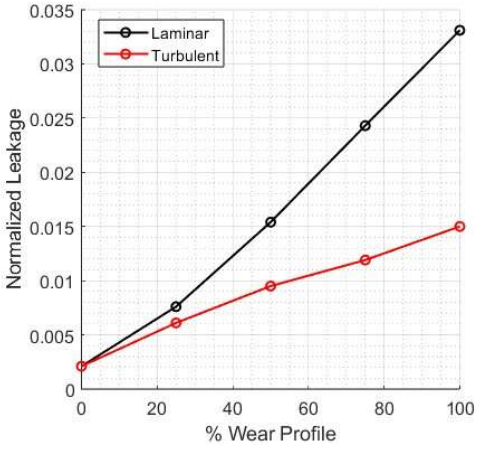


Figure 11: Laminar vs. Turbulent for Varying Clearances

6. CONCLUSION

The study first introduced a framework to incorporate the effect of turbulence in lubricating interfaces. The introduced approach was compared with commercial CFD software, Simerics MP+ for a simplified geometry. Later, the proposed model was validated with experimental measurements on a 444cc/rev commercial axial piston water pump. Finally, a case study was presented to highlight the importance of turbulence effects in lubricating interfaces. The proposed model demonstrates the ability to be used as a predictive tool for extreme operating conditions with innovative geometries and working fluids. Rather than predicting an extremely high leakage rate like the laminar Reynolds equation, the current approach provides a more realistic leakage estimate to avoid false failure metrics during the research and development phase of novel axial piston machines.

REFERENCES

[1] Trostmann E (2000) Tap Water as a Hydraulic Pressure Medium. CRC Press
[2] Ernst M, Vacca A, Ivantysynova M, Enevoldsen G (2020) Tailoring the Bore Surfaces of Water Hydraulic Axial Piston Machines to Piston Tilt and Deformation. Energies, vol. 13, no. 22, Art. no. 22
[3] Ernst M, Ivantysynova M, Vacca A (2022) Shaping the Piston–Cylinder Interfaces of Axial Piston Machines for Running in the High-Pressure Regime with Water as the Hydraulic Fluid. J. Mechanical Engineering Science, vol. 236, no.12, pp. 6851-6872
[4] Hasko D, Shang L, Noppe E, Lefrançois E (2019) Virtual Assessment and Experimental Validation of Power Loss Contributions in Swash Plate Type Axial Piston Pumps. Energies, vol. 12, no. 16, Art. no. 16
[5] Wieczorek U, Ivantysynova M (2002) Computer Aided Optimization of Bearing and Sealing Gaps in Hydrostatic Machines—The Simulation Tool Caspar. Int. J. Fluid Power, vol. 3, no. 1, pp. 7–20
[6] Ivantysynova M, Huang C (2002) Investigation of the Gap Flow in Displacement Machines Considering Elastohydrodynamic Effect. Proc. JFPS Int. Symp. Fluid Power, vol. 2002, no. 5–1, pp. 219–229
[7] Pelosi M, Ivantysynova M (2012) A Geometric Multigrid Solver for the Piston–Cylinder Interface of Axial Piston Machines. Tribol. Trans., vol. 55, no. 2, pp. 163–174
[8] Ransegnola T, Shang L, Vacca A (2022) A study of piston and slipper spin in swashplate type axial piston machines. Tribol. Int., vol. 167, p. 107420
[9] Mukherjee S, Vacca A, Shang L, Sharma A (2023) A Thermal Modeling Approach for the Piston/Cylinder Interface of an Axial Piston Machine Under Asperity Contact. Meccanica, vol. 58, no. 10, pp. 1929-1957
[10] Ng C-W, Pan C H T(1965) A Linearized Turbulent Lubrication Theory. J. Basic Eng., vol. 87, no. 3, pp. 675–682
[11] Lv F, Jiao C, Ta N, Rao Z (2018) Mixed-lubrication analysis of misaligned bearing considering turbulence. Tribol. Int., vol. 119, pp. 19–26
[12] Ransegnola T (2020) A strongly coupled simulation model of positive displacement machines for design and optimization. PhD Dissertation, Purdue University
[13] Ivantysynova M, Huang C, Christiansen S-K (2004) Computer Aided Valve Plate Design - An Effective Way to Reduce Noise. SAE Trans., vol. 113, pp. 162–173
[14] Mukherjee S (2023) A Multi-Domain Thermal Model for Positive Displacement Machines. PhD Dissertation, Purdue University
[15] Hamrock B J, Schmid B J, Jacobson B O (2004) Fundamentals of Fluid Film Lubrication. CRC Press
[16] Ransegnola T, Sadeghi F, Vacca A (2021) An Efficient Cavitation Model for Compressible Fluid Film Bearings. Tribol. Trans., vol. 64, no. 3, pp. 434–453
[17] Patir N, Cheng H S (1978) An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication. J. Lubr. Technol., vol. 100, no. 1, pp. 12–17
[18] Wu C, Zheng L (1989) An Average Reynolds Equation for Partial Film Lubrication With a Contact Factor. J. Tribol., vol. 111, no. 1, pp. 188–191
[19] Patir N, Cheng H S (1979) Application of Average Flow Model to Lubrication Between Rough Sliding Surfaces. J. Lubr. Technol., vol. 101, no. 2, pp. 220–229
[20] Shenoy S B, Pai R (2009) Theoretical investigations on the performance of an externally adjustable fluid-film bearing including misalignment and turbulence effects. Tribol. Int., vol. 42, no. 7, pp. 1088–1100

DEVELOPMENT OF DIGITAL TYPE TAP-WATER DRIVE FLOW CONTROL VALVE

Ryouichi Hamada[1], Wataru Kobayashi[1]*, Masashi Yokota [2], Hiroki Atogami[3]

[1] *Graduate School of Engineering, Okayama University of Science,700-0005 Ridai-cho 1-1, Kita-ku, Okayama-city, Japan*
[2] *Department of Information Science and Engineering, Okayama University of Science,700-0005 Ridai-cho 1-1, Kita-ku, Okayama-city, Japan*
[3] *Graduate School of Science and Engineering, Okayama University of Science,700-0005 Ridai-cho 1-1, Kita-ku, Okayama-city, Japan*

* Corresponding author: Tel.: +81 86 256-9528; E-mail address: w-kobayashi@ous.ac.jp

ABSTRACT

This study concerned with development of digital type tap-water drive flow control valve for water hydraulic systems. In this study we developed a novel type of valve which can generate 10 steps of output flow rate. In addition, a back pressure supply mechanism is applied to the developed valve to assist torque of a stepping motor driving a rotary disk in the valve. The improved valve with the back pressure supply mechanism is compared with conventional digital valve units using typical On/Off valves. As a result, it was confirmed that the size was reduced by 30% and the mass was reduced by 70%.

Keywords: Water hydaurlics, Digital fluid power, Digital valve unit, Flow control valve

1. INTRODUCTION

Water hydraulic systems have been applied to food processing, medical instrument, and semiconductor devices, which require high cleanliness and environmental friendliness shown in **Figure 1** [1]. This is because their working fluid is water and the systems have 100% oil-free characteristics. In particular, tap-water drive systems take much more attention because they need no hydraulic pumps. The systems are likely to be used at home because supply and disposal of water as a working fluid is easy. In addition, there is no danger which generates heat and an electric shock. From these reasons, tap-water drive systems can be a new driving power source.

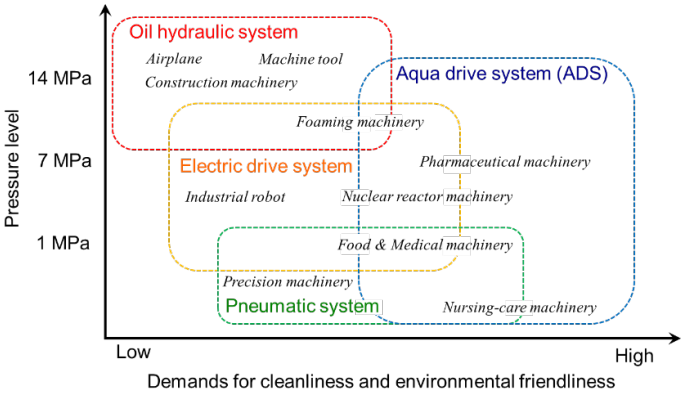
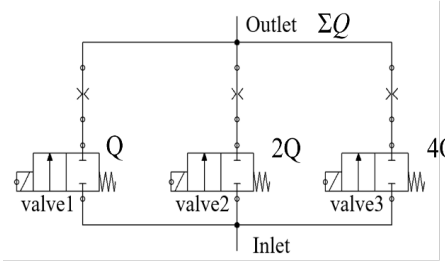


Figure 1: comparison of four driving systems

On the other hand, the systems have following problems; 1) Maximum supply pressure of tap-water in Japan is only 300kPa, 2) Leakage of the systems is relatively large due to low viscosity of water, 3) Few types of water hydraulic elemental equipment have been developed compared to the other driving systems such as oil hydraulic and pneumatic systems. This study focuses on third problem and aims to develop a new component for tap-water drive systems.

In general, there are two type valves for hydraulic systems to control pressure or flow rate. One is On/Off valves and they have characteristics that their operations are only open or close. Although the valves are also low priced, they can only switch the direction of flow but cannot control pressure or flow rate directly. The other is proportional valves for hydraulic systems and they have characteristics that can control valve opening area continuously, which means they can control pressure or flow rate. The valves, however, are high priced among other drive sources. From these problems, when high control performance is required, we have to apply proportional valves.

Recently, digital fluid power has been focused and researched actively [1]-[3]. As related study for oil hydraulic systems, digital valve unit (DVU, for short) was proposed in Europe. DVU can regulate flow rate in incremental steps by using of pulse code modulation (PCM) method [4]. PCM is known as a way of analog-digital conversion and the way has been widely used for compact disks. **Figure 2** shows structure of DVU and obtained outlet flow rate by combination of the valve On/Off patterns. DVU consists of a number of typical On/Off valves and can generate stepwise flow rate by changing combination of On/Off pattern for each valves. For instance, when rated flow rate of three valves are set to Q, $2Q$, and $4Q$, outlet flow rate of DVU can be from 0 to $7Q$, which is like binary codes. Thus, DVU are less costly than proportional valves because it uses only On/Off valves and the valves in general can be gotten inexpensively. However, whole system of DVU may be large size and expensive when the system consists of lots of valves, which means it requires precise incremental steps of outlet flow, in order to make its control performance high. In related study, DVU showed possibility of reduction in energy consumption compared with the energy consumption of a typical 2/2 proportional valve. In addition, it is clear that the concept difficulties are valve states uncertainty and pressure peaks.



0 : closed valve 1 : open valve			Outlet Flow
V3(4Q)	V2(2Q)	V1(1Q)	(ΣQ)
0	0	0	0
0	0	1	1
0	1	0	2
0	1	1	3
1	0	0	4
1	0	1	5
1	1	0	6
1	1	1	7

1 : On
0 : Off
V : valve
ΣQ : outlet flow

Figure 2: Structure of digital valve unit and obtained outlet flow of DVU

2. DIGITAL TYPE FLOW CONTROL VALVE

The aim of this study is development of the digital flow control valve with incremental steps of outlet flow for tap-water drive systems. In particular, to realize the function of DVU with a valve can solve one of above problems for DVU. Also, it is expected that the developed valve can reduce its weight and package size of a whole system. Thus, the valve can provide a new choice of valves for water hydraulic systems.

Figure 3 shows structure and combination of On/Off pattern of a developed valve which we call it digital type flow control valve in this study. The valve consists of a stepping motor, some steel balls, a rotary disk, input and output ports. The output port has a number of orifices corresponding to a

number of On/Off valves in DVU. The steel balls in the rotary disk closed these orifices when the balls overlapped on the orifices. Position arrangement of the balls is set on every 30 degrees on the rotary disk shown in **Figure 3**. Then, the valve can generate incremental steps of outlet flow rate by rotation of the stepping motor. For instance, position 1 which consists of one opening and five closing orifices indicates outlet flow rate of Q and position 2 which consists of two opening and four closing orifices indicates outlet flow rate of $2Q$ and position 3 which consists of three opening and three closing orifices indicates outlet flow rate of $3Q$. Thus, position arrangement of the steel balls with one rotary disk can achieves stepwise outlet flow by using a stepping motor. Note that we can get more steps of flow rate when the rotary disk has different orifice sizes such as A, $2A$, $4A$ and more although this describes only six steps of outlet flow rate which is from Q to $6Q$. Moreover, If the orifice diameter on the rotary disk is equivalent to it on typical On/Off valves consisting conventional DVU, package size and total cost of a whole system can be drastically reduced.

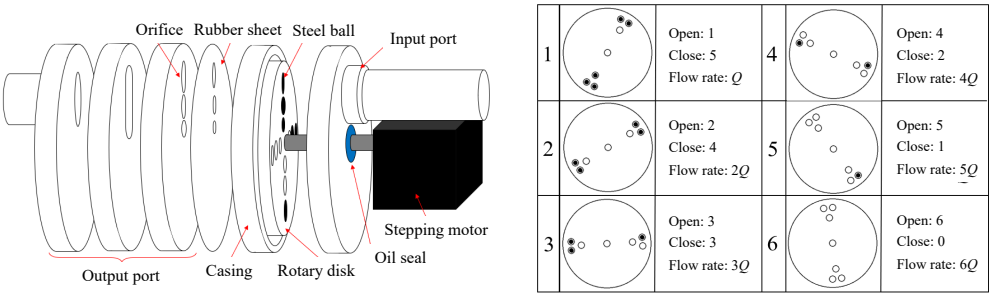


Figure 3: Structure and combination of On/Off pattern for digital type flow control valve

Figure 4 shows a prototype of digital type On/Off valve. The developed valve consists of a stepping motor (ST-42BYG020, Mercury Motor), an acrylic rotary disk with 36 orifices (diameter of 1.8 mm) shown in **Figure 3**, acrylic plates with input and output ports, an oil seal, a rubber sheet (thickness of 0.5 mm) with 6 holes (diameter of 1.4 mm), 15 steel balls (diameter of 3 mm), and input and output connectors. The valve can theoretically generate 7 steps of output flow rate including fully closed. **Figure 5** shows an experimental setup for flow measurement of the developed valve. The setup consists of the valve, a pressure sensor (PSE560-02, SMC Corp.), a flowmeter (FD-XS20, KEYENCE Corp.), and a micro-computer (H8/3664F, Renesas Electronics Corp.) to drive the stepping motor on the valve. Flow measurement is carried out under various supply pressure conditions from 0.05 to 0.30 MPa, which assumes tap-water pressure level and no-load. **Figure 6** shows experimental result for flow measurement of the valve. As seen in the result, the valve can generate incremental steps of output flow regardless of supply pressure and the output flow seems to be proportional to the number of opening orifices on the rotary disk.

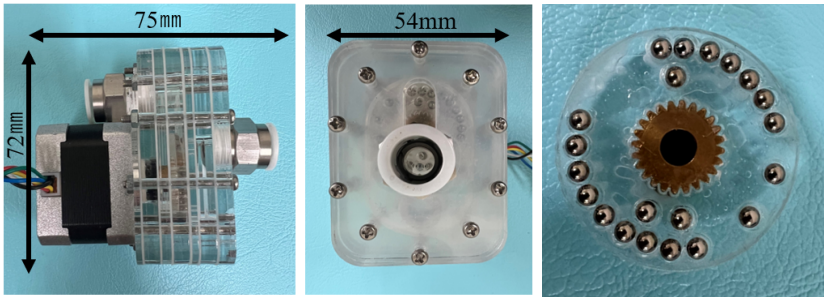


Figure 4: Prototype of digital type flow control valve

The developed valve, however, has a problem that the torque of the stepping motor may be insufficient when number of orifices and steel balls have to be increased in order to increase steps of output flow rate of the valve. To overcome this difficulty, we propose an assist mechanism using back pressure of the valve to rotate the stepping motor.

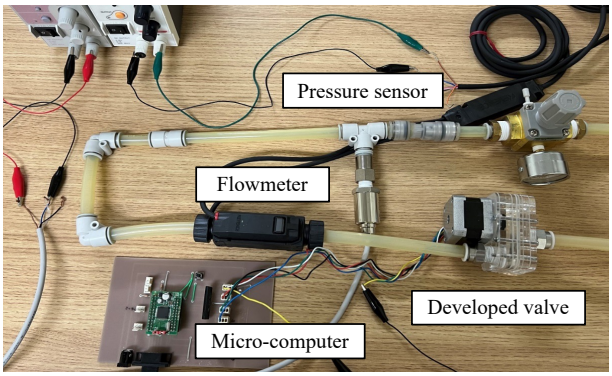


Figure 5: Experimental setup for flow measurement of developed valve

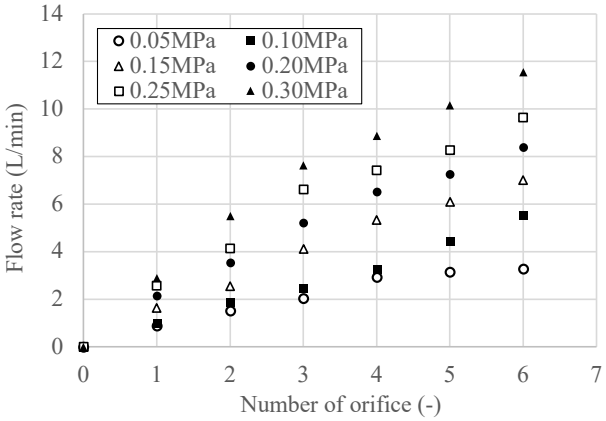


Figure 6: Experimental result (flow measurement of developed valve)

3. IMPROVEMENT OF DIGITAL TYPE FLOW CONTROL VALVE

Improvement of the developed valve is considered as mentioned above. A key concept of improvement is to use back pressure of the valve for torque assist of motor rotation. **Figure 7** describes a schematic diagram of a back pressure supply mechanism of the improved valve. The symbols of the ball valves in the figure express the orifices in the rotary disk of the valve and there exists an additional On/Off valve connecting input and output ports of the flow control valve directly. When the stepping motor connected with the rotary disk rotates, the On/Off valve opens and then back pressure can be generated to balance supply and back pressure around the steel ball. On the other hand, this study aims to reduce weight and package size of the proposed valve and it also needs an additional small-sized bypass valve in the back pressure supply mechanism.

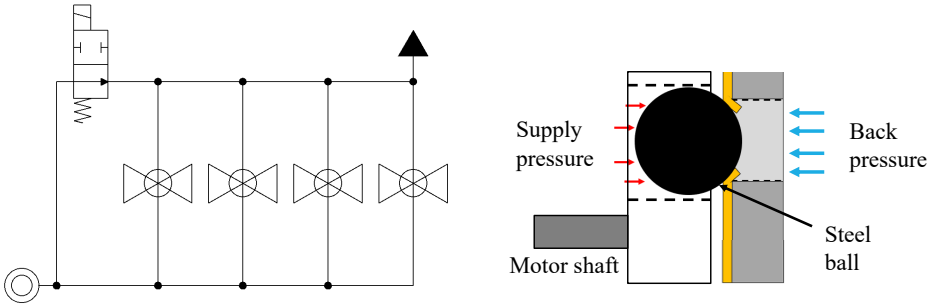


Figure 7: Schematic diagram of back pressure supply mechanism

3.1. Back pressure supply valve

Back pressure supply mechanism should have small-sized and simple structure. In this study, we apply a magnetic holding mechanism to open/close motions of an orifice. **Figure 8** shows a schematic diagram of a back pressure supply valve using magnet balls and coil to move a magnet ball. Exciting the coil momentarily, magnet ball for orifice is moved and overlapped on the orifice, that is, the valve can be closed. Same way for excitation of the coil can open the valve because the magnet ball can rotate and the other magnet ball on the opposite side can pull the ball for orifice.
Figure 9 shows a prototype of back pressure supply valve. The valve consists of three magnet balls (diameter of 3 mm), an enamel wire coil (diameter of 0.2 mm and the number of turns of 200 times), a rubber sheet (thickness of 0.5 mm), casing with input and output ports made by 3D printer, input and output connectors. It has an orifice of 1.8 mm inside and has also relatively small-sized and simple structure.

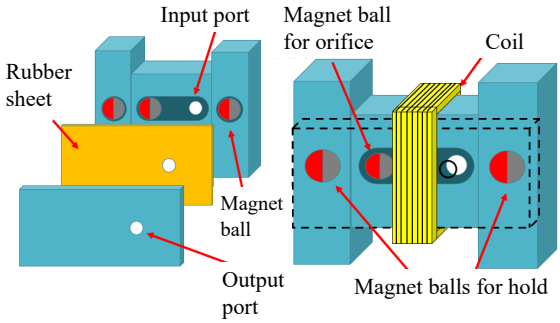


Figure 8: Schematic diagram of back pressure supply valve

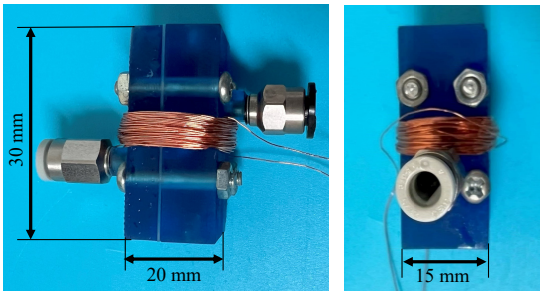


Figure 9: Prototype of back pressure supply valve

3.2. Improved flow control valve with back pressure supply mechanism

The previous developed valve with six steps of output flow has the problem of motor torque and we apply the back pressure supply mechanism mentioned above to overcome the difficulty. Improved flow control valve with the mechanism is tested here. The improved valve has 10 steps of output flow as shown in **Figure 10**. The combination of On/Off pattern for the valve is basically same as the combination for the previous valve in **Figure 3**. A prototype of the improved valve, which is in **Figure 11** has length of 70 mm, width of 70 mm, height of 60 mm and its mass is approximately 400 g. **Figure 11** shows an experimental setup for flow measurement of the improved valve with the back pressure supply valve. The setup consists of the improved valve, a couple of pressure sensors (PSE560-02, SMC Corp.), a flowmeter (FD-XS20, KEYENCE Corp.), and a micro-computer (ESP32-WROOM-32) to drive the stepping motor on the valve. Flow measurement is carried out under various supply pressure conditions from 0.05 to 0.30 MPa, which is same conditions as previous flow measurement in Chap. 2. **Figure 12** shows experimental result for flow measurement of the improved valve. As seen in the result, the improved valve can generate more incremental steps of output flow regardless of supply pressure than the steps of the previous valve which has up to six steps due to shortage of torque of the stepping motor. Thereby, it is confirmed that the additional back pressure supply mechanism can work well. Note that the orifice size of the back pressure supply valve is same as the orifice in the improved valve and it means that the improved valve has totally 8 steps of output flow rate with additional one step.

Table 1 lists comparison of the digital type flow control valve develop in this study with DVU using typical On/Off valves. The On/Off valves used in DVU (USG-3-6-2, CKD Corp.) has same orifice diameter as the one of the developed valve. As a result, the size was reduced by 30% and the mass was reduced by 70%.

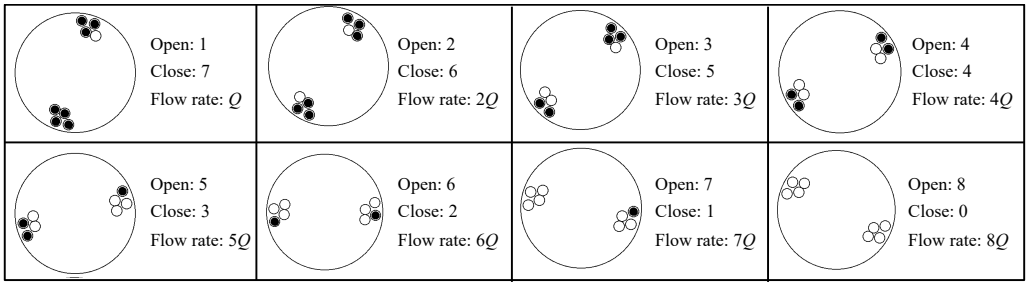


Figure 10: Combination of On/Off pattern for improved valve

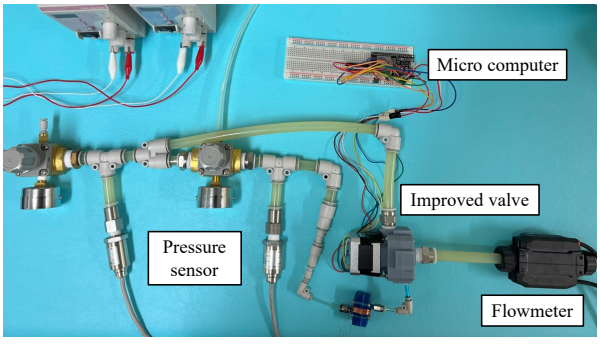


Figure 11: Experimental setup for flow measurement of improved valve

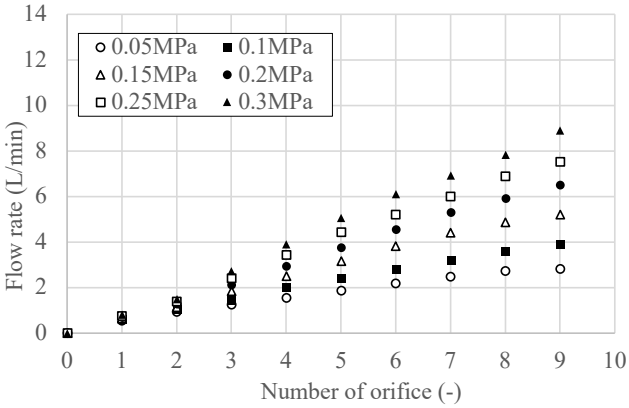


Figure 12: Experimental result (flow measurement of improved valve)

Table 1: Comparison of developed valve with DVU using typical On/Off valves

Comparative parameters	Developed valve	DVU using 9 On/Off valves
Size (L × W × H)	70 mm × 70 mm × 60 mm	80 mm × 75 mm × 75 mm
Total mass	400 g	1350 g
Orifice diameter	1.8 mm	1.8 mm

4. CONCLUSION

This study concerned with development of digital type tap-water drive flow control valve to give another choice of flow/direction control valves, which are generally On/Off valves and proportional/servo valves, for water hydraulic systems. In this study, we developed a novel type of valve which can generate 10 steps of output flow rate. In addition, a back pressure supply mechanism was applied to the developed valve to assist torque of the stepping motor driving a rotary disk in the valve. The improved valve with the back pressure supply mechanism was compared with DVU using typical On/Off valves. As a result, it was confirmed that the size was reduced by 30% and the mass was reduced by 70%. Thus, a compact, lightweight valve with 10 steps of output flow was realized.
This work was supported by JSPS KAKENHI Grant Number JP21K14133.

REFERENCES

[1] Heikkilä, M., and Linjama, M. (2012) Improving damping characteristics of displacement controlled digital hydraulic system. Proceedings of the fifth workshop on digital fluid power 89-102

[2] P. Foschum, A. Plöckinger, and R. Scheidl (2012) Simulation and Experimental Results of PWM Control for Digital Hydraulics. Proceedings of the fifth workshop on digital fluid power 133-151

[3] Matti Linjama (2011) Digital Fluid Power – State of the Art, The Twelfth Scandinavian International Conference on Fluid Power, May 18-20, Tampere, Finland

[4] Essam ELSAED, Mohamed ABDELAZIZ, Nabil A. MAHMOUD (2017) Investigation of A Digital Valve System Efficiency for Metering-In Speed Control Using MATLAB/SIMULINK. Proceedings of 2017 International Conference on Hydraulics and Pneumatics 120-129

Chapter 4

Valves

Characteristic and oscillation tendency study for different seat geometries of the pilot stage of a two-staged pressure control valve

peer reviewed

Martin Kloetzer[12]*, Jürgen Weber[1],

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*
[2] *Rausch & Pausch / Albert-Pausch-Ring 1, 95100 Selb, Germany*
** Corresponding author: Tel.: +49 281/67339; E-mail address: mkloetzer@rapa.com*

ABSTRACT

Two-staged pressure control valves are used in various applications. The usage in semi-active shock absorbers leads to several advantages compared to orifice controlled proportional valves. However, the high dynamic operation and limited assembly space as well as an oscillation tendency requires a detailed understanding and precise layout of the valve, especially the pilot stage. This paper presents a research study of different seat geometries of the pilot stage of a two-staged pressure control valve concerning function parameters and oscillation tendency.

The pilot stage operates as a pressure relief valve with the function to achieve a specific pressure at the pilot seat. The relevant equations to describe the function is the balance of forces combined with the equations of the individual factors (e.g. flow-force). The derivation and determination is done analytically. Just a few parameters, e.g. hydraulic resistance characteristic, will be determined by numerical CFD-simulations and measurements. The relevant description is the pressure-flow-stroke-characteristic (p-Q-x). In general, the study is verified by comparison with measurements and 1D-system simulation.

For pilot stage valves as well as pressure relief valves, there are different possible geometries for the seat and the sealing edge (e.g. ball or cone poppet seat). The various geometries have different advantages and disadvantages regarding performance, stroke dependency, dynamic behaviour as well as component complexity and robustness. This paper demonstrates the impact on different poppet seat geometries regarding the listed factors above. This provides the following benefits: First, simple overview and comparison of various poppet valve concepts regarding basic function. Secondly, a detailed influence analyses of specific design parameters. Additionally, this study can easily be used or extended by the user due to the analytical approach. Finally, this study shows and explains the fundamental functionality and dependencies. This increases the knowledge and application opportunities by helping to design poppet valves.

Keywords: Design Process, Fluids & Tribology, Characteristic and oscillation tendency study, pressure relief valve, geometry-function study of poppet valve design

1. INTRODUCTION

Semi-active shock absorbers are becoming more common within the automotive world. The main requirements on the pressure control valve of the shock absorbers are pressure control accuracy, high dynamic operation and limited packaging. In general, there are two valve types available: direct and pilot operated proportional valve. Direct operated valves or orifice-controlled valves have the advantage of fast dynamics, but the control accuracy of the main piston depends on a high number of

dependencies and component tolerances. This disadvantage is one of the benefits of pilot staged pressure control valves, the control accuracy of the main piston has less dependencies. However, this valve type is more complex in the layout and tends to oscillation at unfavourable operational areas. In sum, two-staged pressure control valve provides advantages and performance once a robust layout without achieved. Therefore, a detailed knowledge and understanding of the relevant characteristic and performance parameters and dependencies are required as well as a comprehension about the valve oscillation tendency.

The two-staged pressure control valve can be separated into 3 sections: main stage, pilot stage and electro-magnetic actor. The main oil flow goes through the main stage. Thereby, the pressure drop will be achieved and controlled. The main piston is not mechanically activated. It is moved by the pressure differences of inlet pressure and counter-control pressure of the pilot stage. This counter-control or pilot stage pressure is created by the pilot stage together with the electro-magnetic actuator, depending on the inlet pressure and the applied current. There are many design options (Valve seat geometries) for the pilot stage of a two-staged pressure control valve available. This paper compares the characteristics of different geometries with an analytical approach and evaluates the oscillation tendency.

2. MECHATRONIC SYSTEM OF THE PILOT STAGE

The pilot stage of a two-staged pressure control valve acts as a pressure relief valve depending on the applied current of the electro-magnet. This system can be described as a mechatronic system and the interference and dependencies can be seen in **Figure 1**:

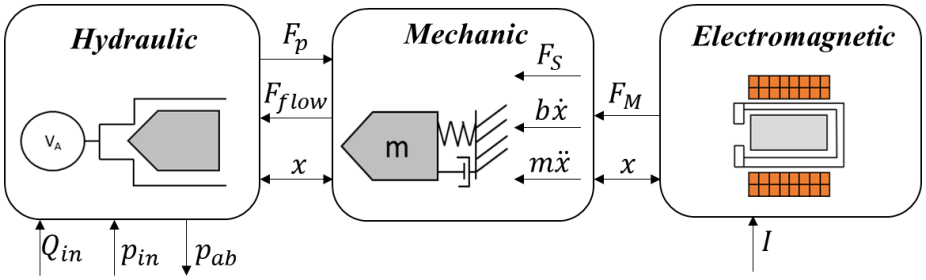


Figure 1: Mechatronic system of a pilot stage

The subsystems of the mechatronic system are Mechanic, Hydraulic and Electro-magnetic. The interaction of those subsystems provides the functionality of the pilot stage: to achieve a certain pressure drop at the valve seat depending on the applied electric current. In the following sections the system equations will be described.

2.1. Hydraulic subsystem

The hydraulic system of the pilot stage is typically a valve seat or metering edge with a chamber volume in front of the seat. As mentioned before, different geometries are possible and will be analysed and compared in this paper. These various geometries define the static and dynamic performance. **Figure 2** shows the different geometries, which will be evaluated:

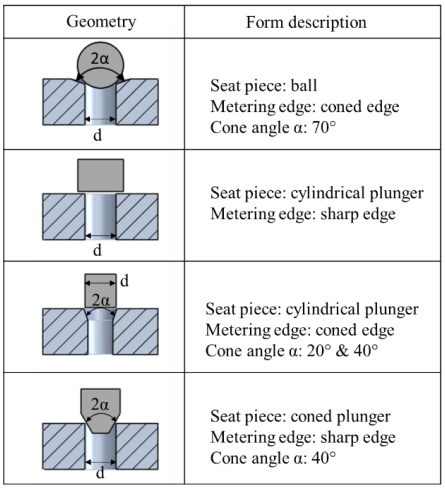


Figure 2: Different valve seat geometries

For this paper, initial pressure area and therefore the diameter d is considered as identical for all geometries. Due to the pressure in the pilot stage inlet chamber, a pressure force on the valve seat piece results, which can be described as in equation (1):

$$F_p = A_p p \tag{1}$$

The pressurized area is dependent on the geometry. For some geometries, the area and stroke interference are a square equation. However, as described in [1], the stroke x for those valves is typically much smaller than seat diameter d. Therefore, the square term is negligible and leads to a linear description of the pressurized area depending on the stroke x:

$$A_p = G_1 + G_2 x \tag{2}$$

Hereby, G_1 is the pressurized area at the initial position ($x = 0$) and can be calculated with:

$$G_1 = A_p(x = 0) = \frac{\pi}{4} d^2 \tag{3}$$

The parameter G_2 describes the linearized factor for the change of the area depending on the stroke, referring to the equation (4):

$$G_2 = \frac{A_p(x) - A_p(x = 0)}{x} \tag{4}$$

The same procedure is applicable for the flow area, which is the cross section between the metering surface and the seat part surface (e.g. coned plunger surface). Therefore, the quadratic term of the stroke x can be neglected and the equation simplified to:

$$A_{flow} = \pi d \sin(\alpha) x = G_3 x \tag{5}$$

This flow area is relevant regarding the flow forces. Those occurs due to the static pressure because of the velocity or the change of the volume flow. This connection is described by Bernoulli's equation and impulse balance [2] which leads to the following equation for the flow force:

$$F_{flow} = k_{GF} \frac{\rho Q^2}{A_{flow}} \tag{6}$$

As mentioned earlier the flow area is dependent on the valve seat geometries. Same applies for the flow coefficient k_{GF}. Typically, the approach is to minimize the occurring flow forces. In this approach the transient term of the flow force is neglected.
Another relevant influence factor is the hydraulic resistance characteristics of the valve seat. Generally, the pressure difference due to a hydraulic resistance is described by the equation (7):

$$\Delta p = \zeta \frac{\rho}{2{A_{flow}}^2} Q^2 \tag{7}$$

Therefore, the resistance coefficient ζ is dependent on several factors (e.g. the viscosity of the fluid, or flow Q) and is usually in the literature characterized as function of the Reynolds number Re and the two constants K_1 and K_2 (8). As seen in the equation (9) of the Reynolds number Re, the influence of the viscosity and therefore the temperature is included.

$$\zeta(Re) = \frac{K_1}{Re} + K_2 \tag{8}$$

$$Re = \frac{Q\ d_h}{A_{flow}\ \nu} \tag{9}$$

The hydraulic diameter d_h is supposed to be proportional to the stroke x [1] and can be described with the perimeter of the flow area P, by the equation (10):

$$d_h = \frac{4\ A_{flow}}{P} \tag{10}$$

The three geometrical factors G_1, G_2 and G_3, mentioned in the equations (3, 4) and (5) as well as ζ will be compared to each other in this paper.
Inserting the equations (8), (9) and (10) into (7) results to the familiar pressure equation:

$$\Delta p(x, Q) = \frac{K_1 \rho \nu}{2 d_h(x) A_{flow}(x)} Q + \frac{K_2 \rho}{2{A_{flow}(x)}^2} Q^2 \tag{11}$$

The resulting pressure difference is mainly dependent on the stroke x and the flow Q. Additionally, the laminar (linear Q-share) and the turbulent (square Q-share) is quite obviously in the equation (11).
Another relevant equation is the pressure build-up description within the inlet volume V_A:

$$\dot{p} = \frac{K}{V_A}\left(Q_{in}(t) - Q_{ab}(t)\right) \tag{12}$$

2.2. Electromagnetic subsystem

Pressure control valves in both versions, direct and piloted controlled, are typically activated by an electromagnetic actuator. Sometimes a reset spring is also implemented into the system. The function is to provide an actuating force on the main piston or the pilot valve seat. The electromagnetic characteristic is very depending on the stroke and remanence gap to the pole part. This is described by the force-stroke curve and is dependent on the layout and design of the magnetic circuit system. It could also be adapted with a compression spring. In general, the overall magnetic force stroke gradient cannot be changed. Therefore, just the operational area of the magnetic actuator should be evaluated and adapted. However, there are three different causes for the operational area: Negative (1), horizontal (2) and positive (3) magnetic force gradient over stroke, which are shown in **Figure 3** in a linear schematic model:

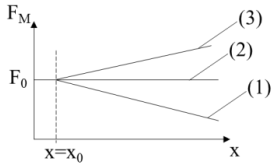


Figure 3: Schematic (linear) model for the magnetic force over stroke curves

The three versions are different in the design and layout. Usually, the magnetic force gradient is manipulated by the magnetic conductance through a pole stage. Considering a starting point at the operation range of the electromagnetic actor at $x = 0$, the initial magnetic force is F_0. The curve of the magnetic force is presumed as a linear behavior, which results in the following equation (13):

$$F_M(I, x) = F_{0,M}(I, x = 0) + c_M x \tag{13}$$

Therefore, the different gradients of raising or falling curve will be described through the sign of c_M. Electromagnetic hysteresis and friction effects are not taken into account. The design with a negative gradient provides the least complex design. This could lead to a commercial benefit in industrial application. The impact on the performance and oscillation tendency will be analyzed within this paper.

2.3. Mechanic subsystem

In some pressure relief valve and also pilot stage system a compression spring for a supporting force and reset-function is used. In this case the spring force and the spring stiffness have a similar behavior as the magnetic force:

$$F_S(x) = F_{0,S}(x = 0) + c_S x \tag{14}$$

Usually in those valve applications, the spring force is the opposite direction of the magnetic force. Therefore, both equation (13) and (14) can be combined into (15):

$$F_{Res}(I, x) = F_{0,Res}(I, x = 0) + (c_M + c_S)x \tag{15}$$

In addition to the spring force, friction and inertial forces are part of the system and need to be considered.

2.4. Complete dynamic mechatronic system

Summing up the shown and explained equations and relations results into the following force balance of the complete dynamic mechatronic system (16):

$$m\ddot{x} + b\dot{x} + F_0 + (c_S + c_M)x + k_{GF}\frac{\rho(Q)^2}{G_3 x} - (G_1 + G_2 x)p = 0 \tag{16}$$

The relevant force scheme with the example of a ball valve seat can be seen in **figure 4**:

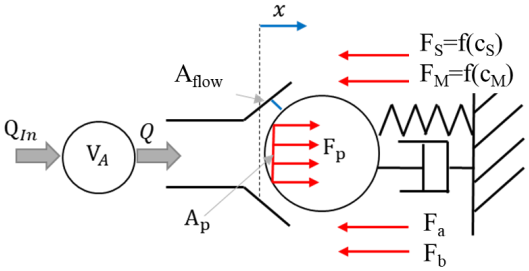


Figure 4: Schematic physical force and influence parameter model of the pilot stage seat

The focus of the paper is the comparison of various geometries and their influence on the mentioned function equation. Hereby the analytical derivation will not be shown in detail and based mainly on [1, 3, 4, 5]. For the following approach the system of the three equations, (16) together with (7) and (12) will be used. Therefore, in order to solve the initial valve problem. following initial conditions are used [1]:

$$p(t=0) = p_0 \tag{17}$$

$$x(t=0) = 0 \tag{18}$$

$$\dot{x}(t=0) = 0 \tag{19}$$

2.5. Stability evaluation by Routh-Hurwitz method

As mentioned before, the stability behavior of pressure relief valves should be evaluated. A very effective method is the Routh-Hurwitz-method for differential equations in the third order [6]. The analytical derivation is based on [1] and will not be shown in detail in this paper due to limited space. Linearization and assumptions are used in the same way. Only the main steps are mentioned below. The force balance (16) is solved for the pressure:

$$p(x) = \frac{m}{G_1}\ddot{x} + \frac{b}{G_1}\dot{x} + \left(\frac{c_S + c_M}{G_1} - \frac{F_0 G_2}{{G_1}^2}\right)x + \frac{2F_0 G_3 k_{GF}}{{G_1}^2}\frac{1}{\zeta(x)}x + \frac{F_0}{G_1} \tag{20}$$

The hydraulic resistance characteristics is kept general for now with $\zeta(x)$. The common simplification of assuming a turbulent behavior is not applied [1]. The equation (20) can be inserted in the pressure build-up equation (12) and describes the reaction of the system to the inlet flow $Q_{in}(t)$:

$$Q_{in}(t) = \frac{V_A m}{G_1 K}\dddot{x} + \frac{V_A b}{G_1 K}\ddot{x} + \frac{V_A}{K}\left(\frac{c_S + c_M}{G_1} - \frac{F_0 G_2}{{G_1}^2} + \frac{2F_0 G_3 k_{GF}}{{G_1}^2}\frac{1}{\zeta(x)}\right)\dot{x} + \sqrt{\frac{1}{\zeta(x)}}\sqrt{\frac{2F_0}{\rho G_1}}G_3 x \tag{21}$$

Considering small strokes, the simplification of $\dot{\zeta}(x) = 0$ is applied. Substituted with simple coefficients, the equation can be modified for easier usage of the Routh-Hurwitz method to:

$$Q_{in}(t) = a_3\dddot{x} + a_2\ddot{x} + a_1\dot{x} + a_0 x \tag{22}$$

The Routh-Hurwitz method implies that a system is stable once all coefficients and determinants of the system matrix M (23) are positive [6].

$$M = \begin{pmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{pmatrix} \tag{23}$$

The coefficients a_0, a_2 and a_3 cannot become negative considering using valid physical values. The coefficient a_1 can be negative and depends on following parameters:

$$a_1 = (c_S + c_M)G_1 - F_0 G_2 + 2F_0 G_3 k_{GF}\frac{1}{\zeta(x)} \tag{24}$$

Analyzing and evaluating the influence parameters, initial force F_0 and G_1 defines the pressure level and is usually an application related given value. Both parameters c_S and c_M describe the modification of the initial force F_0 regarding the stroke x (e.g. magnetic force gradient over stroke). Additionally, the parameters G_2, G_3, k_{GF} and ζ are dependent on the geometry and will be compared in this paper. An extremely interesting relation of this stability criteria and the general pQ-characteristic of those valve type is mentioned in [1] and is evidently visible, when comparing equation (24) with the following pQ-equation (25):

$$\frac{\delta p}{\delta Q} = \frac{\sqrt{\frac{\rho}{2F_0}}}{G_1^{\frac{3}{2}} G_3 \sqrt{\frac{1}{\zeta(x)}}} \left((c_S + c_M) G_1 - F_0 G_2 + 2F_0 G_3 k_{GF} \frac{1}{\zeta(x)} \right) \tag{25}$$

Reviewing both equations (24) and (25), the same term determines the slope and the sign of the equation. Therefore, the simple statement that the pQ-characteristic must be raising in order to achieve valve stability can be determined [1].

The second critical criteria of the Routh-Hurwitz method is the determinant D_2, which may become negative:

$$D_2 = \frac{V_A}{K} \left(\frac{c_S + c_M}{G_1} - \frac{F_0 G_2}{{G_1}^2} + \frac{2F_0 G_3 k_{GF}}{{G_1}^2} \frac{1}{\zeta(x)} \right) \frac{V_A b}{G_1 K} - \left(\sqrt{\frac{1}{\zeta(x)}} \sqrt{\frac{2F_0}{\rho G_1}} G_3 \right) \frac{V_A m}{G_1 K} \tag{26}$$

Especially the hydraulic damping b is an interesting influence parameter, which describes the minimum required hydraulic damping to achieve a stable system [1].

3. VALVE SEAT GEOMETRIES PQX-STUDY

In this section the different geometries are explained and discussed. As introduced and shown in **Figure 2**, the characteristics of five different valve seat configurations (geometry and angle) are evaluated. Following parameters that are listed in **table 1** will be used for illustration purposes of the calculations:

Table 1: List of parameters

Parameter	Value
Valve seat diameter d	2.7×10^{-3} [m]
Density ρ	850 [kg / m]
Mass m	25×10^{-3} [kg]
Bulk modulus K	14000 [bar]
Inlet chamber volume V_A	10×10^{-3} [m^3]
Flow coefficient k_{GF}	0.35 [-]
Initial resulting Force F_0	25 [N]

3.1. Geometrical parameters

Initially, the geometrical parameters of the different geometries G_1, G_2, and G_3 are calculated and compared. The first geometrical influence parameter is G_1, which can be calculated with equation (3). Considering the same resulting force F_0, the same pressure operation level and the same diameter d for all geometries, G_1 is identical. For the parameter G_2 the equation (4) can be used. The different behavior for G_2 of the designs can be seen in **Figure 5**:

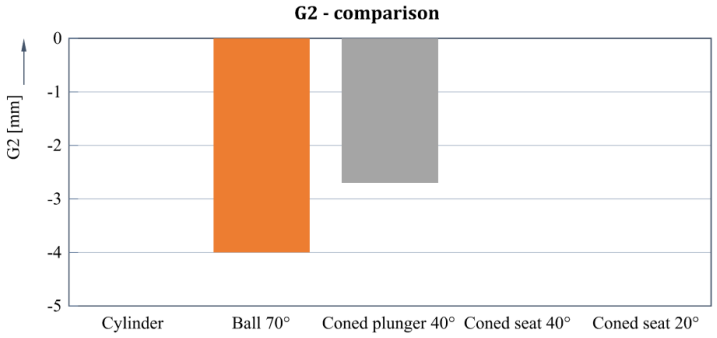


Figure 5: G_2 comparison of the various geometries

Cylinder and coned seat geometries are independent of the stroke. The ball and coned plunger variant decrease the pressurized area with an increasing stroke. The third geometrical parameter G_3 can be calculated with (5) and describes the change of the flow cross section based on the stroke x.

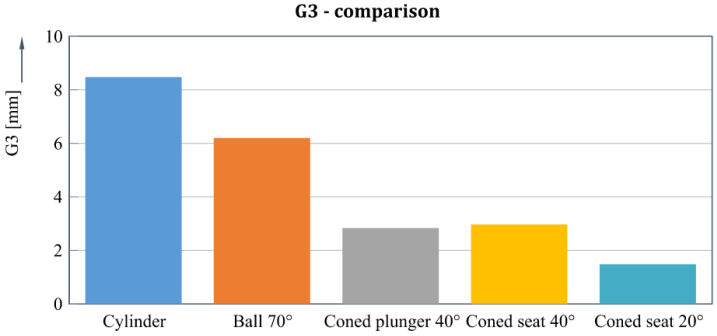


Figure 6: G_3 comparison of the various geometries

Again, the differences between the various geometries are visible.

3.2. PQx-characteristics of the different geometries

The pQx-characteristic is the first step to evaluate and compare the function behaviour of the different geometries for the pilot stage. It shows the resulting pressure difference before and after the pilot stage depending on the flow Q and stroke x. In general, there are four different ways to determine the pQx-curves: CFD-Simulation, measurement on prototypes, via analytical calculation or system-simulation. The last two approaches require reference values for $\zeta(x)$, therefore those are just applicable for recalculation or adaption of existing experience values.
For this paper, CFD-Simulations are used. Below in **Figure 7** the pQx for the ball valve seat is depicted as an example.

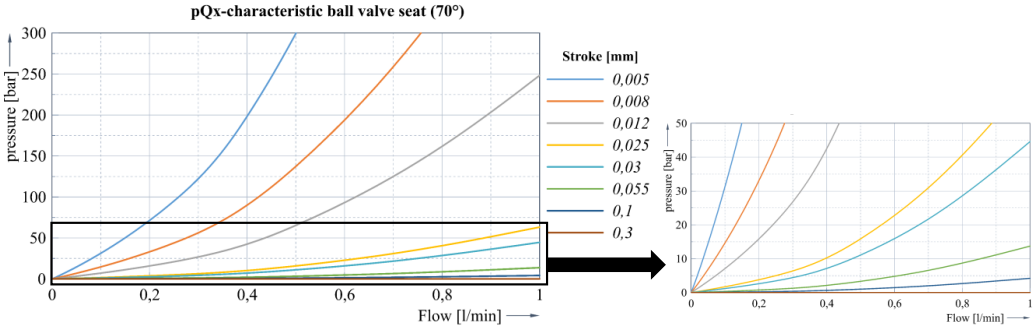


Figure 7: PQx- characteristic for ball valve seat (CFD-Simulation)

Interesting is the comparison of the characteristics for the different geometries. For illustrative purposes following operation point is used: $Q = 0.1\frac{l}{min}$ and $x = 0.1mm$:

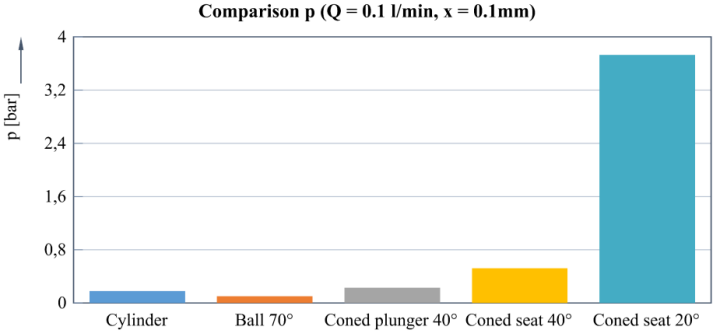


Figure 8: $p\left(Q = 0.1\frac{l}{min}, x = 0.1mm\right)$ comparison

Secondly, the operation point of $Q = 0.5\frac{l}{min}$ and a stroke of $x = 0.03mm$ is used for comparison and shown in the **Figure 9** below:

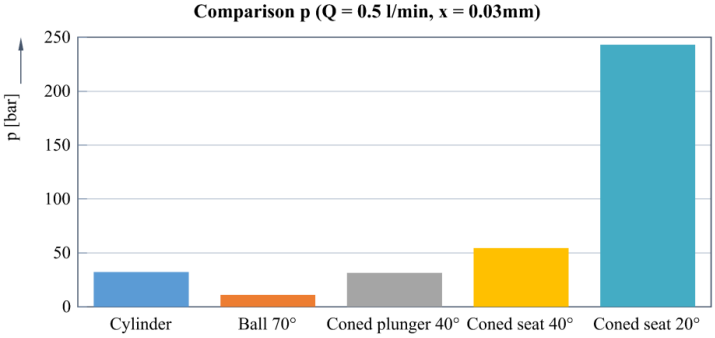


Figure 9: $p\left(Q = 0.5\frac{l}{min}, x = 0.03mm\right)$ comparison

It is visible that the coned seat design, especially with the 20°-angled cone achieve a quite high pressure drop.

3.3. Hydraulic resistance

In the case of a new investigation or new design, the $\zeta(x)$ must be approximated from simulated or measured pQx-characteristic. Therefore, the equation (7) can be used by solving it for the hydraulic resistance $\zeta(x)$ and converting into a function of the Reynolds number $Re(x)$ (27):

$$\zeta(x, Re) = \Delta p(x) \frac{2}{\rho} \frac{A_{flow}(x)^2}{Q^2} = \frac{K_1(x)}{Re(x)} + K_2(x) \tag{27}$$

Below the characteristics of the $\zeta(x, Re)$is shown for the examples of the cylinder (**Figure 10**) and ball (**Figure 11**). First, the laminar (linear) and the turbulent (square) share is quite obviously in the double-logarithm display. Second, the difference between both geometries in terms of quantity of the hydraulic resistance. Third, the fluctuation regarding the stroke is visible. The last mentioned, is important to consider, because the calculation of one operation point is not equal to the complete operation range and results can be different, especially with a big value band of $\zeta(x)$ (e.g. cylinder geometry).

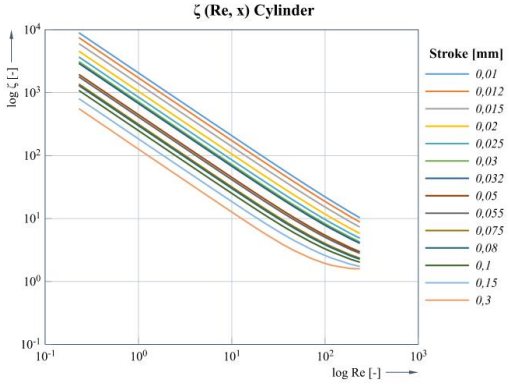


Figure 10: $\zeta(Re, x)$ for the cylinder valve seat geometry

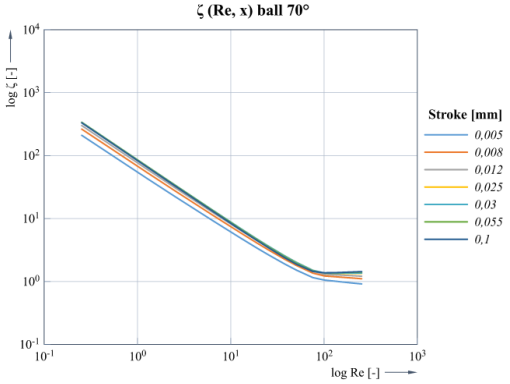


Figure 11: $\zeta(Re, x)$ for the ball valve seat geometry

Again, the results for the different geometries should be compared. First, the comparison of the hydraulic resistance is depicted at the operation point of $Q = 0.1 \frac{l}{min}$ and a stroke of $x = 0.1mm$ (**Figure 12**):

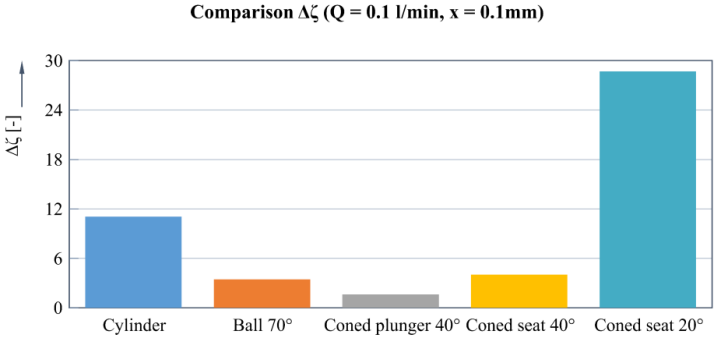


Figure 12: $\zeta(Q, x)$ comparison for the various geometries

Secondly, the width of the band of ζ depending on the stroke at the point of a Reynolds number of 10 is evaluated (**Figure 13**) and shown in the pictures below:

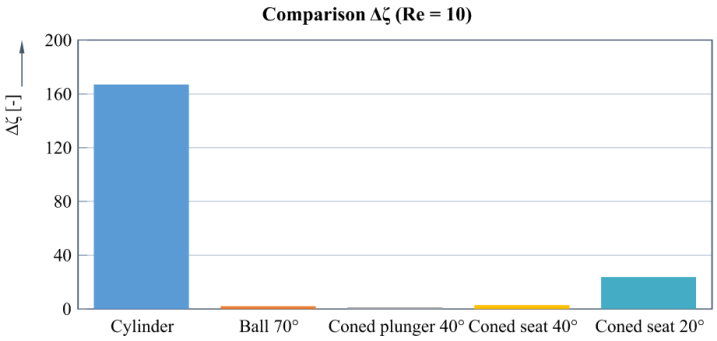


Figure 13: $\zeta(Re = 10)$-width comparison for the various geometries

The cylinder shows a big band in difference to ball and coned plunger, which shown almost no stroke dependency.

From here on, a stroke of $x = 0.1mm$ will be used as reference and boundary condition for further investigations. Therefore, the different values for $K_1(x = 0.1mm)$ and $K_2(x = 0.1mm)$ for the geometries can be determined and compared.

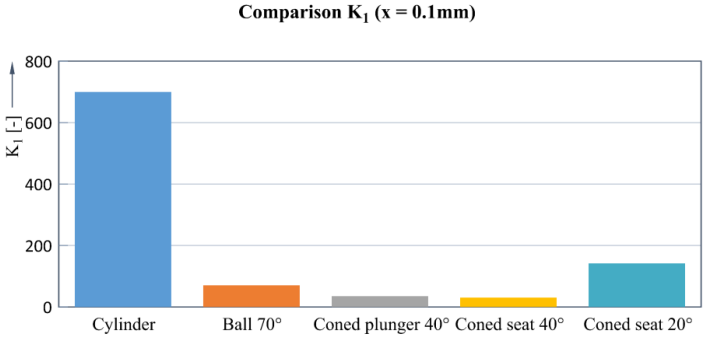


Figure 14: K_1 comparison for the various geometries

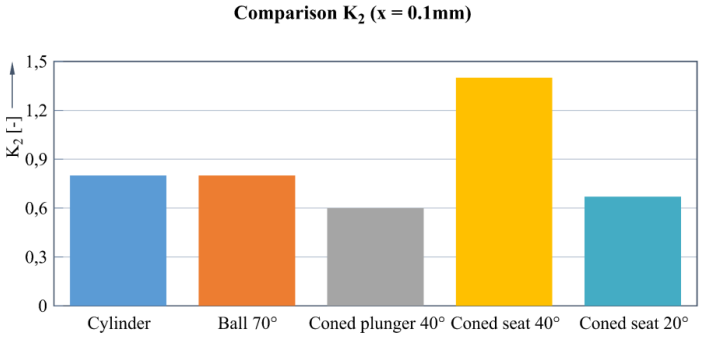


Figure 15: K_2 comparison for the various geometries

3.4. Validation

While using this reference point, an analytical calculation of the pQx-behaviour can be done. This bases on the, from the CFD-approximated, hydraulic resistance characteristic. Due to the limited space, just the pQx-characteristic of the ball seat geometries will be presented in this paper as an example. In the following **Figure 16**, the four curves from the different methods are shown:

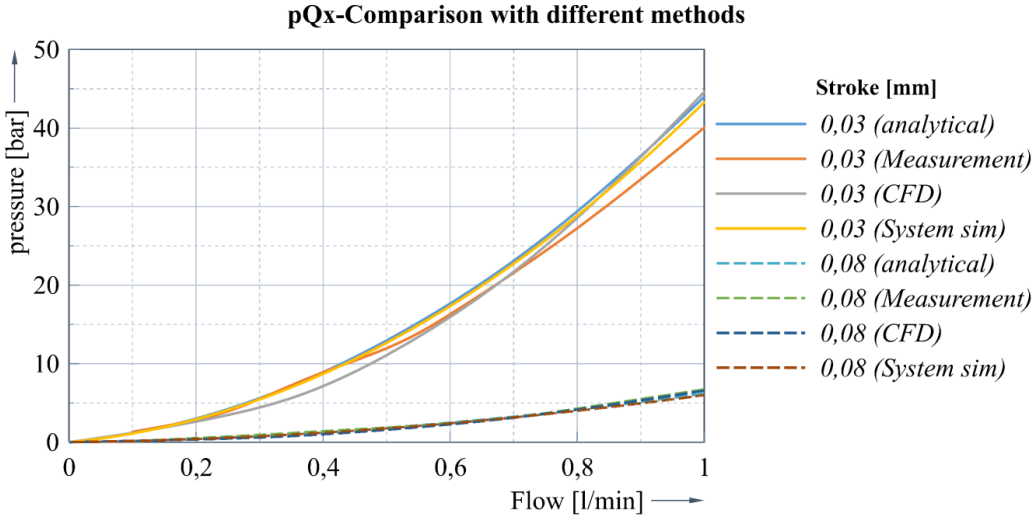


Figure 16: pQx comparison for different methods

The measurements are performed on a hydraulic test bench with a simplified test setup. The stroke of the valve is manually adjusted to the target value for each measurement point. Afterwards the resulting pressure drop based on the flow is measured statically for each operational point. Comparing the curves, all four methods show similar results. That means that the "by-hand" calculation together with approximated values for ζ is accurate enough and can simplified to speed up variant derivations and estimations, if the principial behavior is not changed too much for the reference point.

4. OSCILLATION TENDENCY STUDY

In section 2.5 the stability evaluation method was described with the two requirements that a_1 and D_2 must be positive. Now, this will be used to compare the different geometries by quantifying the oscillation tendency for a pressure relief valve.

First, we will have a look at the a_1 criteria. Reviewing equation (24) the influence parameters are clear and visible. Especially the influence of the amount and the sign of the investigated geometrical parameters can be seen. Examples that have a negative G_2 or a small hydraulic resistance are beneficial. As shown in the section before, there are significant differences between the geometries.

The only undefined parameters within the system and the equation (24) are the stiffnesses $(c_S + c_M)$. Those depends on the design of the electromagnetic actuator and the mechanic system. The magnetic force gradient is in principle negative; however, it can be tuned by design features (e.g. pole stage) to a horizontal or even positive curve. Those design options require a detailed magnetic field layout, usually done with FEM magnetic force simulations. Additionally, those tuning actions lead to more complex components and higher tolerance dependences, especially relevant for high-volume industrial applications. The sum of the resulting force of the electromagnetic can be adapted by a compression spring.

Initially, the basic version without any pole stages or springs and a magnetic force gradient of -40N/mm will be used as an example. The resulting values for a_1 are shown in **Figure 17**:

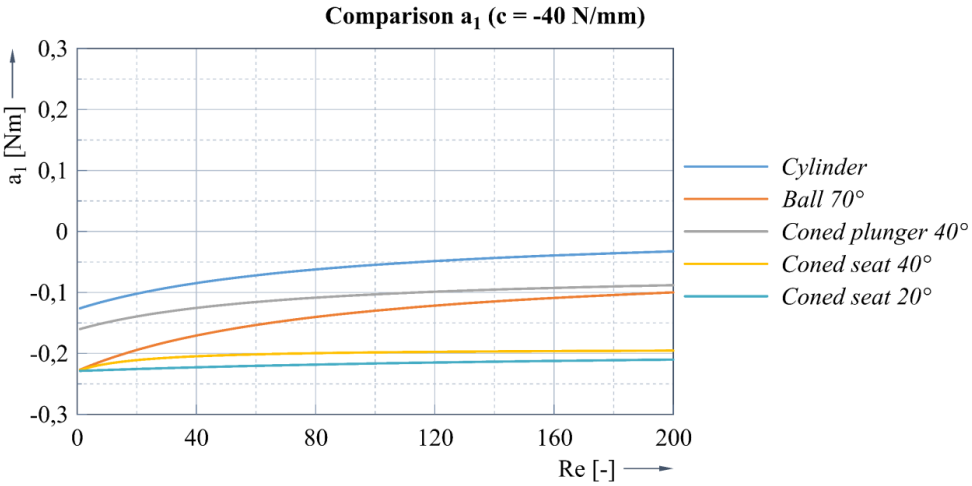


Figure 17: $a_1(c = -40\frac{N}{mm})$ comparison for the various geometries

As depicted the results for of all geometries a_1 is negative, which would mean that valve configuration would tend towards oscillation. However, different values for the geometries can be detected. For more precise research, equation (24) can be solved for the stiffness and with that a critical force gradient – the minimum required valve to fulfill the stability criteria can calculated and is depictured in **Figure 18**:

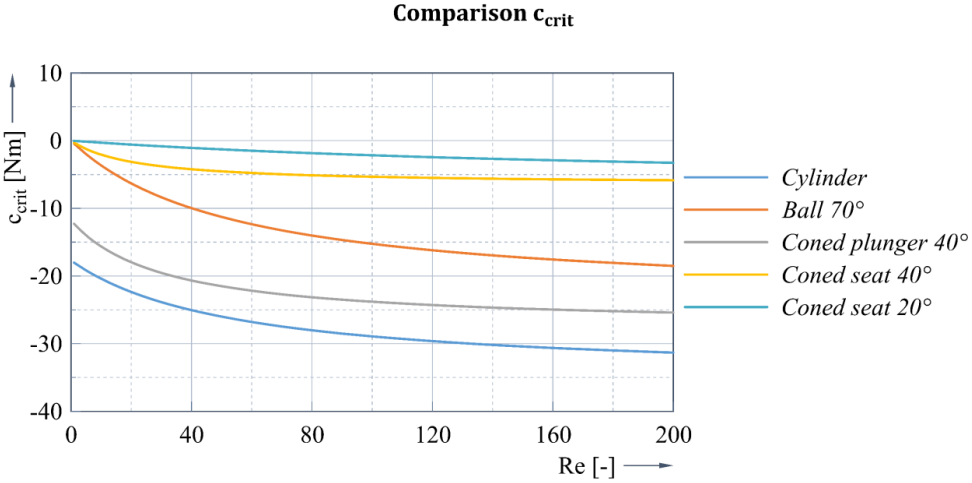


Figure 18: c_{crit} comparison for the various geometries

The ball geometry requires the lowest and the coned geometries the highest minimum magnetic force gradient to fulfill the criteria ($a_1 > 0$). This could be a significant influence parameter to evaluate the layout of electromagnetic actuators, especially in regards of design complexity (e.g. pole stages). Considering the criteria fulfillment for all the geometries at least a horizontal gradient is required ($c_S + c_M = 0$). This kind of parameter influence study can be done extremely fast and easy with the analytically approach, while using the calculation template with. The results are shown in the **Figure 19** below:

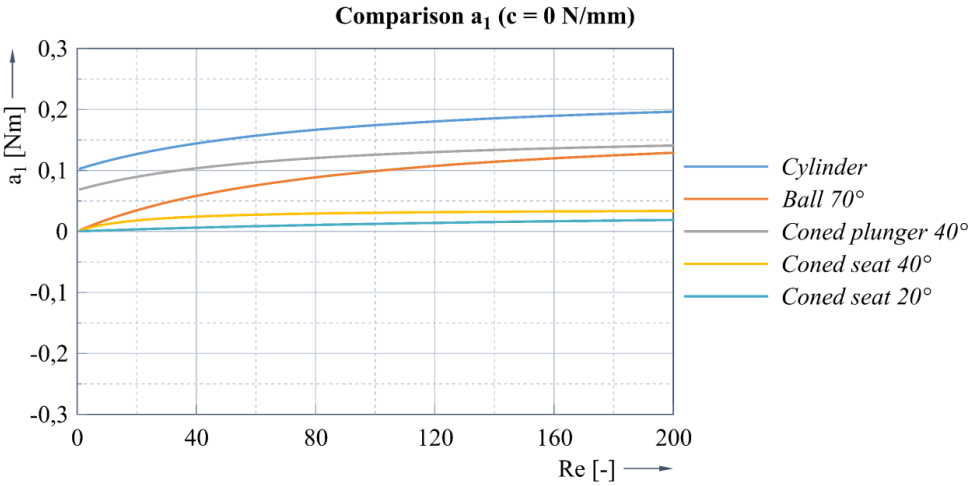


Figure 19: $a_1(c = 0\frac{N}{mm})$ comparison for the various geometries

This time, as expected, all geometries fulfill the first stability criteria. Now, the second criteria D_2 can be evaluated. Hereby, the viscous damping b is a very important parameter. Usually, the damping should be rather small in order to provide good dynamic valve behavior (e.g. valve shift time). At the same time, reviewing equation (25) shows a minimum viscous damping is required to avoid valve oscillation. Therefore, the equation can be solved for b which leads to a description for the critical damping b_{crit}:

$$b_{crit} = \sqrt{\frac{2F_0}{\rho G_1}\frac{K}{V_A}}\sqrt{\frac{1}{\zeta}}\frac{mG_1{}^2G_3}{(c_S + c_M)G_1 - F_0G_2 + 2F_0G_3k_{GF}\frac{1}{\zeta}} \tag{28}$$

Using equation (28) the required damping for all geometries can be calculated and compared. Additionally, the influence parameters can be seen, which can help to optimize the system. The different achieved results for the geometries are depictured in **Figure 20** below:

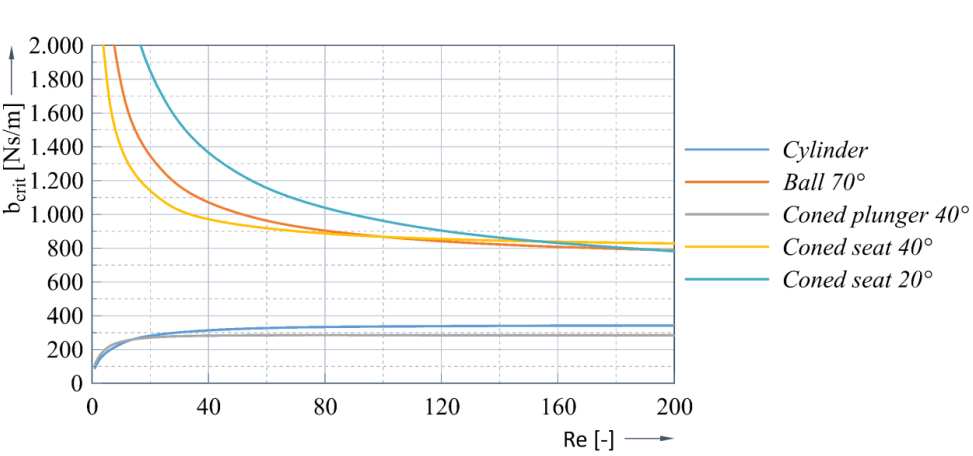


Figure 20: b_{crit} comparison for the various geometries

5. CONCLUSION AND OUTLOOK

In the first section of this paper, the analytical derivation for the different sub-systems of a pilot stage for a two-staged pressure control valve are explained. Several simplifications in form of e.g. linearization of the pressurized area are presumed. The various possibilities for the force-stroke behavior of the electromagnetic actuator are shown and used as influence study later in this paper. Based on the explained equations of the subsystem, the dynamic behavior of the pilot stage is achieved and is described by the force balance (16), pressure equations depending on stroke (20) and finally the differential equation third order of the system (21). The main approach of this paper is to describe and evaluate different possible valve seat geometries.

First, the various characteristics of the geometries were analyzed by the pQx-behavior. Therefore, four different methods can be used: Analytically, CFD-Simulations, 1D-Sytstem simulation, or measurements on samples. However, the goal was to create a manual way with an analytical approach to calculate the pQx-behavior with an acceptable deviation which was achieved. The required geometrical influence parameters can be calculated and compared with the equations (7). The hydraulic resistance is the only parameter which cannot be directly calculated. It was shown that it can be quantified with the equation (27) from CFD-simulation or measurements. The resulting pQx-characteristics of the various geometries shows the different behavior, especially regarding a laminar or turbulent flow. Depending on an industrial usage of the valve, advantages, e.g. regarding higher viscosity influence and with that more temperature dependences can be determined and geometries chosen. The validation of the calculated results has been done by comparing 1D-system simulations, CFD- simulations and test bench measurements.

Secondly, the same geometries of the pilot stage valve seat should be evaluated regarding oscillation tendency. With the explained Routh-Hurwitz method based on equation (21) two main criteria for a_1 (24) and D_2 (26) are shown and the results are depicted. At this point, the analytical approach and the criteria of the equation (24) and (28) help significantly to identify and understand the impact parameters and to evaluate improvements. One further identified influence parameter is the resulting force gradient over stroke, which mainly depends on the magnetic layout. Thereby, the direct relation from the gradient to the function and the oscillation tendency was discovered and verified. This investigation could be extended to focus on the relationship between magnetic force gradient of the actuator and the influence on the performance.

Summarizing this paper, the analytical approach shows in detail the influence parameters which help to understand the performance and the dependencies. This can be used to understand this valve type and should support new developments and improvements on existing applications. In addition, this approach can be repeated by users and a calculation template can be created in a similar way. In general, this investigation should be valid for pilot stages as well as pressure relief valve considering similar boundary conditions.

An extremely important parameter was not considered in this paper: temperature. All the calculations, simulations and measurements have been done at a constant temperature. Especially with a laminar dominant valve design and operation, the temperature has a significant influence due to the viscosity dependency. This approach, which was used and explained in this paper, could be extended by the additional dimension of the temperature. Thereby, different geometrical dependencies for the temperature, overall performance and oscillation tendency could be discovered and quantified.

NOMENCLATURE

A_{flow}	Flow cross section	m^2
A_p	Pressurized area	m^2
a_0, a_1, a_2, a_3	Stability coefficients	
b	Hydraulic damping	Ns/m
b_{crit}	Critcal hydraulic damping	Ns/m
c_{crit}	Critical force gradient	N/m

c_M	Magnetic force gradient	N/m
c_S	Spring stiffness	N/m
d	Diameter	m
d_h	Hydraulic diameter	m
D_0, D_1, D_2	Determinants	
F_{flow}	Flow force	N
F_M	Magnetic force	N
F_p	Pressure force	N
F_S	Spring force	N
F_0	Initial force	N
G_1,	Geometrical parameter	m^2
G_2, G_3	Geometrical parameter	m
I	Electrical current	A
k_{GF}	Flow force coefficient	-
K	Bulk modulus	bar
K_1, K_2	Hydraulic resistance constants	-
M	Routh-Hurwitz Matrix	m^2/s
m	Masse	kg
v	Kinematic viscosity	m^2/s
p	Pressure	bar
P	Perimeter	m
Q_{in}	Inlet volume	l/min
Q_{ab}	Outlet volume	l/min
ρ	Density	kg/m^3
Re	Reynolds number	-
t	Time	s
V_A	Chamber volume	m^3
x	Stroke	m
ζ	Hydraulic resistance	-

REFERENCES

[1] Osterland S (2018) Analytical Description of the Static and Dynamic Behaviour of a Pressure Relief Valve. Symposium on Fluid Power and Motion Control FPMC2018. Bath, UK. DOI:10.1115/FPMC2018-8859

[2] Will D, (2008) Hydraulik, Heidelberg, Germany

[3] Wehner, D, (2008) Modellbasierter Systementwurf am Beispiel vorgesteuerter Druckbegrenzungsventile. PhD Thesis, Institute of Fluid Power, Dresden, Germany

[4] Erhard, M, (2016) Regelverhalten, Gestaltdesign und Robustheit direktgesteuerter Proportionaldruck-begrenzungsventile, PhD Thesis, Institute of Fluid Power, Dresden, Germany

[5] Wobben, G, (1978) Statisches und dynamisches Verhalten vorgesteuerter Druckbegrnzungsventile unter besonderer Berücksichtigung der Strömungskräfte, PhD Thesis, Institute of Mechanical Engineering, Aachen, Germany

[6] Schulz, G, (2007) Regelungstechnik 1: Lineare und Nichtlineare Regelung, Rechnergestützter Regelentwurf, München, Germany

SPONSORED BY THE
Federal Ministry of Education and Research

Finanziert von der Europäischen Union
NextGenerationEU

peer reviewed

DEDICATED DESIGN OF THE FREE JETS FOR ROTARY SLIDE VALVES

Lennard Günther*, Sven Osterland, Danny Staroszyk, Jürgen Weber

Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden
* Corresponding author: Tel.: +49 351 463-39200; E-mail address: Lennard.guenther@tu-dresden.de

ABSTRACT

The paper presents a general analytical method for the determination of the flow angel in hydraulic components like valves and pumps. In this paper it will be illustrated by a rotary slide valve.
The design of the flow geometry e.g. in valves depends on many influencing parameters. A suitable choice of parameters ensures that target variables such as flow characteristics, flow force, but also other effects such as cavitation and noise generation can be improved. An important parameter is the flow angle of the free jet, which forms in each valve directly downstream of the narrow section of the control edge. The resulting free jet influences the flow force as well as cavitation damage. By a suitable choice of the angle of this free jet, the flow force can be reduced by changing the direction of the outgoing impulse. With regard to cavitation, the impact of the free jet can be shifted and thus the cavitation erosion can be shifted or weakened. However, the flow angle is not always known and the influence of the parameter cannot be fully exploited. In the past, numerous elaborate experimental investigations or simulations have been carried out to improve influencing parameters on corresponding target sizes, as just mentioned. In many cases, the decisive factor is the flow angle of the emerging free jet.
This paper deals with the investigation of the flow angle of free jets as well as the technical application on the basis of rotary slide valves. In the first section, geometric factors influencing the flow angle are discussed, as well as the transferability of the results under varying operating conditions (laminar and turbulent). Using a generic minimal model, the behavior of the flow angle with respect to geometric influence factors and operating conditions is investigated by means of CFD. The results are adapted to the real application, a rotary slide valve. Direct adjustment of the flow angle results in a significant improvement in the resistance torque caused by the flow force. It becomes clear how efficient the adjustment of the flow angle can be if the basis of the formation of the free jet is known. Due to the derivation of the relationship with the help of an abstracted minimal model, the knowledge gained can be used in many ways and can also be transferred to other applications in the field of fluid technology. Optimization processes are more efficiently without using elaborated models e.g. driven by CFD.

Keywords: flow angle, free jet, rotary slide valve, design process

1. INTRODUCTION

The flow angle of the free jet plays a decisive role in the functionality and performance of fluid power components. Internal flows, which are common in hydraulics, have a large number of abrupt cross-sectional constrictions. At those narrow points, e.g. at control edges of valves, a free

jet is formed as the flow passes through. This free jet influences the resulting flow force acting on the valve spool [1]. In addition, the free jet has a great influence on the formation and transport of cavitation bubbles. In the shear layers of the free jet, vortices are formed in which the bubbles are created and grow [2]. The bubbles are transported with the flow and collapse if the local pressure increases again, e.g. at stagnation points. In areas close to the wall, this leads to a microjet. That causes erosion damage to the component, shown in [1] and [3]. However, free jets also occur in pumps such as the valve plate of an axial piston pump, during the transition from low pressure to high pressure, which can cause cavitation damage [4]. Depending on the geometry, the free jet aligns itself via the flow angle. The literature on specific angle positions is long. In [1], the angle for individual control positions of a spool valve is defined. Similarly, the flow angles are described for example by [5], [6], [7], [8], where all data on the angle correlate well with an acceptable scatter. In [9] and [10] the clearance at the valve spool and the sleeve are considered, which is important for very small opening ratios.

A common question during the valve and pump development process is: In what way is the flow angle affected by the control edge geometry? This paper deals with the investigation of the flow angle by using a generic minimal model. The aim is a general method about how the flow angle results from geometric parameters and in which operating range it is valid. The transfer of the results to a real application reflects the applicability of the obtained correlation. The implementation is carried out using a rotary slide valve, which is illustrated in **Figure 1**. As a special design compared to conventional rotary slide valves, this one is integrated into a cartridge installation space. The valve can be operated in both flow directions. The illustration shows for example of the flow direction the inlet horizontally and the outlet vertically on the bottom. The valve has three control edges in the horizontal plane, which are offset by 120 °. A complete opening and closing process takes place at a twisting angle φ of 0 ° (fully open) to 60 ° (closed). The aim is to optimize the flow forces by actively influencing the flow angle of the free jet.

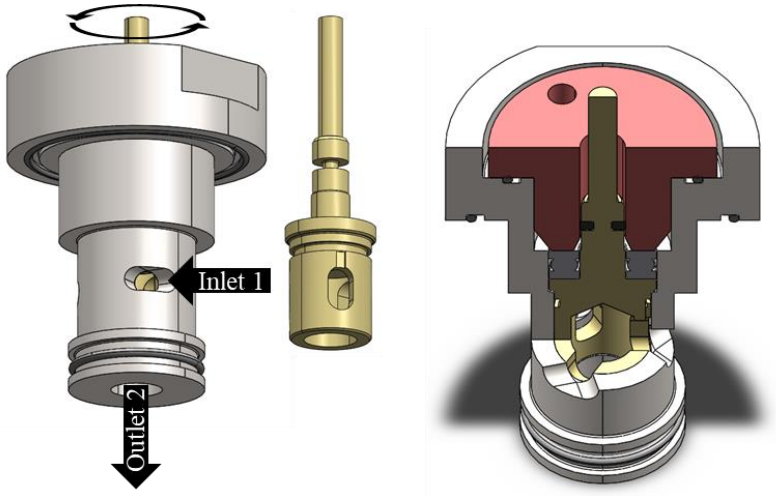


Figure 1: Rotary slide valve, exemplary flow direction 1 → 2 through the assembly (left), valve spool (middle), sectional view through the inlet plane and the vertical center plane (right)

2. METHODOLOGY

The flow force calculation can be done in different ways. A simple and in most cases effective method is the determination of the resulting force using the conservation of momentum. Another advantage is that the dependencies can be easily identified through such an analytical relationship. For a defined control volume Ω, time-dependent (I), incoming and outgoing (II) momentum, as well as pressure forces (III) and shear forces (IV) or volume-dependent forces, such as acceleration forces (V), and body forces (VI) can be taken into account via the boundary S. As shown in Eq. (1), the complete relation of the conservation of momentum is described in integral form.

$$\underbrace{\int_\Omega \frac{\partial \rho \vec{u}}{\partial t} dV}_{\text{I}} + \underbrace{\int_S \rho \vec{u}\,(\vec{u}\cdot\vec{n}) dS}_{\text{II}} = \underbrace{-\int_S p\vec{n} dS}_{\text{III}} + \underbrace{\int_S \bar{\bar{\tau}}\cdot\vec{n} dS}_{\text{IV}} + \underbrace{\int_\Omega \rho \vec{f} dV}_{\text{V}} + \underbrace{\vec{F}_B}_{\text{VI}} \tag{1}$$

For the hydraulic application concerning the rotary slide valve, the following simplification can be assumed, which reduces the expression, **Figure 2**.

Steady state: $\frac{\partial}{\partial t} = 0$

Pressure compensated: $\int_S p\vec{n} dS = 0$

Flows close to the wall are very small compared to the resulting body forces ($\vec{F}_{\bar{\tau}} \ll \vec{F}_B$) [11]: $\int_S \bar{\bar{\tau}}\cdot\vec{n} dS = 0$

Movement of the volume is neglected: $\int_\Omega \rho \vec{f} dV = 0$

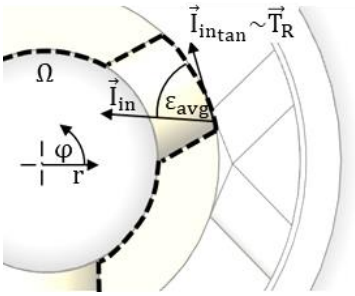


Figure 2: Application of the law of conservation of momentum on the basis of the rotary slide valve, valve spool (yellow), housing (light gray), control volume Ω (black, dashed).

This results in an equation that describes only the dynamic part of the flow. The momentum entering and leaving the control volume determine it.

$$\int_{\check{S}} \rho \vec{u}\,(\vec{u}\cdot\vec{n}) dS = \vec{F}_B \tag{2}$$

Eq. (2) can be explained in cylindrical coordinates, Eq. (3). Thus, it becomes easier to apply to the control volume of the rotary slide valve. The momentum of the outlet is perpendicular to the flow over the control edge. Thus there is no influence in the equations.

$$F_{flow} = \rho \left|\vec{u}_{in}^2\right| \cdot A_{in} \cdot \cos(\varepsilon_{avg}) \tag{3}$$

$$T_R = \int F_{flow}\, dr = \rho \left|\vec{u}_{in}^2\right| \cdot A_{in} \cdot r \cos(\varepsilon_{avg}) \tag{4}$$

Eq. (3) defines the flow force and (4) the resistance torque acting tangential on the valve spool per control edge. As already mentioned, it is easy to see that the resistance torque T_R as well as the flow force F_{flow} depend on the geometric parameters of the flow area A_{in} and the averaged flow angle ε_{avg}. Since the area cannot be used as an influencing parameter in this case because it must necessarily be opened to release the flow cross section. Only the flow angle remains to actively influence the resistance torque of the flow at the control edge / narrow section.

But how does the flow angle behave and how can it be influenced? In the following, the methodology for investigating and derivation will be explained.

Due to the lack of knowledge about the exact behavior of the flow angle, the correlation is examined in a generic minimal model by using CFD. The geometry is selected in such a way that no disturbance effects can influence the flow. The focus is purely on the variation of the flow angle. It is shown in [1], that the flow angle has a significant dependency of the inlet geometry.

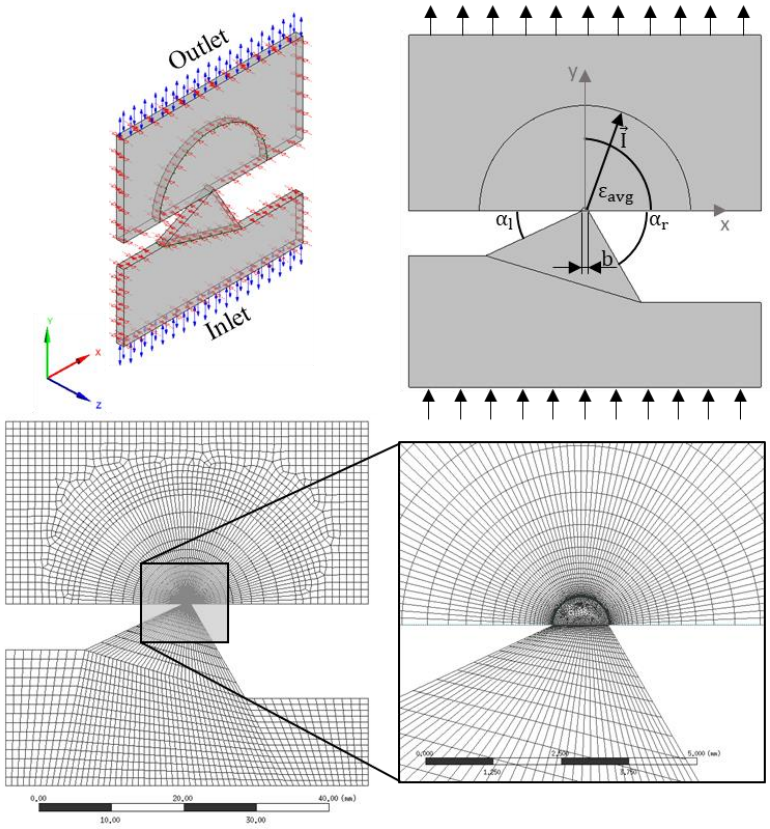


Figure 3: Overview of geometry of the model (top) and representation of meshing (overview - bottom left, bottleneck - bottom right).

Figure 3 illustrates the flow simulation model, which is used to study the flow angle. The geometry is divided into inflow area (lower region), narrow section (center) and outflow area (upper region). The inlet region is variable using the angles α_l and α_r of the left and right walls. Thus, the flow angle of the free jet in the narrow section can be changed. The narrow section has

a fixed width b of 1 mm and is arranged horizontally. The outlet region spans a free space starting from the narrow section in which the jet can spread freely. Due to many symmetrical relationships in technical geometries, a two-dimensional model is used for this investigation. The setup of the simulation model is summarized in **Table 1**. With the aid of this model, a new analytical method will be established to determine the flow angle at a narrow section very well (ch. 3.3, Eq. (13)). The correlation shows a simplified expression between the inlet geometry and the flow angle of the free jet. The comparison of a real application is made in ch. 4 by using the rotary slide valve.

Table 1: minimal model - setup

property	setting	property	setting
geometrical variation		analysis type	
α_l	$\{0 \dots 180\}°$	temporal consideration	steady state
α_r	$\{0 \dots 180\}°$	turbulence	Shear Stress Transport model (RANS)
b	1 mm	convergence	
boundary conditions		convergence criterion	10^{-4}
inlet	opening with $p_{in} = \{10^{-4} \dots 10^{2}\}$ bar	max. number of time steps	200
outlet	opening with $p_{out} = 0$ bar	fluid properties	
wall	no slip wall (geometry related to x-y-plane)	density	863 kg / m³, incompressible
symmetry	in z-direction	kin. viscosity	46 mm² / s
		temperature	40 °C, isotherm

3. FLOW ANGLE OF FREE JETS ON THE MINIMAL MODEL

This chapter deals with the study of the properties and derivation of the flow angle. For a better understanding, the following structure is used.

- general definition of the local and averaged flow angle
- laminar and turbulent influence
- analytical method for the averaged flow angle

3.1. Definition of the flow angle ε

The local flow angle ε_{loc} at a given point in the fluid domain is defined by the velocity components u (x-direction) and v (y-direction), see Eq. (5). To determine the averaged flow angle ε_{avg}, the flow angle ε_{loc} is averaged over the width b at the narrow section, as shown in Eq. (6). For the investigation, the flow angle is calculated at $y = y_0 = 0$, the position of the narrow section.

$$\varepsilon_{loc} = \operatorname{atan}\left(\frac{v}{u}\right) \tag{5}$$

$$\varepsilon_{avg} = \frac{1}{b}\int_{-b/2}^{b/2} \varepsilon_{loc}(y_0, x)\,dx \tag{6}$$

In **Figure 4**, the flow angle ε_{loc} in the narrow section over the opening coordinate x is shown. The left diagram clarifies the profile of the flow angle between the left and right edges of the narrow section. Due to the Neumann (no slip) boundary conditions along a wall, the flow angle at the edges is equal or very similar to the orientation of the wall. On the left wall the geometrical angle α_l is 25 ° and the the flow angle is about 45 °. On the right side the geometrical angle α_r is 60 ° and the flow angle is about 108 ° (related to α_r 72 °). The deviations are caused by a strong rotation of the flow in the corners of the narrow section. The averaged value of the angle is ε_{avg} =79 °.

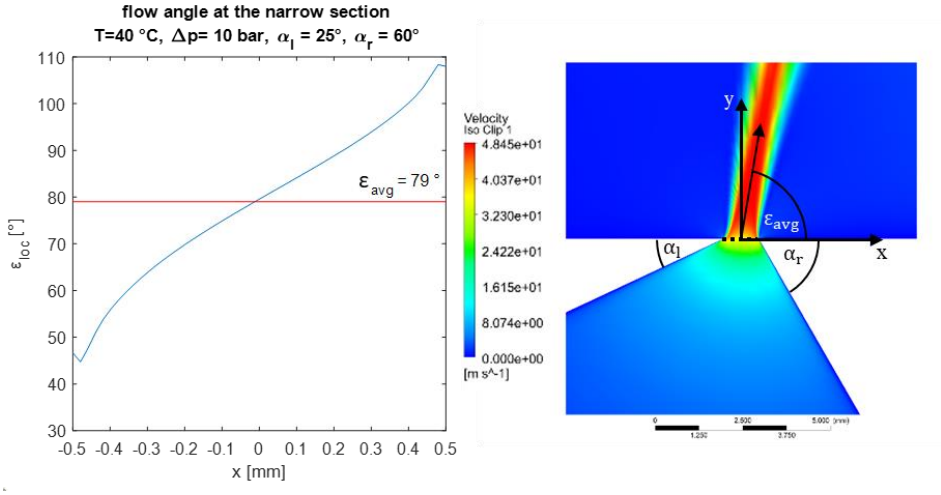


Figure 4: Flow angle ε_{loc} at the narrow section for the case: $\alpha_l = 25$ °, $\alpha_r = 60$ °, $\Delta p = 10$ bar, flow angle plotted over the opening coordinate x (left), velocity field of the free jet around the narrow section (right)

3.2. Laminar and turbulent influence

In this section the range of validity of the calculation of the averaged flow angle ε_{avg} for different Reynolds numbers is investigated. Decisive here is the consideration of a laminar and / or turbulent flow. According to [12] with reference to [13], the critical Reynolds number of a free jet is Re ≈ 30. The Reynolds number is defined here as follows.

$$\mathrm{Re} = \frac{\tilde{u} \cdot l}{\nu} \text{ with } \nu = \frac{\eta}{\rho} \tag{7}$$

The velocity $\tilde{u}$ is defined such as the characteristic length $l = b$ at the narrow section of the free jet. The Eq. (7) is illustrated by following example. A valve of nominal size 08 according to ISO 7368 is opened to 20 % and operates with a volume flow of 15 l / min. According to [14], this area of the valve is strongly in the laminar region of hydraulic resistance. The calculated Reynolds number Re = 138 is well above the critical value of 30. That means the flow is nevertheless predominantly laminar in the gap per se. Due to the superposition of laminar and turbulent flow, the resulting free jet behind the narrow section is nevertheless turbulent. It is clear there is a strictly separation between the laminar and turbulent behaviour of a hydraulic resistance and the characteristics of a free jet.

In the following, it will be shown that a turbulent free jet is almost always present for hydraulic components. This also applies to the minimal model. With the fluid properties of HLP 46 (**Table 1**), the following course of the Reynolds number over the pressure drop is obtained for the geometry, see **Figure 5**. In the diagram on the left, the Reynolds number is plotted against the pressure drop Δp. It includes the marked operating points investigated for this study. In order to gain a better understanding of the influence of temperature on the Reynolds number and thus on turbulence, the characteristic curve for 20 °C and 100 °C is shown in addition to the 40 °C curve. It shows that an undercutting of the limiting Re ≈ 30 (red line) is given even at lower temperatures like 20 °C for pressure drops smaller than 1 bar. A classical hydraulic application has much higher operating points. It shows that a laminar flow is only important in special cases of very low pressure drops and over large narrow sections. However, in such special cases effects like flow forces are negligible compared to static pressure forces. The reason for these cases with very low velocities is the quadratic contribution of the flow force, see Eq. (3).

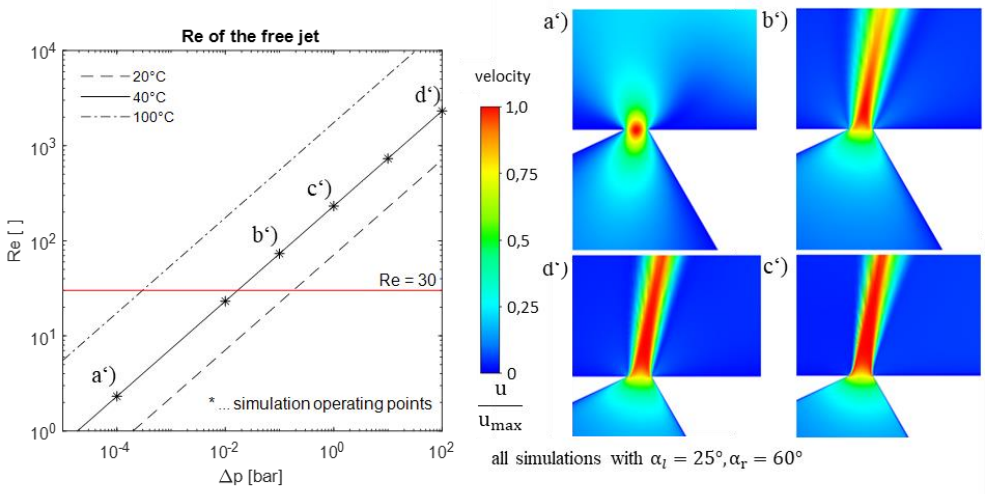


Figure 5: Investigation of the laminar-turbulent influence for the case: "$\alpha_l = 25°$, $\alpha_r = 60°$,, Reynolds number plotted over pressure drop for various temperatures (left), velocity fields of the free jet for Reynolds numbers at a') to d')

On the right side in **Figure 5**, the velocity fields for the operating points a') to d') are shown for the geometrical inlet shape configuration from **Figure 4** ($\alpha_l = 25\,°$, $\alpha_r = 60\,°$). It can be seen how the shape of the free jet is influenced by increasing Reynolds number. For $Re > 30$ the shape of the free jet is fully developed and has a sharp separation between the free jet velocity field and the static environment. For $Re < 30$ the shear layer has almost disappeared and it prevails a laminar flow.

The correlation between the averaged flow angle ε_{avg} against the Reynolds number for various geometrical inlet shape configurations (α_l, α_r) will be discussed in **Figure 6**. The left diagram illustrates the normalized averaged flow angle $\varepsilon_{avg}/\varepsilon_{avg100bar}$ (normalized to turbulent flow) against the Reynolds number. A deviation is recognizable between the turbulent and the laminar flow for all cases. With exception of the case „$\alpha_l = 30\,°$, $\alpha_r = 140\,°$" for $Re > 30$ all averaged flow angles stay almost the same and converge. Only for $Re < 30$ it deviates up to 10 % of the

reference value. The reason for that is the influence of the laminar flow. The special case "$\alpha_l = 30\,°, \alpha_r = 140\,°$" already shows a much higher deviation at larger Reynolds numbers. For better understanding, the right side of the figure shows the configuration for cases a") to d") by increasing Reynolds numbers. Due the very flat inlet the free jet is attached to the wall even if the flows is turbulent. This effect is called Coanda effect, [12] and [15]. For certain inclinations of the free jet to the wall the suction area of the free jet leads to the fact that the free jet is aligned with the wall, [14] and [15]. For inlet shape configurations in this range, larger deviations are to be expected.

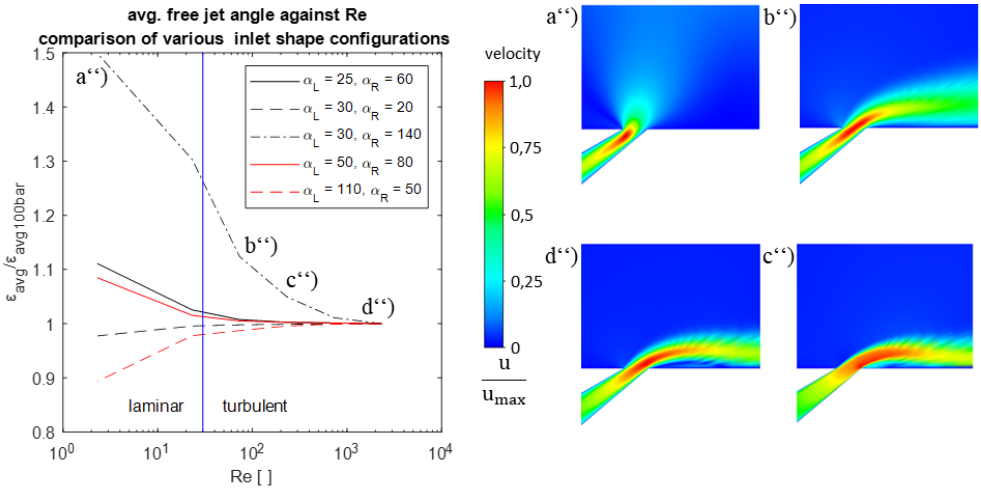


Figure 6: Averaged flow angle ε_{avg} against the Reynolds number for various geometrical inlet shape configurations (α_l, α_r, left), the influence on acute and flat boundary cases (right)

In summery, most important for hydraulic applications is the turbulent free jet. In contrast to the hydraulic resistance description the laminar and turbulent behavior of a flow is always considered in parallel as principle of superposition, as shown in [1]. The current investigation focuses on the free jet angle. The flow resistance of hydraulic components is not taken into account. Thus, it must not be confused with the laminar free jet at this point.

Because of its predominantly turbulent behavior, all further considerations are performed for the turbulent free jet.

3.3. Analytical method for the averaged flow angle

The derivation of the averaged flow angle ε_{avg} for a turbulent free jet is illustrated in **Figure 7**. In a) it can be seen an overview of the geometry near the narrow section including the inlet and outlet momentum. The momentum in the inlet is aligned according to the flow angle. The range of the flow angle is limited through geometrical parameters α_l and α_r. The right side (b) shows the course of the flow angle along the narrow section (see **Figure 4**). Regarding to ch. 3.1 the averaged flow angle lies in the center point of the narrow section. Additionally the course of $\varepsilon(x)$ can be separated in two parts. The first one $\varepsilon(x)^*$ is a point symmetrical graph with its center

point in the coordinate origin. The second part is a constant offset ε_{avg}. In summary the equation for $\varepsilon(x)$ is:

$$\varepsilon(x) = \varepsilon(x)^* + \varepsilon_{avg} \tag{8}$$

It is necessary to set an unproven but plausible assumption to make a reference to the momentum. The absolute value of the momentum over the flow angle in the inlet is uniform. Thus, the momentum can be written as $\vec{I}(\varepsilon) = \vec{e}(\varepsilon)\left|\vec{I}\right|$. In this formula $\vec{e}(\varepsilon)$ is the unit vector with the direction of ε. That means the whole fluid in the inlet moves uniformly towards to the narrow section.

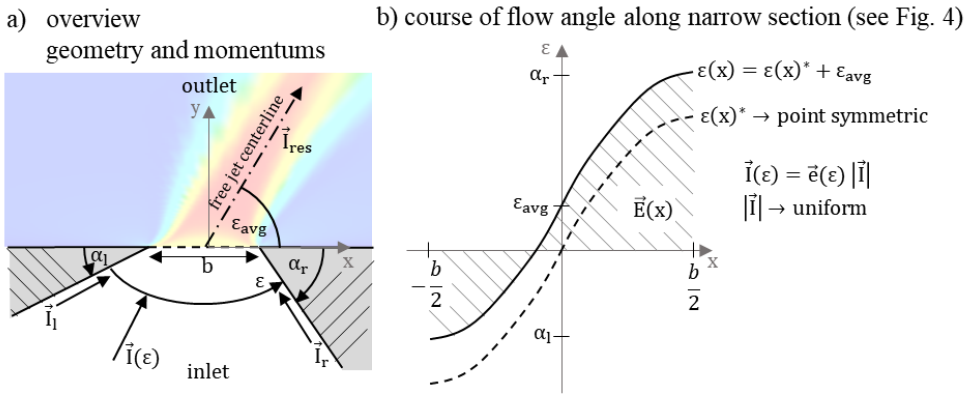


Figure 7: Prediction of the averaged flow angle ε_{avg} – simplified model to consider the incoming and outgoing momentum (a), course of the flow angle along the narrow section (b)

In the following, the relationship between the averaged flow angle and the geometry is derived from the assumptions just mentioned. Regarding to Eq. (6) the first mean value theorem is applied to the whole course of ε according to **Figure 7** b). Integrating over the graph of $\varepsilon(x)$ yields the following expression.

$$\varepsilon_{avg} = \frac{1}{b}\int_{-b/2}^{b/2} \varepsilon(x)\,dx = \frac{1}{b}\left[E(b/2) - E(-b/2)\right] \tag{9}$$

In this case E is the root function of ε. With the possibility to separate $\varepsilon(x)$ such as in Eq. (8), Eq. (9) can be extended and rearranged.

$$\varepsilon_{avg} = \frac{1}{2}\left[\left(\frac{E(b/2)^*}{b/2} + \varepsilon_{avg}\right) - \left(\frac{E(-b/2)^*}{b/2} - \varepsilon_{avg}\right)\right] \tag{10}$$

The expression $(.)^*$ symbolizes the point symmetrical part of the course. The separation remains in the root functions. The first term $\frac{E(\pm b/2)^*}{b/2}$ describes the mean increase of the function for the respective interval $[0, b/2]$ (case 1, left term) and $[-b/2, 0]$ (case 2, right term). The second part is only the offset. Both formula have the same structure like Eq. (8). By means of case separation between the two intervals case 1 and case 2, Eq. (8) can be integrated into Eq. (10).

In case 1 with the interval $[0, b/2]$, a linear equation is set up, which is to be determined at x = b/2. The increase of the function is defined over the whole interval with $\Delta E/\Delta x$. The offset is clarified through ε_{avg}. Case 2 can be described analogously. Eq. (11) shows for both intervals the expression in the form of Eq. (8).

$$\begin{aligned} \tilde{\varepsilon}(b/2) &= \frac{E(b/2)^* - E(0)}{b/2} + \varepsilon_{avg} = k\varepsilon(b/2)^* + \varepsilon_{avg} \text{ for } [0, b/2] \\ -\tilde{\varepsilon}(-b/2) &= \frac{E(0) - E(-b/2)^*}{b/2} - \varepsilon_{avg} = -k\varepsilon(-b/2)^* - \varepsilon_{avg} \text{ for } [-b/2, 0] \end{aligned} \tag{11}$$

The additional factor k is a scaling of the point symmetrical part of $\tilde{\varepsilon}$ for integrating of Eq. (8). With Eq. (11) it is possible to extend Eq. (10) to the following expression.

$$\varepsilon_{avg} = \frac{1}{2}[\tilde{\varepsilon}(b/2) + \tilde{\varepsilon}(-b/2)] \tag{12}$$

Because of the point symmetrical characteristic of $\tilde{\varepsilon}(x)$ as well as $\varepsilon(x)$ the term $\varepsilon(x)^*$ will always eliminate itself. The offset ε_{avg} remains and the Eq. (12) is fulfilled. Thus, Eq. (12) is universal for all point symmetrical equations with a uniform offset.

That means for calculations of the flow angle at the narrow section, only the flow angle at the limiting walls must be known. Regarding to **Figure 4** and the Neumann boundary conditions along a wall the flow angle and the local geometrical angle can be equated. It applies: $\tilde{\varepsilon}(b/2) = \pi - \alpha_r$ and $\tilde{\varepsilon}(-b/2) = \alpha_l$. For a better illustration, Eq. (12) can be extended to a vector sum of the inlet momentum with the reference to the resulting momentum of the free jet. The assumption $\vec{I}(\varepsilon) = \vec{e}(\varepsilon)|\vec{I}|$ with $|\vec{I}|$ as uniform is applied over the inlet area. The **Figure 8** shows the relationship between the momentum of inlet and outlet.

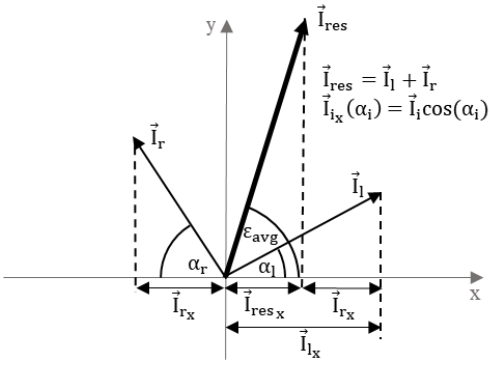


Figure 8: Momentum of inlet ($\vec{I}_l$ and $\vec{I}_r$) in reference to the outlet momentum $\vec{I}_{res}$

For the calculation, the relationship is transferred into Cartesian coordinates. With the expression $\vec{I}_{i_x}(\alpha_i) = \vec{I}_i \cos(\alpha_i)$ all angle dependent momentum can be related to the x component. The **Figure 8** illustrates sum of the x components aligned to the abscissa. According to this principle, Eq. (12) can be transferred into the x component of the Cartesian coordinates, as shown in Eq (13).

$$\cos(\varepsilon_{avg}) = \frac{1}{2}(\cos(\alpha_l) - \cos(\alpha_r)) \tag{13}$$

Equation (13) is the major finding of this work. It describes the correlation between the geometrical angles α_L, α_r and the averaged flow angle ε_{avg}. This equation allows the direct calculation of the flow angle from the geometry of the control edge only. It is independent of any operating conditions (see ch. 3.2). As an example, Eq. (13) gives a flow angle of 78,3 ° when the geometrical data from **Figure 4** are used. This matches very well with the CFD results of 79 °.

Figure 9 shows the comparison of the Eq. (13) and the CFD data over the full range of α_l and α_r. Geometrical impossible configurations, when $\alpha_l + \alpha_r > 180°$, are colored grey in the diagram. It can be seen that for all geometric shapes, the values calculated by Eq. (13) match very well with the CFD results. Only for the limiting cases $\alpha_l \gg \alpha_r$ as well as $\alpha_l \ll \alpha_r$ a slight deviation is observed. The deviations are smaller than 10 % for the special cases and thus acceptable. In another cases $\alpha_l = 0$ °and $\alpha_r = 0$ °, the free jet angle can only vary between 0 ° and 90 ° or from 90 ° to 180 °. If both inlet angles are identical, the free jet angle is always constant 90 °, as shown in the velocity field of configuration b'). This special case enables a symmetrical flow through the inlet geometry and generates a "classic" vertical free jet.

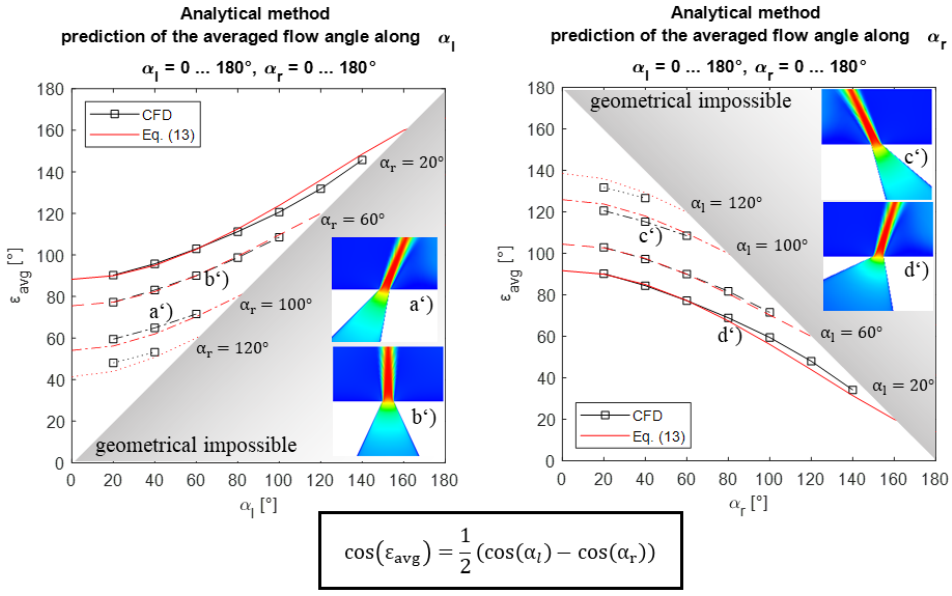


Figure 9: Prediction of the averaged flow angle ε_{avg} – comparison between Eq. (13) and CFD data, plotted over $\alpha_l = \{0 \ldots 180\}$ ° (left) and $\alpha_r = \{0 \ldots 180\}$ ° (right)

In summary, the Eq. (13) shows very good results compared to the minimal model. The simple expression represents a connection to geometrical parameters only and is not related to any operation conditions. The assumptions for this derivation is valid for turbulent flow. Thus, it is useable for almost all applications of hydraulic systems.

4. TRANSFER TO REAL APPLICATION

4.1. Prediction of the flow angle in real applications

In the following, the Eq. (13) is evaluated by using the rotary slide valve. **Figure 10** shows the comparison between CFD data with the results of Eq. (13). In a) the geometrical characteristics of the narrow section is defined. The influencing angle in the inlet of the valve is α_{in}. The angle α_{in} is valid in the range between 0 ° (vertical inlet) to 75 °. The relationship between α_{in} and the angles α_l and α_r of Eq. (13) is illustrated in figure a).

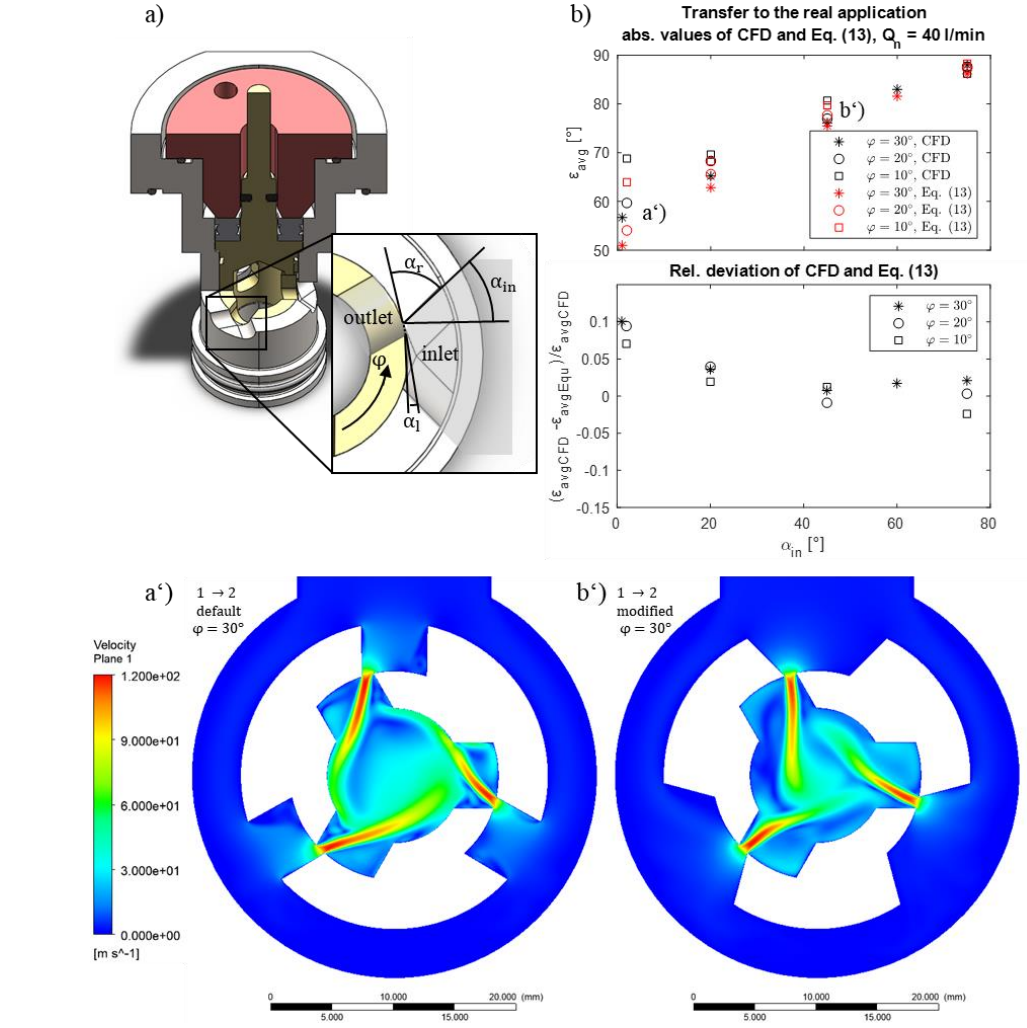


Figure 10: Rotary slide valve, geometrical setup (a), comparison between CFD data and the Eq. (13) along $\alpha_{in} = \{0 \dots 75\}$ ° for different twisting angles $\varphi = \{10,\ 20,\ 30\}$ ° (b), velocity field in symmetrical plane through the narrow sections for different configuration of $\alpha_{in} = \{0\ °,\ 45\ °\}$, $Q_n = 40$ l/min, flow direction 1 → 2 (see **Figure 1**)

The setup of all simulations is similar to **Table 1** and is expected to the parameter variation with α_{in}. In b) the absolute values of the averaged flow angle as well as the relative deviation between

CFD data and Eq. (13) are plotted over α_{in} for different twisting angles φ. The angle α_r corresponds to the inverse behavior of the angle α_{in} ($\alpha_r = 90\,° - \alpha_{in}$). Additionally, α_r and α_l have a variation from 0 ° (almost closed) to - 15 ° (almost fully open) dependent on the twisting angle. The reason for that is the convex shape of the outer surface of the valve spool. In order of the Eq. (13) to be applied, the geometric conditions must be established from **Figure 7**. Regarding to **Figure 9** right side, the graph of the absolute values shows a very similar course. For $\alpha_{in} > 20\,°$ the results between CFD and Eq. (13) match very well and the relative deviation is below 5 %. It proves that the explained calculation method is reliable. In the range of $\alpha_{in} < 20\,°$ the scattering of the values becomes larger between different twisting angles and the relative deviation is the highest in this course with about + 10 %. However, for an engineering application the deviations are still acceptable. The results in general are in line with the statement from **Figure 9**.

For a better understanding how the fluid domain is modified, in a') and b') the velocity field in a symmetrical plane through the narrow sections is shown for different configurations of $\alpha_{in} = \{0\,°,\ 45\,°\}$. The plot a') shows the default configuration with a vertical inlet. The free jets flow much more flatly into the geometry and have large contact regions with the wall. The second one, b'), clarifies the modified inlet shape. The free jets meet in the center and partially dissipate themselves. Thus, the resistance torque can be significantly reduced (see ch. 4.2). It is recognizable how α_{in} influence the flow angle of the free jet.

To sum up, the new method of Eq. (13) allows a reliable prediction of the flow angle of free jets. Same results can be achieved for complex geometries as the rotary slide valve. It is possible to influence actively the jet direction after the narrow section. Looking at other issues such as cavitation collapse regions (erosion regions) or stagnation points in general, the potential for improvement using Eq. (13) is high.

4.2. Improvement of the rotary slide valve

The main goal of the improvement of the valve is the reduction of the resistance torque which is strongly dependent on the flow force. As shown in Eq. (3) the flow force is mainly influenced by the flow angle at the narrow section. With the results of ch. 4.1, the following improvements could be made, as shown in **Figure 11**. The characteristic curves for pressure drop (a) and the resistance torque (b) over φ are plotted for the default (black) and the modified (red) valve design and both flow directions (see **Figure 1**). The direction $1 \rightarrow 2$ is marked as solid line and $2 \rightarrow 1$ as dashed line. In addition, both characteristic curves are simulated by a constant nominal flow rate Q_n. Comparing to pressure drop across φ (a), the results for both designs are similar. Moreover, the gradient is almost identical. There is only an angular offset of 1 ° to 2 °. Significant changes can be seen in the resistance torque. Especially for the flow of $1 \rightarrow 2$, the amount in the maximum range decreases by 65 % from 0.23 Nm to 0.08 Nm. The resistance torque of $2 \rightarrow 1$, on the other hand, is negatively affected. However, this effect is significantly smaller and can be accepted at this point. In order to actively influence the flow direction $2 \rightarrow 1$, the inlet angle of the inner geometry of the rotary slide valve would have to be varied according to this principle. Unfortunately, cost-related changes are not acceptable for cost optimized part usage in industrial applications.

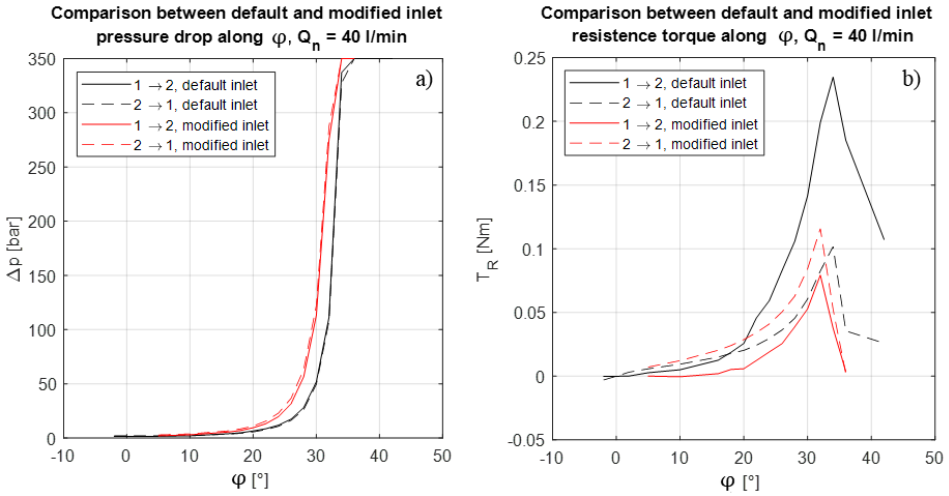


Figure 11: Improvement of the modification, comparison of the pressure drop across φ at constant flow rate (a) and of the resistance torque across φ (c), in each case for both flow directions, $Q_n = 40\ l/min$

In summary, the flow force improvement by using flow angle modification according to Eq. (13) is very effective. The main setup of the geometry is unchanged and only the inlet shape near the narrow section has to be modified. The impacts of this modification are enormous and the production effort is kept within limits.

5. DISCUSSION AND CONCLUSION

In summary, the following conclusions could be drawn with the help of the presented research.

Turbulent-laminar operating conditions

For turbulent flows, the averaged flow angle is independent of the operating point. In the range of laminar flows there is a clearly deviation compared to turbulent flow angles. Due to the low Reynolds number of $Re \approx 30$, this effect is outside the operating conditions of "classical" hydraulic applications. Therefore, the operating range is always in the turbulent range and a behavior independent of the operating point is to be expected.

Analytical method for the determination of the flow angle

The new method according to Eq. (13), the averaged flow angle can be predicted with very good approximation. Slight deviations are only to be expected in ranges of $\alpha_l \gg \alpha_r$ as well as $\alpha_l \ll \alpha_r$, which can amount to a maximum deviation of 10 %. The simple expression represents only a dependency to geometrical parameters and is not related to any operation conditions. The assumptions for this derivation is valid for turbulent flow. Thus, it is useable for almost all applications of hydraulic systems.

Transferability to real applications

The new method of Eq. (13) allows a reliable prediction of the flow angle of free jets which is also applicable for a wide range of hydraulic geometries. It is possible to influence actively the

jet direction after the narrow section. Looking at other issues such as cavitation collapse regions (erosion regions) or stagnation points in general, the potential for improvement using Eq. (13) is high. The main setup of the geometry is unchanged and only the inlet shape near the narrow section has to be modified. The impacts of this modification are enormous and the production effort is kept within limits.

NOMENCLATURE

A	Area	m^2
b	width	m
f	acceleration	m / s^2
F	force	N
I	momentum	N s
l	characteristic length	m
n	normal vector	-
p	pressure	bar
r	radius	m
S	boundary of control volume	m^2
t	time	s
T	temperature	°C
T_R	resistance torque	Nm
u	velocity as vector or as component in x-direction	m / s
v	velocity as component in y-direction	m / s
V	volume	m^3
x	coordinate direction	m
y	coordinate direction	m
α	geometrical angle	°
Δ	difference	-
ε	flow angle	°
E	Root function of ε	
η	dynamic viscosity	Pa s
$\boldsymbol{\rho}$	density	kg / m^3
$\boldsymbol{\tau}$	shear stress tensor	Pa
υ	kinematic viscosity	mm^2 / s
$\boldsymbol{\varphi}$	twisting angle	°
$\boldsymbol{\Omega}$	control volume	m^3

ACKNOWLEDGMENTS

The author would like to thank the institutions Bundesministerium für Bildung und Forschung (BMBF) and Europäische Union - NextGenertionEU for funding the project " Erarbeitung funktionsintegrierter Schaltventile für Maschinen, Anlagen und Fahrzeuge zur Senkung des Energieverbrauchs (funkiS)“with the funding code 01LY2004B.

REFERENCES

[1] N. Gebhardt, J. Weber (2020) Hydraulik – Fluid-Mechatronik. Grundlagen, Komponenten, Systeme, Messtechnik und virtuelles Engineering. Dresden, January 2020, Dresden, Germany

[2] S. Osterland, L. Günther, J. Weber (2022) Experiments and Computational Fluid Dynamics on Vapor and Gas Cavitation for Oil Hydraulics, TU Dresden, 2022 Dresden, Germany
[3] S. Osterland, L. Müller, J. Weber (2021) Influence of Air Dissolved on Hydraulic Oil on Cavitation Erosion, TU Dresden, 2021 Dresden, Germany
[4] R. Ivantysyn, A. El Shorbagy, J. Weber (2018) Schlussbericht – Smart Pump – decentralized control for vessel engine. Dresden, 2018, Dresden, Germany
[5] M. Dietze, et. al. (1996) Messungen und Berechnungen der Innenströmung in hydraulischen Sitzventilen. Technische Hochschule Darmstadt, 1996, Darmstadt, Germany
[6] M. Kipping, et. al. (1997) Experimentelle Untersuchungen und numerische Berechnungen zur Innen strömung in Schieberventilen in der Ölhydraulik. Technische Hochschule Darmstadt, 1997, Darmstadt, Germany
[7] C. Latour, et. al. (1996) Strömungskraftkompensation in hydraulischen Sitzventilen. RWTH Aachen, December 1996, Aachen, Germany
[8] M. Ristic, et. al. (2000) Dreidimensionale Strömungsberechnungen zur Optimierung von Hydraulikventilen bezüglich der stationären Strömungskräfte. RWTH Aachen, 2000, Aachen, Germany
[9] K. Wanne, et. al. (1965) Messung und Untersuchung der axialen Kräfte an ölhydraulische Steuerschiebern. Technische Hochschule Stuttgart, 1965, Stuttgart, Germany
[10] M. Lechtschewski, et. al. (1994) Untersuchung der Abhängigkeit der Strömungskraft vom Hub des Ventilschiebers und der Druckdifferenz. Institut für Werkzeugmaschinen und Fluidtechnik, TU Dresden, 1994, Dresden, Germany
[11] S. Osterland, J. Weber (2016) A numerical study of high pressure flow through a hydrsulic pressure relief valve considering pressure and temperature dependent viscosity, bulk modulus and density. TU Dresden, 2016 Florianópolis, Brazil
[12] H. Schlichting, K. Gersten (1996) Grenzschicht-Theorie. Springer-Verlag, 1996, Bochum, Germany, ISBN: 3-540-55744-x 9th edition
[13] E. N. Andrade, et. al. (1931) The velocity distribution in a liquid-into-liquid jet. The plane jet. Proc. Phys. Soc. London, 1931, London, UK
[14] J. Liu, A. Sitte, J. Weber, (2022) Invesitgation of temperature on flow mapping of electrohydraulic valves and corresponding applications, ASME/BATH 2022., UK
[15] E. Truckenbrodt (1980) Fluidmechanik – Band 2 Elementare Strömungsvorgänge, dichteveränderliche Fluide sowie Potential und Grenzschichtrömungen. Springer-Verlag Berlin Heidelber New York. München, Germany, ISBN 3-540-10135-7 2nd edition

SIMULATION OF GAS LEAKAGE ON BALL SEAT VALVES

peer reviewed

Felix Fischer[1]*, Katharina Schmitz[1]

[1]*Institute for Fluid Power Drives and Systems, Campus-Boulevard 30, 52074 Aachen*

* Corresponding author: Tel.: +49 241 80-47725; E-mail address: f.fischer@ifas.rwth-aachen.de

ABSTRACT

Gas leakage is a critical issue in various industrial applications that utilize ball seat valves. This is especially relevant for hydrogen applications, due to its high reactivity. This paper presents the application of a model for analyzing liquid leakage on ball seat valves applied to gases and the experimental validation of these results. The research objective is to enhance the understanding of the leakage mechanisms and provide valuable insights for improving the design and performance of ball seat valves. The simulation model considers valve geometry, surface roughness, and material properties to predict the leakage behavior accurately. The simulation method is based on the contact mechanics model developed by Persson [1]. It considers the surface roughness via the two-dimensional spectral density. The model is validated experimentally by comparing the simulated leakage rates with experimentally measured values for seats with different surface roughness and anisotropy. This way, the validation shows whether the leakage model can be applied to gases.

Keywords: Simulation, Leakage, Tribology, Gas, Valves, Pneumatics

1. MOTIVATION

Metallic sealing in gas applications is a highly specialized and crucial technology known for its ability to create gas-tight seals in demanding environments. This method relies on metal-to-metal interfaces to achieve tightness, effectively preventing the escape or ingress of gases. These applications encompass various industries, including aerospace, automotive, petrochemical, and power generation, playing a pivotal role in ensuring safety and operational reliability.

One notable advantage of metallic sealing is its capability to achieve hermetic tightness, which is essential in applications where gas leakage is unacceptable, such as in spacecraft, vacuum chambers, and certain industrial processes. Furthermore, these seals demonstrate exceptional resistance to corrosion, temperature and chemical exposure, making them suitable for environments where gases may be corrosive or chemically reactive. In addition to their longevity and reliability, metallic seals have a longer service life compared to many other sealing methods, reducing maintenance and replacement costs.

In the hydrogen industry, the advantages of metallic sealing are particularly noteworthy. Given the unique properties and challenges associated with hydrogen gas, metallic seals offer a reliable solution. Hydrogen is known for its small molecular size, which allows it to permeate through materials that might be impermeable to other gases. When soft polymer seals are exposed to hydrogen, pressure changes can lead to fractures of the seals due to hydrogen contained in the material. Metallic seals, with their gas-tightness, are crucial in preventing hydrogen leakage, especially in applications involving hydrogen storage, transportation, and fuel cells [2].

Another notable aspect is the absence of PFAS (per- and poly-fluoroalkyl substances) such as PTFE (polytetrafluoroethylene) in metallic sealing. PTFE, a commonly used material in traditional gasket

seals, has come under critique recently due to its long-term chemical stability. Due to the spread of microscopic PFAS particles, residues of PFAS can be found in human bodies. Especially, a high concentration of PFOA (perfluorooctanoic acid), a PFAS used in the production of PTFE, is known to accumulate permanently in human bodies. The effects of PFAS on the environment and human health have yet to be examined [3,4].

Simulations are essential for assessing the leak tightness of metallic seals. They provide a cost-effective way to evaluate seal performance early in the design process, predict how seals will perform under various conditions, and identify potential weaknesses or areas susceptible to leaks. Simulations also help to optimize seal designs, reduce costs, and accelerate development.

The study examines the gas leakage in metallic seals, focusing on ball seat valves, see **Figure 1**, as an example. These valves serve as commonly used check valves and provide an adequate model for metallic seals due to the easy ball and seat exchange in an experimental setting. This allows for the study of the varying effects of materials and surfaces under consistent conditions. The findings and approaches outlined in this study can be applied to other configurations of metal seals.

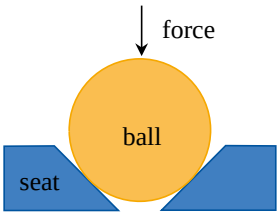


Figure 1: Sketch of a ball seat valve

The paper is structured as follows: In section 2 the state of the art of leakage simulations for metallic seals as well as for gas leakage is presented. A short introduction to the contact mechanics theory developed by Persson, on which the simulation model used in this paper is based, is given in section 3. The simulation method as well as the experimental set-up for the validation are introduced in section 4, while the results of both the simulation and experiment are discussed in section 5. This paper concludes with a summary and outlook.

2. STATE OF THE ART

Leakage simulations of metallic seals have experienced significant progress recently, covering various aspects. One noteworthy development includes a study that created an experimental setup to investigate fluid leakage in metallic seals, which specifically concentrated on water leakage between a steel sphere and a steel body with a conical surface. The experimental findings align well with a corresponding theoretical model, which predicts that plastic deformations may reduce the leak rate by a factor of about 8 [5,6].

In another study, Zhang et al. analyzed fluid leak rates across metal sealing surfaces by creating fractal models for both the contact and leakage processes. They refined the model to describe the seal-contact interface of two metal rough surfaces. In this study, the authors used fractal theory to mathematically model the leakage channel under the contact surface. They also analyzed how factors such as fractal structure, surface material, and gasket size affect the contact and leakage processes. To simulate sealed ring gaskets, numerical simulations were used [7].

Additionally, in another study, Tikhomorov, Gorlenko, and Izmerov used a mathematical modeling approach to simulate leakage through mechanical seals, accounting for waviness and roughness by treating the microscopic net of flow channels through the contact area as a porous medium [8].

These advancements contribute to better comprehension and improved predictive abilities for administering leakage in metallic seals.

The leak tightness of seals against gaseous media has not yet been studied widely. Most of the knowledge gained from research on metallic seals in general is also applicable to gases, but there are still some relevant differences.

A study calculated the leakage in gas labyrinth seals compared to the leakage for liquids. The research revealed that bulk-flow models (BFMs) for gas labyrinth seals often inaccurately predict mass flow. Labyrinth seal mass flow prediction was analyzed using computational fluid dynamics (CFD) to evaluate the effects of tip clearance and operating conditions, and an updated kinetic energy transfer coefficient was derived from the data to improve the accuracy of the Neumann leakage equation.

In another paper, Huon et al. analyzed the gas leakage for polymer seals. The leak rate depends on the interfacial surface roughness and viscoelastic properties of the rubber. The authors present a theory for gas flow that considers both diffusive and ballistic flow.

In the same paper, these models have been compared to experiments. To create random surface roughness, the rubber O-rings undergo sandblasting. Notably, while gas leakage occurs, there is no such issue observed for water-filled barrels. This is attributed to capillary effects resulting from Laplace pressure or surface energy [9].

3. THEORETIC BACKGROUND

The simulation method presented in this paper is based on Persson et al.'s percolation method and contact mechanics.

This contact mechanics model is rooted in a theoretical framework that addresses the intricate dynamics of rough surfaces in contact. This model employs principles from solid mechanics and surface science, providing a comprehensive understanding of contact mechanics. At its essence, the model recognizes the intrinsic roughness of real physical surfaces by characterizing these irregularities through statistical descriptors like power spectral density (PSD) functions, usually denoted as $C(\boldsymbol{q})$, with $\boldsymbol{q}$ being the two-dimensional surface roughness wave vector [10].

Elastic deformation of the contacting materials due to an applied load is a crucial factor. To analyze this deformation, contact area, and pressure distribution, Hertzian elasticity theory is employed. The model utilizes integral equations derived from Green's function formalism to connect contact pressure, deformation, and adhesion while accounting for the nonlinearity induced by surface roughness. With this formalism, a modified contact pressure relation $p_C(u)$ can be derived, which connects the local contact pressure p_C with the average separation of the surfaces u [11].

Material properties, such as elasticity and Poisson's ratio, are crucial in characterizing how materials respond to external loads and deformations. Moreover, statistical mechanics principles are utilized to describe the distribution of contact points and the probability of adhesive interactions, which enhances the model's effectiveness in describing the macroscopic mechanical behavior of rough surfaces [10].

The leakage theory developed by Persson et al. is based on the concept of percolation theory. It treats the evolution of the pressure distribution with increasing resolution of the surface roughness as a diffusive process in which the resolution replaces the function of time. This approach allows a detailed analysis of how surface roughness affects the effective contact area and the leakage rate.

In their work, they explain the leakage mechanism of the sealing interface based on their contact mechanics theory and two-dimensional percolation theory. This theory predicts that when the relative

area ratio of the apparent and the real sealing surface Π is about 0.42, a leakage channel will be formed at the sealing interface, and the fluid will flow from the high-pressure side to the low-pressure side. In this theory, the real contact area (and accordingly Π) is not only a function of the contact pressure p_C respectively u, but also of the magnification ζ. The magnification denotes a cut-off of the PSD and it indicates up to which wave vectors q are considered in the evaluation. Hence, ζ can be treated as the resolution of the contact. The ratio Π decreases as ζ increases. This allows for the definition of a critical magnification ζ for every contact pressure, at which the contact area percolates. The theory then predicts the average separation of surfaces at a magnification slightly above this critical magnification. The leakage can then be calculated using the Bernoulli equation for a thin gap [12]. In general, the Bernoulli equation is valid for both compressible and incompressible fluids.

Alternative to the critical junction approach, which has been used in this paper, Persson et al. also propose the effective medium approach, which calculates the effective conductivity of the mesh of channels that percolate the contact area. This method reaches very similar results to the critical junction method [13]. The predictions by Persson et al.'s contact theory have been validated experimentally for liquid leakage at polymeric and metallic seals multiple times [5,14].

4. METHODOLOGY

This section introduces the methodology of the study presented in this work. It is separated into two sections. The first section introduces the simulation used for calculating the leak rate. In the second section, the test rig used for experimental validation of the simulation is presented.

In this work, a total of six different seats with varying surface properties have been analyzed simulationally and experimentally. The leakage of these valves has been evaluated for relative air pressures up to 7 bar. These are the conditions usually found in pneumatic systems.

4.1. Simulation

The simulation implemented for this work is based on the contact mechanics model by Persson et al. presented in section 3. The leakage simulation developed in this paper can be divided into two sub-steps: a microscopic and a macroscopic one.

The microscopic simulation examines the deformation of the roughness peaks and the resulting network of microscopic channels. The length and shape of these channels result from the topography of the surfaces and the contact pressure distribution determined in the macroscopic simulation using Persson contact theory. The exact surface structure is incorporated statistically in the form of the 2D power spectral density.

In the macroscopic simulation, the contact pressure distribution and thus the contact width are determined. This calculation is based on the applied pressures and forces, the geometry of the valve, and the material properties of the seat and sealing body. The simulation method is based on analytical equations for the contact pressure distribution.

To calculate the leakage, the flow through the contact area has to be determined using a flow model. For this purpose, the critical resolution method is used in this paper. Once the model has been successfully developed, the results can be validated with the measurements from the next subsection.

Surface analysis

The first step of the simulation is the microscopic analysis. In this step, the provided surfaces are analyzed and the contact pressure relation $p_C(u)$ is calculated based on these measurements. For this reason, the two-dimensional PSDs are calculated for every surface. The surface data has been

measured by optical microscopy with a vertical and lateral resolution of up to 270 nm. An estimate for the true 2D-PSD can be calculated by a two-dimensional version of Welch's method [15].

A depiction of the two-dimensional PSD of the first of the six seats can be seen in **Figure 2**. On the left, a two-dimensional representation of the upper right quadrant (all components of $\boldsymbol{q}$ are positive) of the PSD can be seen. On the right-hand side, the comparison of the PSD in the direction of the fluid leakage, the radial direction, and the azimuthal (tangential) direction, which is orthogonal to the radial direction (direction of fluid flow), is shown.

Based on the two-dimensional PSD, it is possible to calculate the radially averaged PSD $C(q)$, as well as the Peklenik number γ, which represents the degree of the anisotropy. In general, γ will be a function of ζ [16]. The exact value of γ is dependent on the orientation of the surface. In this work, the local orthonormal coordinate system is chosen, such that the x-axis aligns with the radial direction of the seat. Thus, a value of γ below indicates, that the surface asperities are mainly aligned in the azimuthal direction. This is very typical for technical surfaces of seats because they are usually manufactured by turning and the primary contribution to the roughness are the grooves left over by this manufacturing process. The remains of these grooves can usually even be seen after post-processing

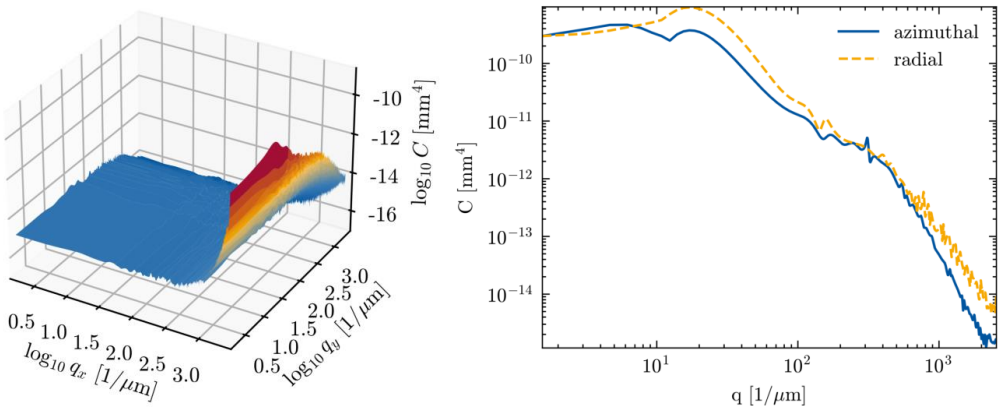


Figure 2: Two-dimensional PSD of seat 1

The two most important variables to describe the roughness of any given surface are the two-dimensional root mean square roughness S_Q and the Peklenik number γ. Both can be calculated either directly from the surface topography in spatial coordinates or from the two-dimensional PSD in the wave vector space. The consistency of the PSD calculation can be verified by comparing the results of both calculations. In further evaluations of the surfaces, only $C(q)$ and $\gamma(\zeta)$ are needed.

The radially averaged PSD of seat 1, as well as the function for γ can be seen in **Figure 3**.

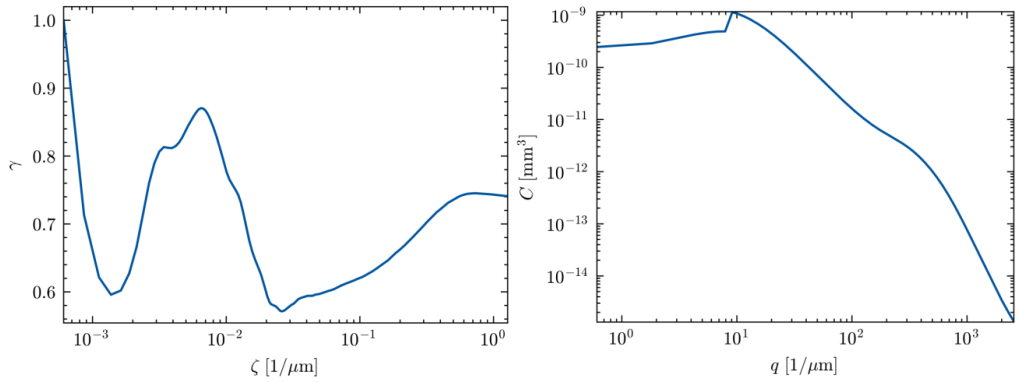


Figure 3: Radially averaged PSD and the magnification-dependent Peklenik number of seat 1

The values of S_Q and γ (= $\gamma(q_{max})$) of the surfaces analyzed in this work are listed in **Table 1**.

Table 1: Properties of the surfaces

Measured surface	S_Q [µm]	γ	Measured surface	S_Q [µm]	γ
Seat 1	2.5	0.74	Seat 4	1.6	0.50
Seat 2	2.7	0.94	Seat 5	1.0	0.52
Seat 3	5.2	0.40	Seat 6	1.9	0.39

Using these PSDs, it is possible to calculate the contact pressure distribution p_C and the fraction of the real contact area Π as a function of ζ for all contact pressures for each surface. In **Figure 4**, graphs of these quantities for seat 1 can be seen. The ratio Π is depicted for a contact pressure of 55 MPa.

Based on these prerequisites, it is possible to calculate the pressure flow factors ϕ_P and shear flow factors ϕ_S as defined by Patir and Cheng [17]. The flow factors indicate the influence of the surface roughness on the fluid dynamics as described by the average Reynolds equation. The shear flow factors indicate the amount of fluid transported in case of a relative motion of the rough surfaces, while the pressure flow factors indicate the change of the flow through the system compared to the expected flow through a smooth system with the same geometry. Patir and Cheng present in their work an approximative formula for both flow factors ϕ based only on S_Q and γ. The result flow factors differ from the results predicted by the Persson theory, which takes into account γ and $C(q)$ [18].

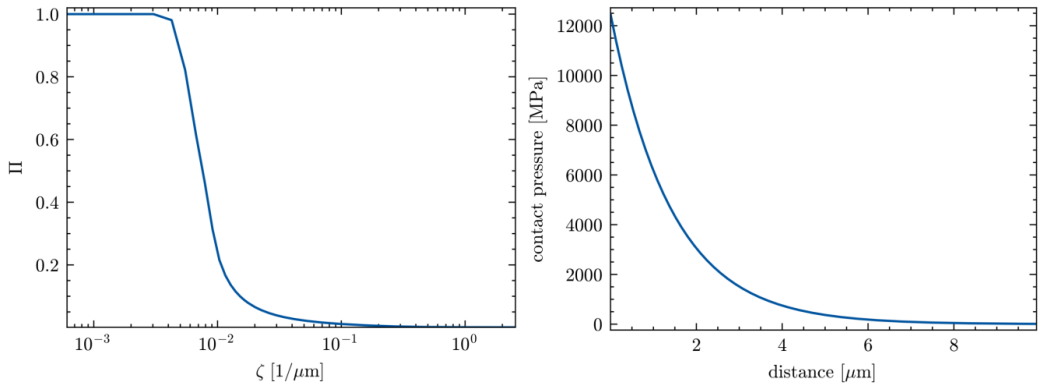


Figure 4: Ratio of real contact area vs. apparent contact area and the contact pressure relation of seat 1

A comparison between the shear flow factors and the pressure flow factors according to the Persson

theory and the approximation by Patir and Cheng can be seen in Figure 5.

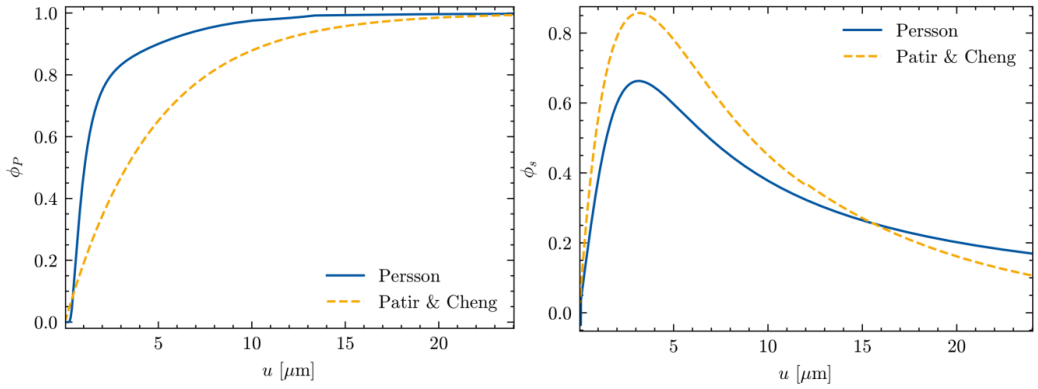


Figure 5: Pressure and shear flow factors of seat 1

In case of leakage at static ball seat valves, the shear flow factors are not needed for further calculations. However, the pressure flow factors are directly proportional to the leakage of the valve. Due to the dependence of ϕ_P on the contact pressure, which can be mapped to the distance u bijectively, as well as the dependence of the leakage on the width of the contact area, the contact pressure distribution must still be calculated. This calculation is presented in the following section.

Contact pressure distribution

The contact pressure distribution of the ball and the conical seat must be calculated based on the geometry and the material properties of the bodies at contact. These macroscopical contact properties are, however, still dependent on the microscopic contact pressure relation. The elasto-plastic contact between the ball and the seat can be separated into two different contributions. On the one hand, there is the elastic deformation of the macroscopic smooth bodies. On the other hand, there is the elasto-plastic deformation of the surface asperities. For low contact pressures, it can be assumed, that the almost whole elastic energy can be encompassed by the elastic deformation of the surface.

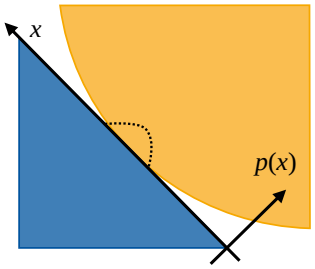


Figure 6: Coordinates used for expressing the contact pressure distribution on the seat (shown in blue)

Using this approximation, and fitting the contact pressure relation seen in **Figure 4** by an exponential function of the form $p_C(u) = p_{\max} \exp(-u/u_0)$, it is possible to calculate the contact pressure distribution analytically. The contact pressure distribution along the flank of the ball seat valve can be calculated by the following Gaussian bell curve [5]:

$$p(x) = \frac{1}{2\pi R \cos\theta \sin\theta} \frac{F}{\sqrt{2\,\pi\, R u_0}} \exp\left(-\frac{(x - x_0)^2}{2 R u_0}\right) \tag{1}$$

In this equation, F is the total normal force acting on top of the ball, R is the radius of the ball, θ is

the angle of the slope of the seat, x is the position along the slope (see **Figure 6**), and x_0 is the position of the maximum of the curve.

The geometry of the test object used in this work, as well as the material parameters, are listed in **Table 2**. The geometry and the material are equal for all seats examined in this work.

Table 2: Parameters of the test objects

Parameter of ball	Value	Parameter of seat	Value
Radius R	2 cm	Inner radius	7.5 mm
Young's modulus	190 GPa	Young's modulus	200 GPa
Poisson's ratio	0.28	Poisson's ratio	0.28
Yield strength	650 MPa	Yield strength	550 MPa
		Slope of the flank	$\pi/4$ (= 45°)

The calculated contact pressure distribution for seat 1 can be seen in **Figure 7**. Using this contact pressure relation, the maximal contact pressure can be read directly from the equation. However, the width of the contact area cannot be defined trivially. The Gaussian curve in this model does never reach zero and the areas of low contact pressure are not relevant for the seal. Thus, in this work, the contact area is arbitrarily chosen as the double standard deviation of the bell curve. This choice allows the definition of a consistent method of calculating the width based on the given parameters. Other possible choices would be, for example, the half-width height of the curve. In either case, there is a linear relation between these parameters. The width itself enters linearly into the leakage.

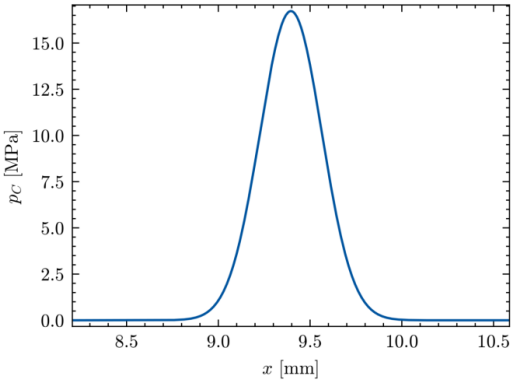


Figure 7: Pressure distribution of seat 1 at a relative gas pressure of p = 5.1 bar

The mass flow rate $\dot{m}$ can be calculated by using the following equation, based on the Hagen-Poiseuille equation for under-critical gases with constant viscosity, which is a valid approximation for air at small pressures:

$$\dot{m} = p\,\rho(p)\frac{u^3}{12\,\eta}\phi_P\frac{R\,sin\,\theta\,cos\,\theta}{a_C}$$

In this equation, p is the relative gas pressure, and a_C is the calculated width of the contact. The density ρ is, according to the ideal gas law, a function of both the pressure and the temperature. In this case, due to the small leak rate, the process is assumed to be isothermal at the environmental temperature.

With these steps, it is possible to calculate the leakage of the valves. The results of this simulation will be validated experimentally; the next section introduces the experimental setup used in this work.

4.2. Experiment

The simulation presented in this paper is validated experimentally using a test rig designed at ifas for this purpose. This test rig allows for the calculation of the leak rate by measuring the decrease of the relative pressure inside a closed cavity connected to the closed valve over time. A circuit plan of the test rig used in this work can be seen in **Figure 8**.

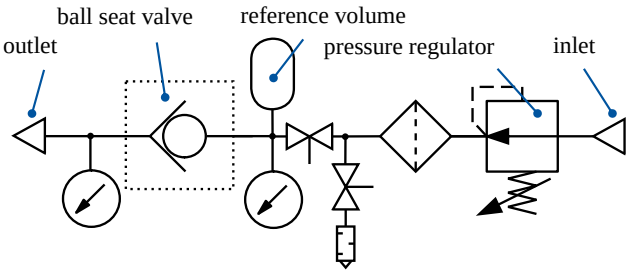


Figure 8: Circuit plan of the test rig

To measure the pressure drop-off due to the leakage through the valve correctly, it is important to ensure that the other connections to the reference volume are leak-tight compared to the seat valve. In this work, this has been assured by measuring the leak rate on a perfectly tight reference valve. The pressure drop-off has been measured five times for every seat with a new mounting process between every single measurement. The pressure curve of a single measurement on seat 5 can be seen on the left-hand side of **Figure 9**. The dotted red lines indicate the cut-offs of the evaluated time. The right-hand side shows the relation between the relative gas pressure and the mass flow rate. The relation between the mass flow rate and relative pressure can be calculated by the following formula, which can be derived from the ideal gas law:

$$\dot{m} = \frac{V}{R_s T} \frac{\mathrm{d}p}{\mathrm{d}t} \tag{2}$$

In this equation, V represents the closed reference volume, R_S is the specific gas constant of air, and T the (constant) temperature of the system.

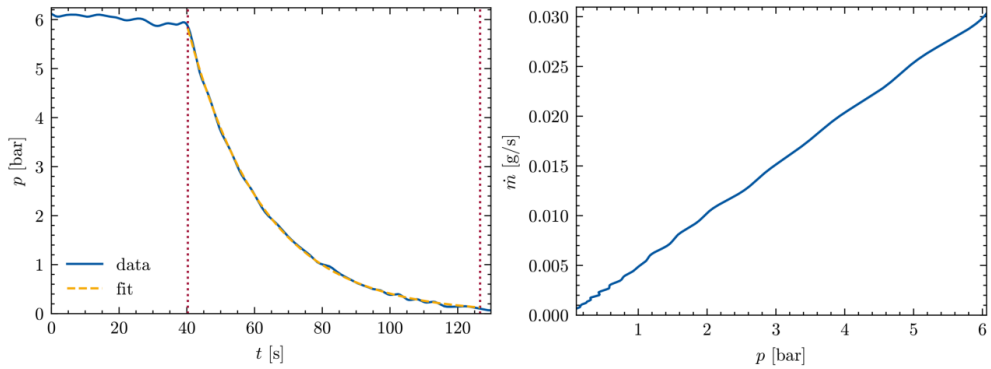


Figure 9: Pressure drop-off and leak rate of the first seat

The derivative of the pressure can be found by applying an exponential fit to the graph of *p(t)*, as seen on the left-hand side, and taking the analytical derivative of this fit. The bijective form of this relation allows for the calculation of the leak rate for every relative pressure for each measurement. Then, these results are averaged for a chosen set of relative pressures over the whole set of measurements,

and the mean, as well as the uncertainty of the leak rate, can be found.

5. RESULTS AND DISCUSSION

A comparison of the measured and the simulated leak rates can be seen in **Figure 10**. It shows the leak rates for seat 6 (left) and seat 3 (right), which have similar anisotropy, but very different roughness and different leak rates, see **Table 1**.

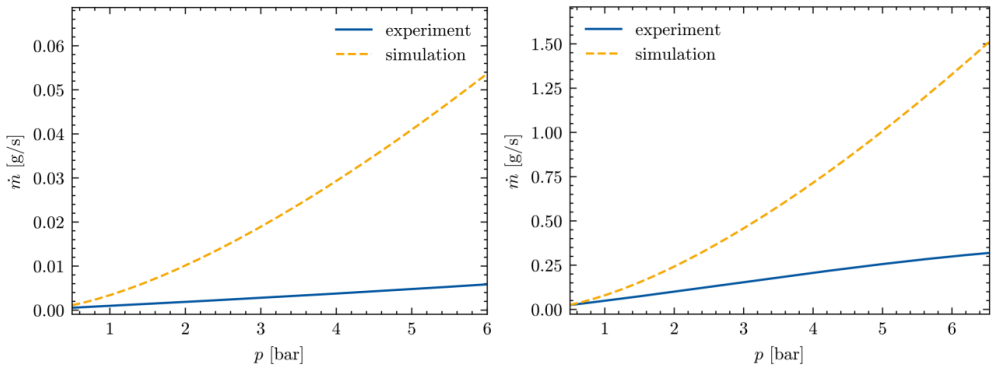


Figure 10: Comparison of the simulated and the experimental leak rates for seats 6 and 3

In both cases, the simulation overestimates the leak rate for high relative pressures and accordingly high contact pressures. In the case of seat 4 on the left-hand side, the leakage is underestimated for low gas pressures.

The simulated and measured leak rates at p = 4 bar are listed in **Table 3.**

Table 3: Leak rates at p = 4 bar

Measured surface	Simulation [mg/s]	Experiment [mg/s]	Measured surface	Simulation [mg/s]	Experiment [mg/s]
Seat 1	2.5	20.2	Seat 4	13.0	9.9
Seat 2	2.7	23.8	Seat 5	1.0	5.0
Seat 3	702.0	206.8	Seat 6	31.8	3.8

The reason for this overestimation is the flow law implemented in this simulation. The simulation tool has originally been developed for liquid leakage, rather than gas leakage. Due to the circumstances of the fluid, it could be possible, that the flow conditions are turbulent, i.e. the Reynolds number is too high, and in this case, both the flow factor method and the Hagen-Poiseuille equation are not applicable. If the diameter of the microscopic channels is small compared to the mean free path of the gas molecules, i.e. the Knudsen number is too large, traditional fluid mechanics is no longer applicable [19]. Therefore, further adjustments are needed to apply this model to gas leakage.

6. SUMMARY AND OUTLOOK

In this paper, a simulation method based on Persson's theory to simulate the leakage of liquids through metallic seals has been presented and applied to gases. The simulation results have been compared to experiments. It was shown, that the simulation model overestimates the leakage up to a factor of ten. Therefore, it can be concluded, that the flow model used in the simulation does not apply to gases under the given conditions.

In future works, the flow used in this model must be adjusted for gases. For example, alternative flow

models based on the free molecular flow could be applied to this problem [9,19]. Further research is needed to describe the gas leakage accurately.

NOMENCLATURE

Acronyms

BFM	Bulk flow model
CFD	Computational fluid dynamics
ifas	Institute for fluid power drives and systems
PFAS	Per- and poly-fluoroalkyl substances
PFOA	Perfluorooctanoic acid
PSD	Power spectral density
PTFE	Polytetrafluoroethylene

Roman Symbols

a_C	1D contact area	m
$C(\boldsymbol{q})$	Surface roughness power spectrum	m^4
$C(q)$	Radially averaged $C(\boldsymbol{q})$	m^4
F	Normal force	N
$\dot{m}$	Mass flow rate	kg/s
p	Relative gas pressure	bar
p_C	Contact pressure	Pa
p_{max}	Exponential fitting parameter	Pa
R	Radius of the ball	m
R_S	Specific gas constant of air	J/K/kg
S_Q	Root mean square roughness	m
T	Temperature	K
t	Time	s
u	Average separation of surfaces	m
u_0	Exponential fitting parameter	m
V	Reference volume	m^3
$\boldsymbol{q}$	Surface roughness wave vector	1/m
q	Radial component of $\boldsymbol{q}$	1/m
x	Position on the slope of the seat	m
x_0	Location of the maximal contact pressure	m

Greek Symbols

γ	Peklenik number	1
ϕ_P	Pressure flow factor	1
ϕ_S	Shear flow factor	1
η	Viscosity	Pa s
ρ	Gas density	kg/m^3
θ	Angle of the slope of the seat	rad
Π	Ratio of the apparent and the real contact area	1
ζ	Magnification	1/m

REFERENCES

[1] Persson BNJ. Contact mechanics for randomly rough surfaces. Surface Science Reports

2006;61(4):201–27.

[2] Yamabe J, Nishimura S. Hydrogen-induced degradation of rubber seals 2018:769–816.

[3] Domingo JL, Nadal M. Human exposure to per- and polyfluoroalkyl substances (PFAS) through drinking water: A review of the recent scientific literature. Environ Res 2019;177:108648.

[4] European Chemicals Agency. Annex XV Restriction Report: Per- and polyfluoroalkyl substances (PFASs); Available from: https://echa.europa.eu/documents/10162/f605d4b5-7c17-7414-8823-b49b9fd43aea (5 July 2023).

[5] Fischer F, Schmitz K, Tiwari A, Persson BNJ. Fluid Leakage in Metallic Seals. Tribology Letters 2020;68(4):1–11.

[6] Persson BNJ. Leakage of metallic seals: role of plastic deformations. Tribology Letters 2016;63(3):42.

[7] Zhang Q, Chen X, Huang Y, Chen Y. Fractal modeling of fluidic leakage through metal sealing surfaces. AIP Advances 2018;8(4).

[8] Tikhomorov VP, Gorlenko OA, Izmerov MA. Simulation of leakage through mechanical sealing device. IOP Conf. Ser.: Mater. Sci. Eng. 2018;327:42047.

[9] Huon C, Tiwari A, Rotella C, Mangiagalli P, Persson BNJ. Air, Helium and Water Leakage in Rubber O-ring Seals with Application to Syringes. Tribol Lett 2022;70(2).

[10] Persson BNJ. Theory of rubber friction and contact mechanics. The Journal of Chemical Physics 2001;115(8):3840–61.

[11] Persson BNJ. Relation between interfacial separation and load: a general theory of contact mechanics. Phys Rev Lett 2007;99(12):125502.

[12] Lorenz B, Persson BNJ. Leak rate of seals: Effective-medium theory and comparison with experiment. The European Physical Journal E 2010;31(2):159–67.

[13] Persson BNJ. Interfacial fluid flow for systems with anisotropic roughness. Eur Phys J E Soft Matter 2020;43(5):25.

[14] Angerhausen J, Woyciniuk M, Murrenhoff H, Schmitz K. Simulation and Experimental Validation of Translational Hydraulic Seal Wear. Tribology International 2019;(123):296–307.

[15] Welch P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. 1967;15(2):70–3.

[16] Scaraggi M, Angerhausen J, Dorogin L, Murrenhoff H, Persson BNJ. Influence of anisotropic surface roughness on lubricated rubber friction: Extended theory and an application to hydraulic seals. Wear 2018;410-411:43–62.

[17] Patir N, Cheng HS. Application of Average Flow Model to Lubrication Between Rough Sliding Surfaces. Journal of Lubrication Technology 1979;101(2):220–9.

[18] Persson BNJ, Scaraggi M. Lubricated sliding dynamics: flow factors and Stribeck curve. Eur Phys J E Soft Matter 2011;34(10):113.

[19] Antman SS, Marsden JE, Sirovich L. Microflows and Nanoflows. New York: Springer-Verlag; 2005.

DEVELOPMENT AND TESTS OF A HYDRAULIC SWIVEL DRIVE WITH HYDROSTATIC BEARINGS

Lutz Müller[1]*, Jürgen Weber[1], Jannis Rühe[2], Frank Biller[2]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*
[2] *Homrich Maschinenbau GmbH, Nieland 3, 23611 Bad Schwartau*

* Corresponding author: Tel.: +49 351 463-33705; E-mail address: lutz.mueller@tu-dresden.de

ABSTRACT

Hydraulic swivel drives from SÜDHYDRAULIK by Homrich Maschinenbau GmbH are widely used for different applications. Especially for oscillating movements with high frequencies and for integration of functions as thrust bearings, the axial load bearing capacity of the drives had to be improved. To achieve this goal, a hydrostatic thrust bearing was designed for the high load series swivel drive SP 15 D. In a first attempt the design was focused on achieving the highest possible load at operating pressure of the swivel drive with an acceptable leakage flow. The analytic calculations were executed according to common literature, but supplemented with FEM simulations and test data to consider elastic deformations. The test of the prototype showed a good agreement with the calculation results.

Keywords: Hydrostatic Axial and Radial Bearing, Swivel Drives

1. INTRODUCTION

Hydraulic swivel drives are widely used for different applications ranging from the movement of simple cargo doors once in some weeks to highly complex testing machines. Especially for testing setups and testing machines the rotational movement has to change the direction with very high frequency of sometimes more than 10 Hz applying also high load moments to test materials or components.

A further challenge is the rotation angle. For the high-speed series it is a maximum of 110 degrees but particularly at testing machines the rotation angle can be very small. Often it is not possible to use separate bearings to compensate axial and radial forces, which leads to additionally high loads on the bearing structure of the swivel drive. Thus, not only a few practical operating conditions of swivel drives can be destructive for rolling bearings, which are typically used in these applications.

To extend the operating limits of the swivel drives resulting from conventional rolling bearings, hydrostatic bearings were developed to carry the axial loads. The bearing configuration and geometry was designed according to commonly available literature, mostly summarized by Weck and Brecher [1], whereas this summary does mostly rely on Optitz [2]. But in the real application the deformations of the bearing geometry due to pressure and loads have to be calculated with FEM to ensure a maximum load bearing with small oil through-flow.

For verification of all calculations a comprehensive test program was passed, that ensured the operating parameters of the newly developed bearing system.

2. BEARING DESIGN

2.1. Design Considerations and Calculations

The swivel drive SP15D has very high torque and can withstand also high radial loads. Therefore it is equipped with the rolling bearings for a high radial load capacity. But the axial loads, such a swivel drive could handle, were limited due to a lubrication concept with a restricted ability for self-centring of the rotary piston.

To improve the axial load capacity, a hydrostatic thrust bearing was integrated with the aims not to change the existing rolling bearing setup, not to change dimensions and manufacturing processes of the existing case design and obtain the highest axial load capacity which is possible. A further restriction was, to achieve a leakage flow of the hydrostatic thrust bearing which should be not more than the one, which was required by the well approved lubrication concept.

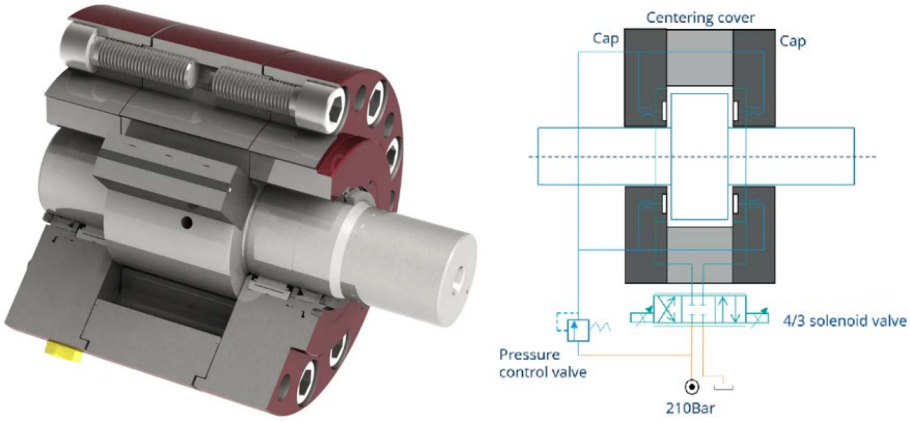


Figure 1: Swivel drive SP 15 D and typical hydraulic layout of the thrust bearing

$$d_m = \frac{d_o + d_i}{2} \quad (1)$$

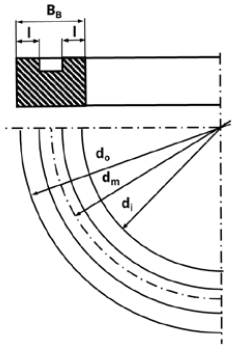
$$b = 2\pi d_m \quad (2)$$

$$A_{eff} = \pi d_m (B_B - l) \quad (3)$$

$$Q = \frac{p b h^3}{12 \eta l} \quad (4)$$

$$R = \frac{p}{Q_L} = R_P + R_K = \frac{p - p_P}{Q_L} + \frac{p_P}{Q_L} \quad (5)$$

$$R = \frac{12\eta l}{b h^3} + \frac{128 \eta l_k}{\pi d_k^4} \quad (6)$$



$$p_P = p \cdot \frac{R_P}{R_K + R_P} \quad (7)$$

$$R_P = \frac{p_P}{Q} \quad (8)$$

$$R_P = \frac{12\eta l}{b h^3} \quad (9)$$

$$h_Q = \sqrt[3]{\frac{12 \eta l Q}{p b}} \quad (10)$$

$$F_{ax} = p_P \cdot A_{eff} \quad (11)$$

$$P_{bear} = p \cdot 2 Q_L \quad (12)$$

Figure 2: Geometry of the bearing grove and design calculations (according to [1] and [2])

Initially the deformation of the shaft was estimated to be much less than the usual manufacturing tolerances of the internal gaps of the swivel drive. This means, the final design of the axial bearing groves could be established as one continuous grove with constant width B_B with sealing edges of equal width l.

For the design of the bearings pocket pressures regulating device at first different concepts were analysed including a membrane type valve. But, since the final design of the bearing should be easy to set up and to maintain but also able to deal with all typical applications of the swivel drives, for instance also only low levels of fluid cleanliness, the concept of capillary restrictors was preferred.
According to [1] and [2] the design calculations are given as shown in ***Figure 2***. These show only one difficulty.

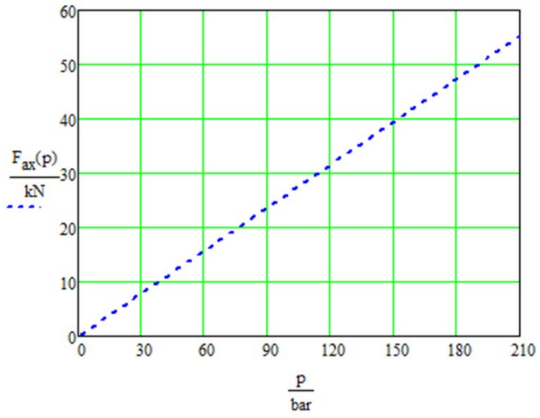


Figure 3: maximum axial load of an ideal hydrostatic bearing which could be realized in a SP15D swivel drive

For optimum work performance of the bearing the resistance ratio $\xi = \frac{R_K}{R_{P0}} \approx 1$ should be nearly 1 [1]. To achieve this, at first, the manufacturing tolerances were measured and a predesign was calculated but the confidence in these data was questionable because the swivel drives structure was suspected to show certain elastic deformation under the high pressures which are normally applied to the bearings and the rotary piston.
Ideally with no deformation and therefore constant capillary and bearing resistances one can calculate the maximum load the bearing would resist by equations (1) to (11) resulting in a maximum axial load F_{ax} depending on the supply pressure p which can be seen in ***Figure 3***.
But in reality the swivel drive will show a distinct pressure dependent deformation, which could be calculated by a FEM simulation with limited effort. But the influences of manufacturing and assembly processes on the practically achievable bearing gap heights were neither possible to be calculated, nor be possible to be measured at the assembled operational swivel drive on another than a fluid mechanic way.
Thus, an experiment was setup to measure the supply pressure dependent volume flow and the resulting symmetrical one-sided bearing gap height.

2.2. Measurement Setup and Results

A test stand according to **Figure 4** (left side) was set up to measure the supply pressure dependent volume flow through the axial hydrostatic bearing. The real test stand can be seen in ***Figure 6***. Since no capillaries were installed yet, nearly the complete pressure drop results from the bearing geometry.

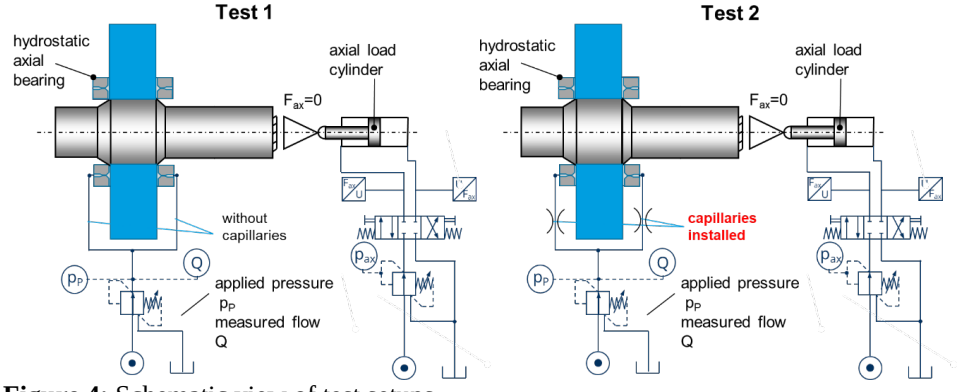


Figure 4: Schematic view of test setups

Because the resistance of the bearing pocket can be calculated with $R_P = \frac{p_P}{Q} = \frac{12\eta l}{bh^3}$, the integral gap height of the bearing can be calculated from the known dynamic viscosity (oil type and measured temperature), the bearings sealing strip width l and B according to the bearings mean circumference b. The mean gap height, which is resulting from these measured data is very precisely evaluated with the following equation. $h_Q = \sqrt[3]{\frac{12\eta l}{Qpb}}$

A big advantage of this procedure is that the behavior of the gap height is obtained for the complete functionally assembled swivel drive, only lacking the capillaries.

As can be seen in ***Figure 5*** left, the flow rate without capillaries is not linear but follows very precisely a quadratic equation. One could suspect non laminar flow behavior but a check of Reynolds-numbers lead to the assumption that this behavior is due to elastic deformation of the swivel drives structure due to the internal pressure at the operational conditions of the bearing. To check this thesis a FEM-model of the complete swivel drive was set up. Simulations were executed for two pocket pressures of 105 bar and 210 bar. The results are presented in the next section.

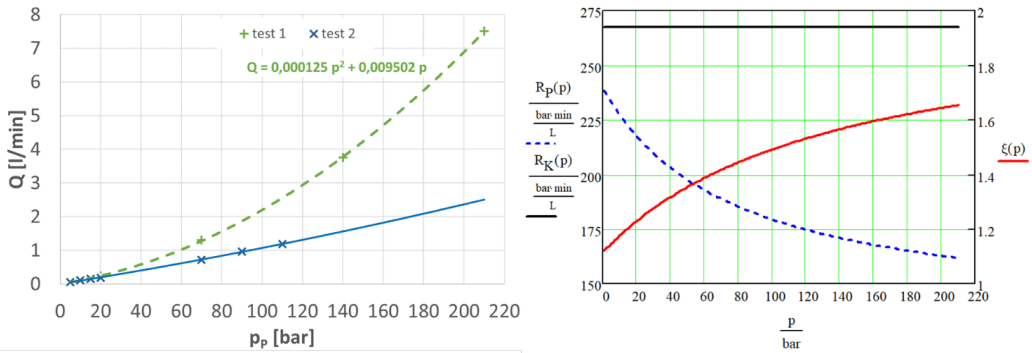


Figure 5: Left: measured flow characteristic for test 1 and 2 accord to **Figure 4** and **Figure 6** and right: the calculated resistances of the bearing pocket (R_P), used capillary (R_K) and resulting ξ

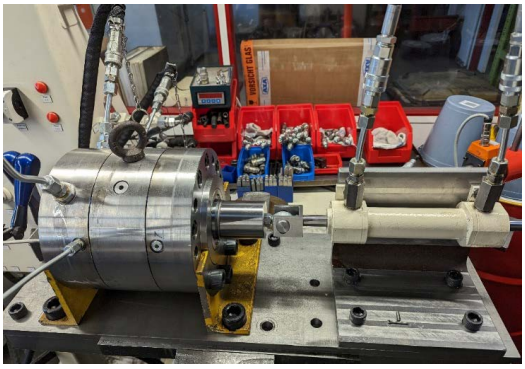


Figure 6: Swivel drive SP 15 D at the test stand for test 1 and 2

But with test 1 already done, it is possible to calculate the supply pressure depending bearings pocket resistance $R_P(p)$ according to equation (9) in ***Figure 2*** as can be seen in ***Figure 5*** on the right side. From the starting value at approximately 1 bar where $R_P(p)$ = 238,5 bar min / L, one can also determine the capillary design. In the current example the capillary resistance was set to be ξ = 1,12.

For test 2 these capillaries were installed in the swivel drives and the through flow test was repeated. The result this test can be seen as blue crosses and blue line in ***Figure 5***. It shows a nearly perfect linear behavior. But since the thesis was not yet proven if the deformation of the swivel drive by internal bearing and load pressures really is the root cause for the pressure dependent gap height, an FEM simulation of the deformation of the structure by internal pressures was done and presented in the next chapter.

2.3. Refining the Design Calculations with FEM

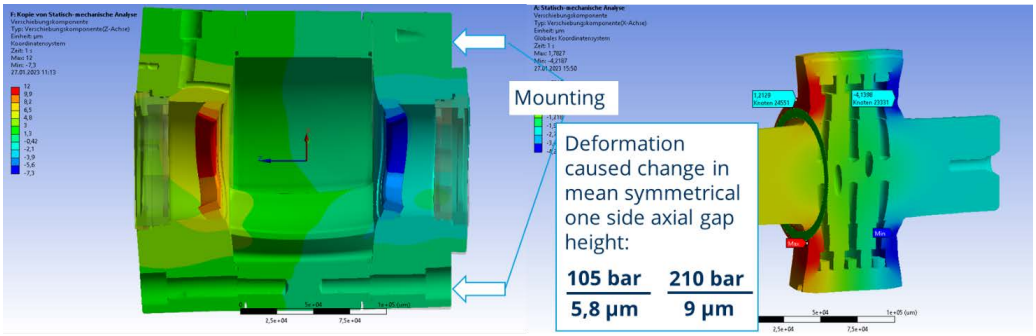


Figure 7: Simulation of the complete deformation of the swivel drive at two pocket pressures

The deformations of the swivel drive under service pressure and pocket pressures of 105 bar and 210 bar are to be seen in **Figure 7**. A typical one sided, symmetrical bearing gap height of 16,5 µm without any loads lead to a deformation which is not straight linear any more up to 210 bar pocket pressure. The value of initially 16,5 µm gap height as well as the calculated pressure dependent gap height (shown as green + in ***Figure 8***) according to equation (10) in ***Figure 2***.

The mean gap height from the FEM-calculated deformation and the approximation formula resulting from these simulations are shown in ***Figure 8***. It is clearly to be seen that the bearing gap height (o FEM) and the calculated bearing gap height from measurement data (+ measurement) show the similar behavior depending on the supply and internal pressure, respectively. That means the deformation of the swivel drives shaft and housing around the hydrostatic bearing is clearly the root cause for the gap behavior.

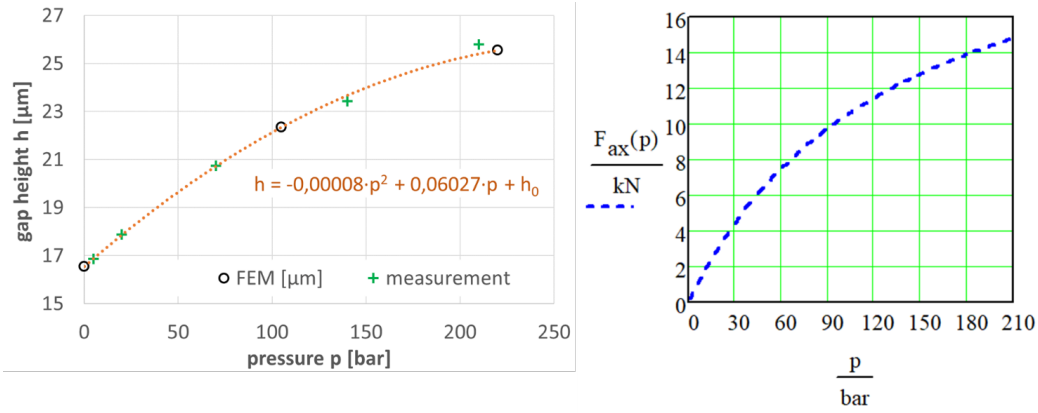


Figure 8: left: Comparison of measured and simulated symmetrical one side bearing gap height right: Calculated maximum axial bearing load

With these data available it is possible to calculate the maximum axial forces the hydrostatic bearing is able to withstand. Therefore the axial force $F_{ax}(p)$ was calculated by using formula (10) whereas p_T was calculated by (7) and A_{eff} results from the geometry of the bearing (3).
The results are given for different bearing supply pressures in the diagram on the right side of **Figure 8**. With a maximum supply pressure of 210 bar a maximum axial force of 15 kN can be

compensated. Since the pressure rating of the swivel drive is given by 300 bar, this value could be further extended to 16,75 kN.

From the equations (5) and (6) given in **Figure 2** also the expected leakage flow of one ($Q_L(p)$) and both two bearing pockets ($2 \cdot Q_L(p)$) can be calculated by the approximation function which was derived from the FEM simulation results shown on the left side of ***Figure 8***. Comparison between the graphs of volume flow of the complete bearing both in ***Figure 9*** and in the left diagram in ***Figure 5*** shows a very good agreement, which is resulting from the very good agreement between FEM simulated bearing gap deformation and measurement of that one shown on the left side of **Figure 8**. To verify the design calculations, which are already assisted by test data and FEM simulations, a further test run was executed and will be discussed in the next chapter.

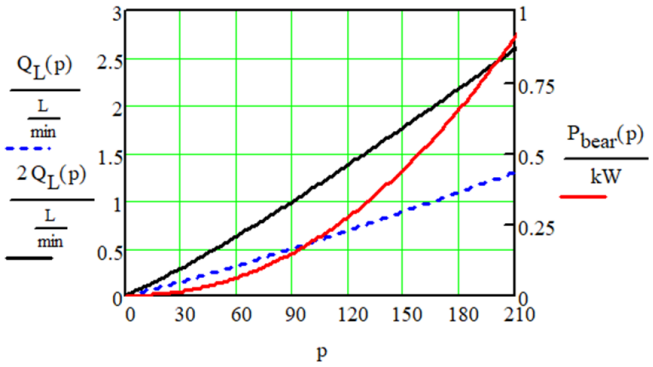


Figure 9: Leakage flow of one (Q_L) and both two bearing pockets ($2 \cdot Q_L$) calculated from FEM deformation and test data

3. EXPERIMENTAL VERIFICATION OF THE DESIGN

For the verification of the design the swivel drive SP15D was set up on the test bench where axial loads could be applied in both directions as can be seen in ***Figure 10*** and ***Figure 11***.

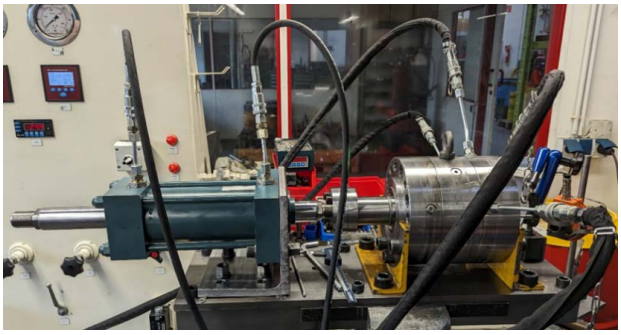


Figure 10: Swivel drive SP 15 D at the test stand for test 3 and 4

In test case 3 and 4 the startup pressure for the swivel drives rotary piston was measured for different axial loads which were applied using a linear cylinder. With this setup it is possible to obtain the axial load limit of the designed thrust bearing because the start-up pressure will significantly increase as soon as mixed or even contact friction will appear when the applied axial force exceeds the load force limit of the hydrostatic bearing. For the test cases 3 and 4 the bearing supply pressure was constantly at 120 bar. For this pressure, according to right diagram in **Figure 8**, one would expect an axial load force of 11,4 kN which could be sustained by the hydrostatic bearing.

The results of these measurements can be seen in **Figure 12**.

At axial load forces up to 12 kN, no increase of the start-up-pressure can be recognized. But further increase of the load lead also to an increase of the start-up-pressure. These data confirm the bearing design to be correct and verify the axial load limit of the bearing according to the right diagram in **Figure 8** too.

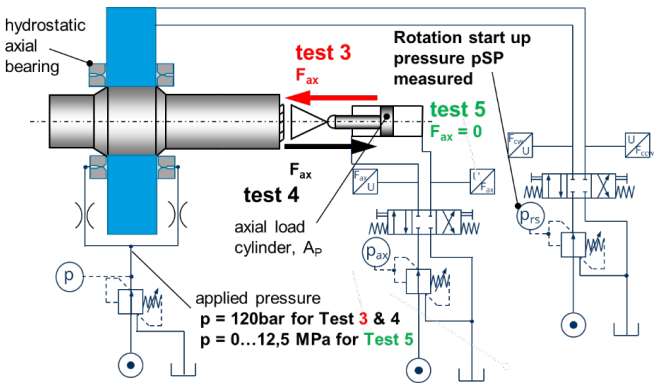


Figure 11: Schematic of the executed Tests 3 and 4

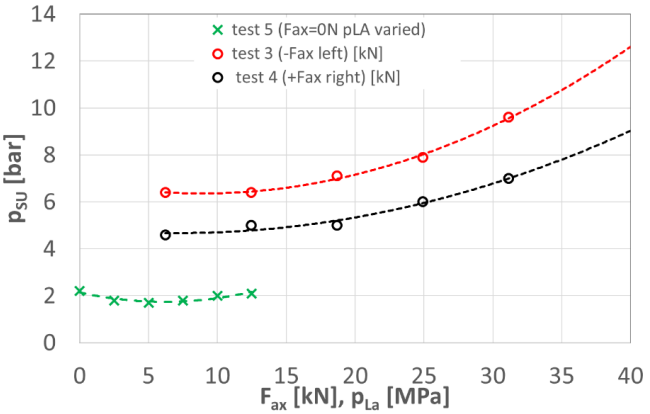


Figure 12: Measurement results from Tests 3 and 4 with a bearing supply pressure of 120 bar

For comparison, the start-up-pressure for test 5 is given. This is essential also for verifying a slight draw back of the test stand. In this test case there was no axial load force applied but the pressure of the hydrostatic bearing was varied. One can see a significantly lower start-up-pressure and also a difference in start-up-pressures between test 3 and 4 which lead to the assumption that the connection between the linear cylinder and the swivel drive is too ridged and the internal friction of the linear cylinders sealing system leads to the high start-up-pressures.
These could be either reduced by calibrated measurements or prevented by a more sophisticated bearing concept. Unfortunately, before the writing of this paper this was not possible to achieve but will be done afterwards to achieve a further development.

4. SUMMARY AND OUTLOOK

In this paper a new hybrid bearing concept for swivel drives was developed, using and verifying analytical design calculations for the hydrostatic axial bearing design whereas the radial bearings were well approved rolling bearings. Because the hydrostatic thrust bearings had to be integrated in the already well approved and reliable bearing concept, the remaining design space lead to limits which had been evaluated and improved by FEM simulations. With these and by test data, the design calculation of the load bearing capability and the required power consumption of the feed flow for the hydrostatic bearings were verified.
Simultaneously to the measurements, the axial bearing was implemented into a swivel drive of an industrial test stand application with over half a year of good experiences.

The design concept presented here was further developed to a not only axial but full hydrostatic axial and radial bearing concept of a new high speed swivel drive HSR14 which is of the same size as the SP15D. This one went already through the design process and is currently set up on a test stand for evaluation.

Supported by:



Federal Ministry
for Economic Affairs
and Climate Action

on the basis of a decision
by the German Bundestag

The presented research activities are part of the project "HYDROTACT – Hydrostatic Bearings for Svivel Drives" (Ref. No. **KK5023209RF1**), which was funded by the German Federal Ministry for Economic Affairs and Climate Action within the ZIM programme.

NOMENCLATURE

A_{eff}	mean effective load bearing area	mm²
b	mean bearing circumference	mm
B_B	width of the hydrostatic bearing	mm
d_i	inner bearing diameter	mm
dK	capillary diameter	mm
d_m	mean bearing diameter	mm
d_o	outer bearing diameter	mm
Fax	axial force	kN
h	mean bearing gap height	µm
h_Q	mean bearing gap height calculated from measurement data	µm
l	sealing edge width of the bearing	mm
l_K	capillary length	mm
p	supply pressure	bar
p_P	bearing pocket pressure	bar
P_{bear}	power required for bearing operation	kW
Q	volume flow	l min^{-1}
Q_L	one sided bearing pocket volume flow	l min^{-1}
R	resistance	bar min l^{-1}
R_K	capillary resistance	bar min l^{-1}
R_P	bearing pocket resistance	bar min l^{-1}
η	dynamic viscosity	kg $(m\ s)^{-1}$

REFERENCES

[1] Weck, M., Brecher, C. (2006) "Werkzeugmaschinen – Konstruktion und Berechnung", 8th edition, ISBN 3-540-22502-1, Springer Berlin Heidelberg New York

[2] Opitz, H. (1969) "Aufbau und Auslegugn Hydrostatischer Lager und Führungen und Konstruktive Gesichtspunkte bei der Gestaltung von Spindellagerungen mit Wälzlagern", Bericht über VDW-Konstrukteur-Arbeitstagung an der RWTH Aachen, Februar 1969

Chapter 5

Tribology

Tribological Design by Molecular Dynamics Simulation – The Influence of Polar Additives on Wall Slip and Bulk Shear

Seyedmajid Mehrnia*, Maximilian Kuhr, and Peter F. Pelz

Chair of Fluid Systems, Technical University of Darmstadt, Darmstadt, Germany

* Corresponding author: Tel.: +49 6151 16-27107; E-mail address: seyedmajid.mehrnia@tu-darmstadt.de

ABSTRACT

This study employed Molecular Dynamics (MD) simulations to examine the tribological impact of polyalkylmethacrylate (PAMA), a polar lubricant additive known for its role as a Viscosity Index (VI) improver, when combined with the non-polar lubricant polyalphaolefin (PAO) 6. Examining the solid-lubricant interface in a confined liquid between iron surfaces with a Couette flow, the research delves into molecular interactions, emphasizing mechanisms governing wall slip for both non-polar and polar molecules. Notably, for non-polar molecules, a singularity in slip length is observed with a molecular-scale gap height resulting in an infinite slip length. However, the addition of polar additives eliminates slip, leading to increased friction. Furthermore, in terms of bulk shear, the introduction of polar additives reduces shear thinning as temperature increases. This dual observation highlights the substantial impact of polar additives on both wall slip behavior and bulk shear properties in the lubricant system.

Keywords: Slip length, Viscosity index improver, Couette flow, PAO 6

1. INTRODUCTION

This paper explores the concept of tribological design through the application of Molecular Dynamics (MD) simulations, with a focus on the influence of polar and non-polar lubricants on wall slip and bulk shear. Using MD simulations, we investigate the behavior of polar additives within the lubricant film, specifically examining their effects on two critical aspects of tribological performance: wall slip and bulk shear.

Wall slip, characterized by the relative motion between the lubricant and solid surfaces, significantly impacts lubrication efficiency and reliability. By studying the molecular interactions at the solid-lubricant interface, we aim to elucidate the mechanisms underlying wall slip and identify strategies to mitigate its occurrence.

Furthermore, it is crucial to analyze the influence of polar additives, such as polyalkylmethacrylate (PAMA), on bulk shear and understand its underlying mechanisms. The presence of these additives can alter the rheological properties of the lubricant film, affecting its ability to sustain load-bearing conditions. Through MD simulations, it investigated the structural and dynamic properties of the lubricant film in the presence of the polar additives, aiming to comprehend their influence on bulk shear and develop insights into optimizing lubricant formulations.

In the investigation of non-polar molecules in lubrication, several studies have explored the phenomenon of wall slip and bulk shear with hydrocarbon oils interacting with metal surfaces [1-5]. For instance, a slip model was proposed for linear hydrocarbon lubricants confined between metal atom surfaces, considering specific surface parameters and shear viscosity [4]. However, this model is limited to ambient temperatures, lacking consideration for the temperature impact on the slip length. In another study, a linear hydrocarbon model with the United-Atoms (UA) model was

employed to simulate iron with Body-Centered Cubic (BCC) lattice sliding walls, utilizing different temperatures for the fluid, though the effects on wall slip were not explicitly addressed [5].

Ewen et al. investigated the impact of surfactant type and coverage on the slip behavior for n-hexadecane liquid, demonstrating that higher slip length corresponds to lower friction [6]. Furthermore, an MD simulation of a hexadecane oil film sheared between two metal surfaces was conducted, exploring the effects of shear rate, film thickness, and surface energy on wall slip [7].

Exploring wall slip and bulk shear phenomena, particularly in the presence of polar lubricant additives, is a captivating yet insufficiently explored aspect in lubrication science. While existing literature provides glimpses into the potential enhancement of slip with specific additives, a comprehensive understanding of variations in slip length and their intricate relationships with polar lubricant additives remains an uncharted domain.

A study [8] investigated how the addition of a friction-modifying substance affects liquid slip, demonstrating a decrease in friction within hydrodynamic contacts. Another investigation [9] explored the impacts of polar Molecules of crude oil in a MD Simulation study, shedding light on their effects on various properties. Additionally, the study delved into the microscopic mechanism of spontaneous imbibition affected by different polarity molecules. The research provided insights into the complex interactions and behaviors of polar molecules within the context of crude oil, contributing to a deeper understanding of their impact on the studied system. In another examination [10], VI improvers were studied, highlighting challenges faced by oil manufacturers but not specifically addressing slip length in hydrocarbon blends.

Before exploring more details about the simulation, it is essential to provide an explanation of the mechanism of the interface system and its crucial parameters. The application of the apparent and real measurements concept proves particularly insightful in the field of MD tribology, providing a valuable framework for understanding the behavior of fluid films between parallel walls. In scenarios where relative sliding occurs with velocity U between the walls separated by distance h the characterization of apparent shear rate $\dot{\gamma}_{\mathrm{app}} = U/h$, and apparent shear stress $\tau_{\mathrm{app}} = \mu\,\dot{\gamma}_{\mathrm{app}}$ as establishes a practical foundation. Notably, the constancy of the real shear stress across the channel height and its equivalence to the wall shear stress ($\tau(z) = \tau_w$) in stationary Couette flow illuminate the prevalence of molecular (viscous) forces in the overall flow field, particularly under conditions of a small Reynolds number. This fundamental understanding underscores the significance of viscous forces in dictating the shear stress distribution, a critical aspect in the study and analysis of tribological phenomena within confined spaces [1,2].

The relationship between bulk shear and wall slip can be expressed through Navier's slip $\tau = k\,u_s$, and Newton's shear $\tau = \mu\,\dot{\gamma}$ equations. These relationships allow us to establish a connection between the slip velocity u_s and the shear rate $\dot{\gamma}$. Specifically, combining Navier's slip equation with Newton's shear equation, we obtain $u_s = \lambda\,\dot{\gamma}$, where λ represents the slip length, cf. Figure 1. This expression highlights the interrelation between slip velocity and shear rate, providing a means to quantify the wall slip based on established tribological principles.

Examining Newton's constitutive equation $\tau = \mu\dot{\gamma}$ for shear and Navier's boundary condition $\tau_{\mathrm{w}} = k u_{\mathrm{s}}$ in the context of phenomenological tribology and rheology proves beneficial. The parameters viscosity μ and friction factor k, or equivalently, slip length:

$$\lambda = \mu/k \tag{1}$$

, are dependent on the shear rate $\dot{\gamma}$, temperature T, and molecular properties, characterized by the (weight-averaged) molecular weight M. Wall slip and bulk shear are function of

$$\text{shear } \mu = \mu(\dot{\gamma}, T, M), \quad \text{slip } \lambda = \lambda(\dot{\gamma}, T, M, \text{solid}). \tag{2}$$

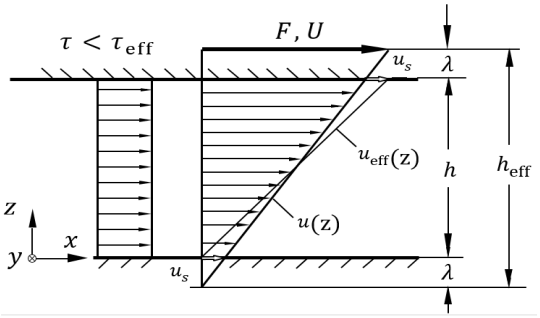


Figure 1: Alteration in the effective gap height caused by the slip length ($h_{\mathrm{eff}} = h + 2\lambda$).

The liquid under investigation in this research is polyalphaolefins (PAO) 6, a base component of synthetic lubricants used in various technical applications. PAO 6 is a hydrocarbon oil primarily composed of branched hydrocarbon molecules, 1-Decane. In addition, Viscosity Index (VI) improvers were used to achieve high-VI hydraulic and gear oils for improved start-up and lubrication at low temperatures. PAMA has long been known in the industry as a key additive to formulate lubricant oils, particularly in the area of green technology-based additives. The structure of 1-Decane tetramer as PAO 6 and PAMA, generated using Avogadro software, are shown in Figure 2.
This paper aimed to extend understanding by exploring Coutte flow under different temperatures, with a particular emphasis on investigating how VI modifiers additves interact with and impact slip and bulk shear mechanisms.

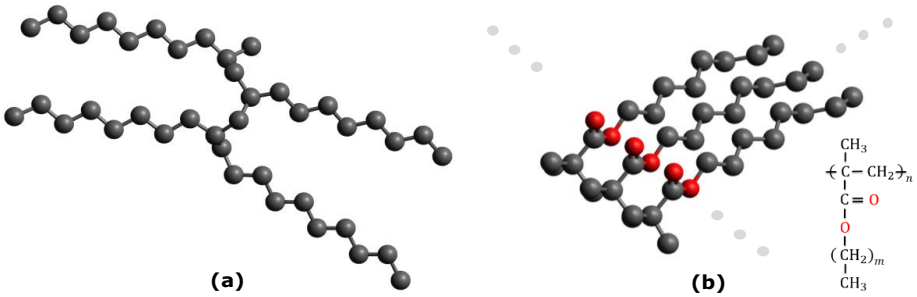


Figure 2: Schematic carbon structures, including (a) a 1-Decane tetramer and (b) a poly alkyl methacrylate (PAMA) comb shape, depicted without their hydrogen atoms.

2. SIMULATION

In this study, MD served as the primary tool for simulating the lubricant-iron system. The classical equation of motion governing a system of atoms was solved to derive the time evolution of the system. To compute the forces within the system, it was imperative to define potential energy functions. The Verlet algorithm, recognized for its precision in simulating atomic motions, was employed as the most accurate numerical scheme in our MD simulation. Additionally, a "neighbor list" numerical scheme was implemented to enhance computational efficiency.

Drawing upon concepts from statistical mechanics, macroscopic considerations were described by classifying microscopic states using an NVE ensemble. To optimize computational time, the United Atom (UA) model was applied specifically to carbon atom groups. Furthermore, a Langevin thermostat was utilized to assist the lubricant atoms in precisely attaining the desired temperature during the simulation.

The Avogadro software was employed for the design of all models, which were subsequently assembled and optimized using the Packmol software. The liquid molecule model adopted in this study incorporates a full atom style, encompassing features such as bond stretching, bending, and torsion. The integration of MD equations of motion utilized the velocity-Verlet algorithm with a time-step of 1.0 fs, specific to the selected force field. The MD simulations were carried out using the open-source LAMMPS molecular dynamics simulator, and a constant NVE integration method was applied to update the velocity and positions of atoms at each time step. The initial macroscopic variables—number of atoms (N), system volume (V), and system energy (E)—were held constant to align with the microcanonical ensemble, ensuring a system trajectory consistent with this ensemble [2,3].

To attain equilibrium, the system underwent relaxation during the initial 0.05 ns of the simulation, ultimately reaching a thermodynamic steady state. Throughout all modeling scenarios, temperature control was uniformly maintained through a Langevin thermostat, mimicking conductive heat flux through the walls to facilitate effective heat removal from the fluid.

Previous research has demonstrated the United Atom (UA) Coarse-Grain models' accuracy in representing hydrocarbon behavior under diverse pressures. The force field selection for MD simulations is critical, and for branched hydrocarbon oil, the NERD-Force-Field was chosen for its applicability to alpha olefins [2]. In the coarse-graining process, CH3, CH2, and CH groups were simplified as single interaction sites. The NERD Force-Field employs an LJ-potential for non-bonded interaction sites, excluding intramolecular interactions within 3 bonds of the same molecule. This non-bonded interaction is characterized by a standard Lennard-Jones 12-6 potential

$$E_{\mathrm{LJ}}(r) = 4\varepsilon \left[\left(\frac{\sigma}{d}\right)^6 - \left(\frac{\sigma}{d}\right)^{12}\right] \ , \ d < d_c \tag{3}$$

where d is the distance among two pseudo atoms, ε is an energy parameter, σ is specified as the length at which the intermolecular energy within the two pseudo atoms is zero. Thus, it could be called van der Waals radius. ε is the well depth and a measure of attraction between two pseudo atoms. In addition, d_c known as cutoff radius for the LJ-interaction potential is set to 1.15 nm. where the interaction parameters for unlike pairs are calculated through the Lorentz-Berthelot combination rules (Eq. 3 and 4) [3]. Non-bonded LJ potential parameters are tabulated in Table 1.

$$\sigma_{ij} = \frac{\sigma_{ii}+\sigma_{jj}}{2} \tag{4}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \tag{5}$$

Table 1. Non-bonded LJ potential parameters [3].

UA group	ε in meV	σ in Å
CH	3.42	3.85
CH_2	3.95	3.93
CH_3	8.96	3.91
Fe	40.98	2.32

For bond stretching and angle bending interactions harmonic potentials are employed:

$$u_r = \frac{k_r}{2}(r - r_0)^2 \tag{6}$$

$$u_\theta = \frac{k_\theta}{2}(\theta - \theta_0)^2 \tag{7}$$

With r_0 and θ_0 being the equilibrium bond length and angle. The torsional potential is described using the following equation:

$$u_\phi = V_0 + V_1(1 + \cos\phi) + V_2(1 - \cos 2\phi) + V_3(1 + \cos 3\phi) \tag{8}$$

The standard Coulombic interaction potential is employed to describe the charges of atoms in the system. This interaction potential governs the electrostatic forces between charged particles and is fundamental in modeling the interactions between atoms in MD simulations. The Trappe force field was employed for modeling polar atoms in the simulation. All the parameters for the intramolecular potential energy functions can be found in Table 2.

$$u_c(r) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \tag{9}$$

Table 2. Intramolecular potential energy functions parameters [3].

Bond stretching potential	
$k_r = 8.31568\ eV/Å^2$	$r_0 = 1.54$ Å
Bond bending potential	
$k_\theta = 5.39\ eV/rad^2$	$\theta_0 = 1.99\ rad$
Torsional potential	
CH_x-CH_2-CH_2-CH_y	
$V_0 = 0\ meV$	$V_1 = 31\ meV$
$V_2 = -23\ meV$	$V_3 = 68\ meV$
CH_x-CH_2-CH-CH_y	
$V_0 = 122\ meV$	$V_1 = 34\ meV$
$V_2 = 12\ meV$	$V_3 = -250\ meV$

In the context of simulating the walls, the approach involved the option of using simple spring potentials or opting for a more advanced force field specifically designed for metals. For this simulation, the Embedded-Atom Method (EAM) potential was used to model the iron walls. The EAM is a popular choice for modeling the force field of metals and alloys. It was developed by Daw and Baskes [2] and is based on the density-functional theory for calculating the properties of realistic metal systems. Embedding energy coupled with a pair potential was used to model the bulk material. In studies involving the shearing of hydrocarbon lubricants by molecular dynamics, a common time step employed was 1.0 fs [2].

3. RESULTS

The density distribution depicted in Figure 3 illustrates the distinct molecular layering of three different branched PAO molecules. The regions proximate to the surfaces form an adsorption layer characterized by heightened fluctuations and oscillations in mass density profiles. The observed peak splitting near the surfaces in the mass density profile is significant and can be attributed to the reduced repulsion area of the united-atom (UA) atoms, allowing these pseudo atoms to closely approach the walls and potentially form stronger bonds with the metal atoms. The measured mass density of PAO 6 at ambient temperature is 0.835 g/mL. The average densities of the molecular structures from the MD simulation presented in Figure 3 align with the experimentally measured density.

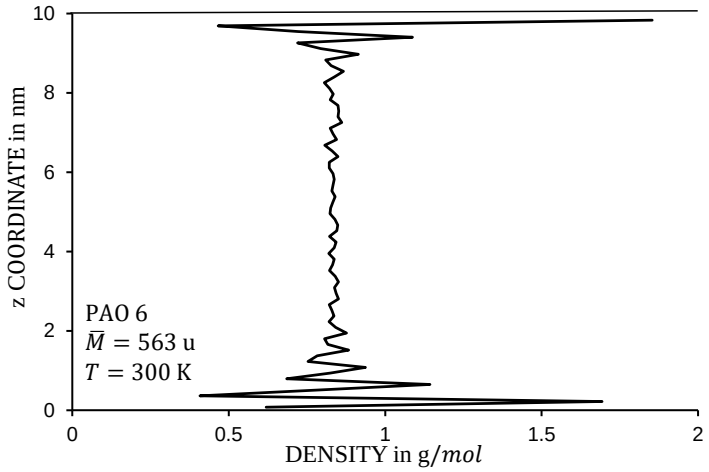


Figure 3: Calculated mass density profile of PAO 6 oil

Contemporary alkene-based lubricants, exemplified by PAO oils, exhibit a dual nature of Newtonian and non-Newtonian behaviors. At low shear rates, viscosity remains constant, forming what is known as the Newtonian plateau. However, beyond a critical shear rate $\dot{\gamma}_c$, there is a swift decrease in viscosity, indicating a shift towards shear-thinning behavior. It's crucial to recognize that the rheological properties of these lubricants are significantly influenced by their molecular structure [14]. Incorporating molecular weight into models, such as the Mark-Houwink relation and the Rouse Model, proves invaluable in establishing a connection between the critical shear rate and the relaxation time of polymeric fluids [15].

Originally designed for short molecules chains, the Rouse model has found applicability in hydraulic fluids, which are characterized by relatively unentangled polymer chains. The model establishes an inverse relationship between critical shear rate $\dot{\gamma}_c$ and the relaxation time r_t of molecules under equilibrium conditions $\dot{\gamma}_c \propto 1/r_t$. Relaxation time is the time for the return of a perturbed system into equilibrium.

The Einstein-Debye equation [11], combined with $\dot{\gamma}_c = 1/r_t$, facilitates the calculation of the relaxation time based on molecular weight and temperature:

$$r_t = \frac{\mu M}{\rho R T} \tag{10}$$

For short-length molecules such as PAO oils, For molecules with short lengths, like PAO oils, both the relaxation time r_t and dynamic viscosity μ exhibit proportionality to molecular weight $\mu \propto M$, $r_t \propto M$. When the product of shear rate and relaxation time is significantly less than 1, $\dot{\gamma} r_t \ll 1$, there is no observable shear rate dependence of shear and slip. In the MD simulation, the relaxation time for the initial Newtonian plateau is determined by the inverse critical shear rate within a 10 ns timeframe, as can be seen in Figure 4. An increase in the number of molecule branches and molecular mass leads to an elevated relaxation time. Additionally, it's noteworthy that higher temperatures result in a higher critical shear rate and, consequently, a shorter relaxation time [2,3].

Figure 4 is shown calculated viscosity profile of PAO 6 oil The fluid's shear viscosity was determined by calculating the ratio of shear stress to the apparent shear rate $\mu = \tau / \dot{\gamma}$ [2]. Utilizing the general equation of the stress tensor for many-body interaction potentials under periodic boundary conditions, the shear stress was computed. In a rheometer, the measured Newtonian viscosity μ_0 of PAO 6 at 300 K was approximately 38 mPa s. Considering that PAO 6 oil is a mixture of different olephins, consist of 1-Decane tetramers, somewhat imprecise comparison can be made.

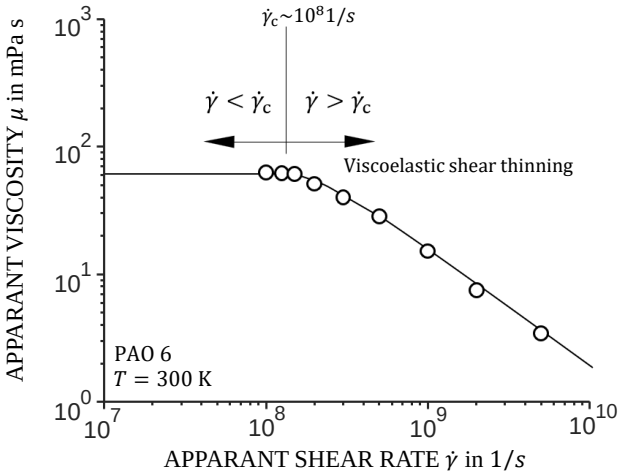


Figure 4: Calculated viscosity profile of PAO 6 oil

3.1. Validation

The accuracy of MD results is closely linked to the force field governing atom interactions, achieving a robust convergence, and the effective implementation of microcanonical ensembles. Additionally, the selection of an appropriate time step for MD simulations plays a crucial role at the simulation's initiation, with the force field determining the requisite time step. Generally, if there is excessive movement among atoms or molecules during two consecutive integrations, the simulation lacks credibility. Table 3 has tabulated a comparison of simulation parameters that shows a good agreement with Ref. [11].

Table 3: Comparison of calculated parameters

Parameter	Ref. [11]	Present work
Density in g/mol	0.832	0.841
Viscosity in mPa.s @ 300 K	38	44
Relaxation time in ns	10.3	9.6

3.2. Wall slip

The relationship between the slip length and the effective molecular length a is evident, as discussed by Pelz et al. [1] through the developed generalized Eyring model for wall slip and Bocquet [12]. This means the slip length scales with the effective molecular length a. It is crucial to emphasize that the assertion $\lambda \sim a$ holds true only under the condition $h \gg a$, an assumption consistently adhered to study by Ref. [1]. However, in cases when $h \sim a$ (length scales in the order of nanometers), the slip length becomes a function of the gap height h as clarified in Figure 5.

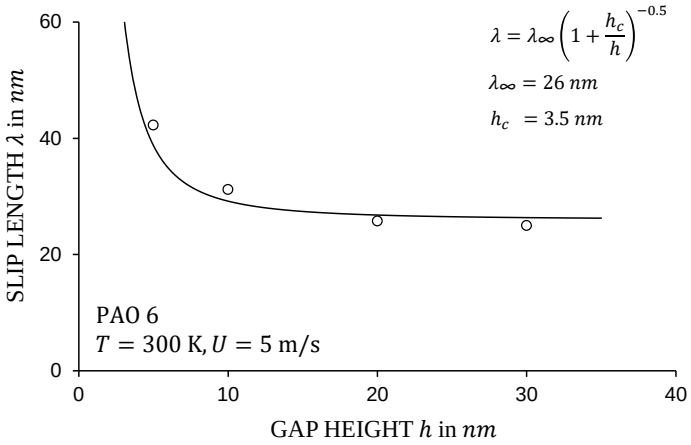


Figure 5: Slip length as a function of gap height when $h \sim a$ for PAO 6 in interaction with iron atoms. As the apparent shear rate is below the critical shear rate, the plotted slip length is independent of the shear rate.

Figure 5 shows the calculated values of slip length λ versus gap heights h and asymptotic expansion Eq. (11) as solid lines. We assumed the critical wall distance $h_c = a = 3.5\ nm$ for PAO molecules. For $h \to 0$ the slip length becomes infinite. This singularity is intuitively understandable, given that for $h \to a$, the slip velocity approaches half of the relative velocity ($w \to U/2$) due to symmetry considerations. The adhesion forces between molecules in the bulk are significantly greater than those between non-polar molecules near the wall and the iron atoms of the wall. Consequently, for $h \to a$, the shear rate approaches zero, resulting in $\lambda \to \infty$.

For $h \to \infty$ we expect a constant value of the slip length $\lambda \to \lambda_\infty$. Hence, an asymptotic expansion reads

$$\lambda = \lambda_\infty \left(1 + \frac{h_c}{h}\right)^{-0.5} \tag{11}$$

This outcome aligns with the findings by Thompson [13], where the functional behavior of λ implies the existence of a universal boundary condition at a solid–liquid interface. When scaling λ by its asymptotic limiting value λ_∞ and shear rate $\dot{\gamma}$ by its critical value $\dot{\gamma}_c$, the data collapses onto an asymptotic relation, described by the equation $\lambda = \lambda_\infty (1 - \dot{\gamma}/\dot{\gamma}_c)^{-0.5}$. The observed behavior suggests that in close proximity to a critical shear rate $\dot{\gamma}_c$, the boundary condition can markedly influence flow behavior at macroscopic distances from the wall, an experimental phenomenon recurrently observed in numerous polymeric systems. This substantial slip is not readily captured by existing phenomenological models. Furthermore, it implies that for flows near $\dot{\gamma}_c$, minor alterations in surface properties can induce significant fluctuations in the apparent boundary condition [13].

In the context of polar additive molecules, their polar functional groups exhibit an affinity for the fluid molecules. This interaction can give rise to a thin layer of fluid molecules in proximity to the surface, commonly referred to as the boundary layer or adsorbed layer. The boundary layer possesses distinct characteristics when compared to the bulk fluid, leading to alterations in viscosity and flow behavior. In this specific scenario, slip values were exceedingly small, approaching zero. This occurs when the fluid molecules strongly bind with the polar molecules, causing them to adhere closely to the surface. Such minimal slip results in heightened friction and increased resistance to flow, leading to reduced flow rates and elevated pressure drop.

Figure 6 shows the presence of a black-marked 1-Decane tetramer molecule situated alongside a metal wall, while being influenced by the presence of polar lubricant additives. This figure provides a side view of the simulation box, depicting a 1-Decane tetramer molecule positioned alongside a stationary surface. Notably, the angle and rotation of the marked black molecule near the stationary

wall are highlighted, offering a comparative analysis between PAO 6 liquid and the blended lubricant PAO 6+VI. The figure vividly illustrates translational rotation in the z-direction, predominantly influenced by the reduced density in close proximity to the stationary wall.
A distinct observation is the molecule's tail flipping while the opposite end remains tethered to the stationary surface. It's noteworthy that the marked black molecule is entangled between polar molecules (oxygen) adhered to the metal surface, adding an additional layer of complexity to the molecular interactions observed in the simulation. This entanglement contributes to the dynamic behavior and rotation of the molecule within the confined space.
As a consequence of this entanglement, there is an increase in friction values, and the slip is observed to be close to zero for the blended lubricant with additives. This phenomenon emphasizes the impact of molecular interactions on the lubrication properties, ultimately influencing the frictional behavior and slip characteristics of the lubricant blend.

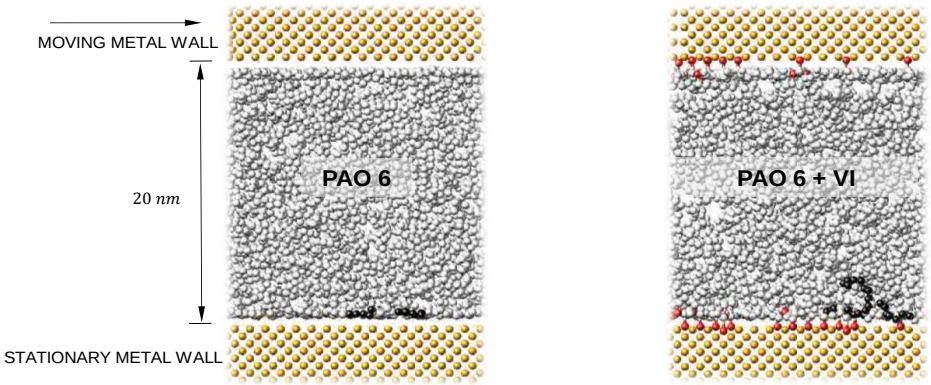


Figure 6: Side view of a 3D simulation box depicting a black marked PAO 6 (1-Decane tetramer) molecule along a metal wall (Fe atoms), with and without PAMA (polar lubricant additives) at 5 ns. Oxygen atoms are highlighted in red, while carbon united group atoms are shown in gray.

3.3. Bulk shear

The presence of PAMA as a VI improver has a significant impact on viscosity, especially at elevated temperatures. The interaction between polar molecules and carbon groups, a characteristic feature of PAMA, plays a crucial role in influencing the viscosity behavior of the lubricant.
To comprehend the role of a VI improver, envision it as a coil spring initially coiled up in a ball at lower temperatures. In this compact configuration, the VI improver has minimal impact on the viscosity of the oil. As temperatures increase, the VI improver exhibits an intriguing behavior— it expands or extends its arms, essentially enlarging its molecular structure. This transformation prevents the oil from thinning out excessively at higher temperatures. The coil spring analogy serves as an effective visualization tool: at low temperatures, the VI improver remains in a condensed state, resembling a coiled spring with minimal influence on oil viscosity. However, as temperatures rise, the VI improver "uncoils," extending its arms to resist excessive thinning of the oil. This property is particularly advantageous in high-temperature environments, as it helps maintain the viscosity of the lubricant within desired limits.

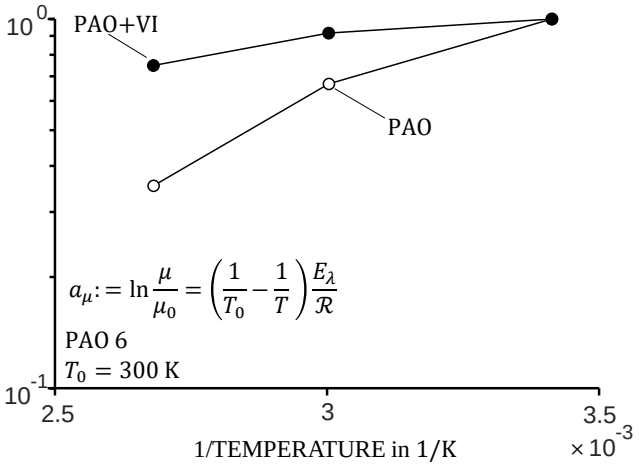


Figure 7: Arrhenius plot of PAO 6 and PAO6+VI to show the effect of polar additive

The Arrhenius plot depicted in Figure 7 provides a visual representation of the impact of incorporating PAMA as a VI additive. This additive plays a crucial role in enhancing the lubricant's viscosity, particularly at elevated temperatures, by dynamically regulating the molecular structure in response to temperature variations. As illustrated in the Arrhenius plot, the viscosity of the lubricant undergoes a notable increase as temperature rises. This is a direct result of the effective action of PAMA as a VI improver. The molecular adjustments facilitated by PAMA prevent excessive thinning of the lubricant at higher temperatures, contributing to enhanced viscosity.

4. CONCLUSION

In the realm of simulations involving confined lubricants, diverse methodologies are employed to assess slip. One such approach entails analyzing the velocity profile of lubricant molecules in proximity to solid surfaces to determine slip velocity. The earlier discussion on singularity in slip length is also touched upon, providing a comprehensive overview of the multifaceted influences in confined lubricant systems. Notably, findings from this analysis indicate minimal slip values, primarily attributed to the presence of polar atoms on wall surfaces. The calculation of viscosity for confined polar lubricant additives combined with PAO oil between metal surfaces in MD simulations demands a consideration of interactions among lubricant additives, PAO oil, and metal surfaces. This involves applying a shear gradient, computing the stress tensor, and analyzing stress autocorrelation to estimate viscosity. The resultant viscosity values underscore the substantial impact of polar molecules on viscosity, chiefly stemming from coil expansion. The concept of coil expansion posits that a polymer maintains a coiled conformation at lower temperatures and subsequently expands as solubility increases at higher temperatures, leading to heightened viscosity. Additionally, the interaction between polar molecules and carbon groups contributes to the increased viscosity observed.

ACKNOWLEDGMENT

We kindly acknowledge the financial support by the German Research Foundation (DFG) within the Collaborative Research Centre (CRC) 1194 "Interaction of Transport and Wetting Processes" - Project-ID 265191195, sub-project C06. The computing time was granted by Lichtenberg HPC computer resources at TU Darmstadt. We appreciate the Hessian Competence Center for High-

Performance Computing – funded by the Hessen State Ministry of Higher Education, Research, and the Arts – for helpful advice.

REFERENCES

[1] Pelz P, Corneli T, Mehrnia S, Kuhr M (2022) Temperature-dependent wall slip of Newtonian lubricants. Journal of Fluid Mechanics 948 A8

[2] Mehrnia S, Pelz P (2021) Slip length of branched hydrocarbon oils confined between iron surfaces. Journal of Molecular Liquids 336 116589

[3] Mehrnia S, Pelz P (2022) Tribological design by Molecular Dynamics simulation– The influence of molecular structure on wall slip and bulk shear. Journal of Chemical Engineering & Technology 46 (1) 95-101

[4] Savio D, Fillot N, Vergne P, Zaccheddu M (2012) A Model for Wall Slip Prediction of Confined n-Alkanes: Effect of Wall-Fluid Interaction Versus Fluid Resistance. Tribology Letters 46 11-22

[5] Ghaffari M, Zhang Y, Xiao S (2017) Molecular dynamics modeling and simulation of lubricant between sliding solids. Journal of Micromechanics and Molecular Physics 2 2

[6] Ewen J, Kannam S, Todd B, Dini D (2018) Slip of Alkanes Confined between Surfactant Monolayers Adsorbed on Solid Surfaces. Langmuir 34 3864-3873

[7] Jabbarzadeh A, Atkinson J, and Tanner R (1999) Wall slip in the molecular dynamics simulation of thin films of hexadecane. Journal of Chemical Physics 110 2612-2620

[8] Choo JH, Forrest AK, Spikes HA (2007) Influence of Organic Friction Modifier on Liquid Slip: A New Mechanism of Organic Friction Modifier Action. Tribology Letters 27 239-244

[9] Wang S, Wang J, Liu H, Liu F (2021) Impacts of Polar Molecules of Crude Oil on Spontaneous Imbibition in Calcite Nanoslit: A Molecular Dynamics Simulation Study. Energy & Fuels 35 (17) 13671-13686

[10] Cusseau P, Vergne P, Martinie L (2020) Relationship between Film Forming Capability and Rheology of Lubricants with VI Improvers. Aus Wissenschaft und Forschung / TAE-Plenarvorträge

[11] Liu P, Lu J, Yu H, Ren N (2017) Lubricant shear thinning behavior correlated with variation of radius of gyration via molecular dynamics simulations. J. Chem. Phys. 147 084904

[12] Bocquet L, Barrat J-L (2007) Flow boundary conditions from nano- to micro-scales. Soft Matter 1235 3(6) 685

[13] Thompson P, Troian S (1997) A general boundary condition for liquid flow at solid surfaces. Nature 389 360-362

NUMERICAL STUDY ON ABRASIVE WEAR OF RECIPROCATING SEALS UNDER MIXED LUBRICATION CONDITIONS

peer reviewed

Jiehao Wang[1], Jing Li[1]*, Yaobao Yin[1]

[1] *School of Mechanical Engineering, Tongji University, Caoan Road 4800, 201804 Shanghai*

* Corresponding author: Tel.: +86 18964895619; E-mail address: cynthia_li@tongji.edu.cn

ABSTRACT

The reciprocating seals of the actuator are under mixed lubrication conditions, where rough peak contact and oil film lubrication coexist. The seal wear during operation has a severe impact on its sealing performance. The actual wear phenomena usually include many forms, such as abrasive wear, adhesive wear, etc. The classical Archard model based on adhesive wear theory is commonly used for seal wear analysis, while research on abrasive wear is still rare. This study establishes a three-body micro-contact model of the piston rod, abrasive and wear ring based on fractal theory and finite element simulation to analyze the contact characteristics of abrasives in the sealing separation. The influence of fractal surface parameters, abrasive size, shape, and material hardness on abrasive motion, contact, and friction stress are investigated. Next, the percolation channel model of the sealing surface is established to investigate the mechanism of topography characteristics on the leakage after wear. The calculation model and method for reciprocating seal leakage can be obtained based on percolation theory. This model and theoretical analysis aim to clarify the wear failure life assessment, full life cycle management, and anti-degradation design of reciprocating seals.

Keywords: Reciprocating seal, Mixed lubrication, Abrasive wear, Surface topography, Percolation theory

1. INTRODUCTION

The reciprocating seal is a prevalent sealing mechanism in hydraulic systems, extensively applied across diverse fields such as aerospace, engineering machinery, and the chemical and pharmaceutical industries. This seal type is highly valued for its low leakage rate, superior reliability, and broad applicability. The core components of a reciprocating seal typically include an O-ring, a wear ring, and various auxiliary components. During the seal's operation, relative motion is produced between the wear ring and the piston rod surface, leading to friction [1,2]. The reciprocating seal's performance is contingent upon maintaining its sealing capability during the reciprocating motion while minimizing friction to limit wear.

Abrasive wear is a phenomenon wherein external hard particles, protruding objects, or surface rough peaks cause surface material to dislodge during friction [3]. There are three main types of abrasive wear: (1) Two-body abrasive wear occurs when abrasive particles move relative to a solid surface, producing scratches or furrows on the surface according to the direction of abrasive movement; (2) Low-stress abrasive wear is observed in a pair of friction pairs when the rough peak of the hard surface rides the abrasive particle on the softer surface; (3)Three-body abrasive wear occurs when external wear particles or metal particles move between two friction surfaces, generating extremely high contact stress and causing the surface material to spall [4].

Reciprocating seals, typically employing contact sealing, are particularly prone to abrasive wear due

to the working conditions. The seal gap allows the percolation of lubricating oil, placing the sealing surface in a mixed lubrication state, where the compressed oil and rough peaks bear the normal load, as shown in **Figure 1**. Long-term operation of the hydraulic system can produce tiny metal particles that flow into the sealing system with the oil, leading to a medium rich in impurities. The primary failure mode is severe sealing surface wear, leading to percolation channel formation and significant leakage. This can easily result in substantial leaks and rapid seal failure, drastically reducing the service life of the reciprocating seal [5]. The amount of wear produced by abrasive wear depends on the abrasive particles' hardness, strength, shape, sharpness, and size. Wear of the seal alters the topography of the sealing surface, producing separation in areas where continuous contact clusters form on the original sealing surface. These gaps can create percolation channels through the sealing surface, resulting in seal failure and leakage [6,7]. Thus, investigating the surface topography, deformation, friction, and wear mechanism of the reciprocating seal end face under abrasive wear conditions is of substantial importance.

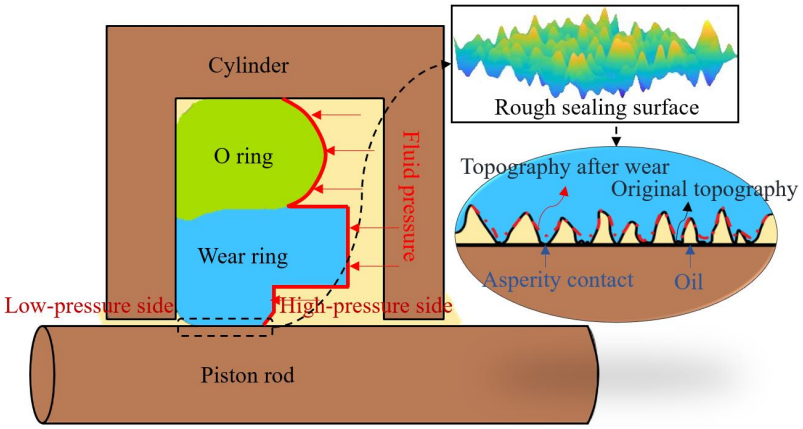


Figure 1: Working condition of the reciprocating seal

The sealing surface topography evaluates the seal's processing and manufacturing and helps explore the nature of lubrication, contact, and wear mechanisms. The fractal theory provides a practical tool for modelling the contour curve of a rough surface [8]. The advantage is its independence from the measurement scale, and the simulated curve exhibits regularity, making it suitable for mathematical model calculations of friction and wear under mixed lubrication. Under abrasive conditions, the friction pair often transitions from two- to three-body wear [9,10]. It has been observed that abrasive particles become embedded in the polytetrafluoroethylene (PTFE) friction pairs used for sealing, causing wear on the surface of the metal piston rod. Finite element and numerical analysis can effectively analyze the sealing surface's motion mode, stress-strain, and pressure changes during wear.

The influence mechanism of the rough surface profile parameters of the wear ring and topography evolution following wear on sealing performance should be considered. However, the current study calculates the leakage rate by studying the oil film thickness of the sealing surface using the average Reynolds equation [11]. This approach only allows for an estimate of the leakage state and fails to reflect the sealing nature of the continuous contact area on the sealing surface to prevent oil leakage. In this study, we propose combining percolation theory with the topography evolution of the sealing surface, constructing a sealing leakage channel to determine the flow rate at the critical constriction, and calculating the leakage amount. When combined with the friction and wear mechanism, this method enables the construction of a microscale reciprocating sealing leakage analysis method.

2. NUMERICAL MODEL

2.1. Fractal theory

Every machined surface, when magnified sufficiently, reveals a degree of roughness. This is true for the sealing surface as well, and as such, when sealing surfaces make contact, they only do so at specific points. The non-contact areas connect to form a percolation channel, the fundamental source of leakage. On a microscopic scale, the stress and deformation of the contact points and the percolation channel's size and distribution are intimately linked to the surface topography.

Currently, the usage of fractal theory to study topography relies on the Weierstrass-Mandelbrot (W-M) function. This function exhibits randomness, self-similarity, and continuity and is non-differentiable at all points, which aligns perfectly with the traits of machined surface contours [12,13]. The expression for the three-dimensional fractal surface model in the Cartesian coordinate system is:

$$z(x,y) = L\left(\frac{G}{L}\right)^{D-2}\left(\frac{\ln\gamma}{M}\right)^{1/2}\sum_{m=1}^{M}\sum_{n=0}^{n_{\max}}\left(k\gamma^{n}\right)^{D-3} \times\left\{\cos\phi_n - \cos\left[\frac{2\pi\gamma^{n}\left(x^2+y^2\right)^{1/2}}{L}\times\cos\left(\tan^{-t}\left(\frac{y}{x}\right)-\frac{\pi m}{M}\right)+\phi_n\right]\right\} \tag{1}$$

where the variables x, y, z represent the horizontal coordinates, vertical coordinates, and surface height, respectively; M represents the number of superimposed wave layers in the fractal surface; L is the sample length; γ^n denotes the spatial frequency; n is the frequency base number; ϕ_n is a random number between 0 and 2π for the superposition of waves; D is a fractal dimension, represents the number of fractal cones; G is the scale parameter, indicating the relative height of the surface features. This model provides a comprehensive mathematical representation of the surface topography, allowing for the nuanced characterization of a rough surface's topography characteristics. When M is set to 1 and y to 0, the three-dimensional fractal surface model can be simplified to obtain a two-dimensional contour function of the fractal surface in the X-Z plane, which is

$$z(x) = L\left(\frac{G}{L}\right)^{D-2}\left(\frac{\ln\gamma}{M}\right)^{1/2}\sum_{n=0}^{n_{\max}}\gamma^{(D-3)n}\left[\cos\phi - \cos\left(\frac{2\pi\gamma^{n}x}{L}-\phi\right)\right] \tag{2}$$

Since the roughness of the piston rod is significantly smaller than that of the wear ring, the sealing contact surface of the reciprocating seal can be approximated as the extrusion contact between a fractal surface and a smooth, rigid surface. The following assumptions are adopted to simplify the contact process [14]: (a) the fractal surface is isotropic; (b) the contact at the asperities follows the Hertzian contact hypothesis, and the asperities are spherical; (c) the axial deformation between the asperities is ignored. The curvature radius of a single asperity can be determined as follows:

$$R = \left|\frac{1}{\left|d^2z/dx^2\right|_{x=}}\right| = \frac{l^{(D-1)}}{4\pi^2G^{(D-2)}\left(\ln\gamma\right)^{1/2}} \tag{3}$$

where l represents the base wavelength of the asperity on the reference plane. Once the contact is quantified, contact stress distribution can be determined using classical contact mechanics models.

2.2. Abrasive wear

In this study, only the wear of the wear ring is considered while disregarding the wear of the piston rod. There are three main wear mechanisms: (a) The normal load will force the abrasive pressure on the surface, and the surface will be cut through the furrow action of the abrasive in relative motion.

(b) The abrasive particles move between the two contact surfaces, like cutting tools, causing wear on the softer surface; (c) The abrasive particles move quickly between the two contact surfaces, producing an impact force on the friction surface and resulting in wear. Generally, the rod speed of the actuator is not too large, and the wear forms are mainly the first two. The Archard model quantitatively describes the correlation between the wear volume of material and variables such as load, sliding distance, and hardness [15]. This model has been extensively employed in studies related to seal wear due to its simplicity and accuracy under certain conditions, and its expression is

$$W_a = \frac{K_a F_N s}{H_m} \tag{4}$$

where W_a is the wear volume, K_a is the wear coefficient, F_N is the nominal load, s is the sliding distance, and H_m is the hardness of the softer material. The Archard's model can also be modified to calculate the wear depth at a specific location, that is

$$h_w = k_a p_c s \tag{5}$$

where $k_a = K_a / H_m$ is the non-dimensional wear coefficient, p_c is the asperity pressure. In addition, a finite element analysis (FEA) model for Archard wear can be established by incorporating the APDL command flow into the contact pair of the piston rod and wear ring. The operating conditions for the FEA model are depicted in **Figure 2**. The wear volume on the rough surface of the wear ring can be calculated to obtain the topograph evolution of the wear ring's surface over time.

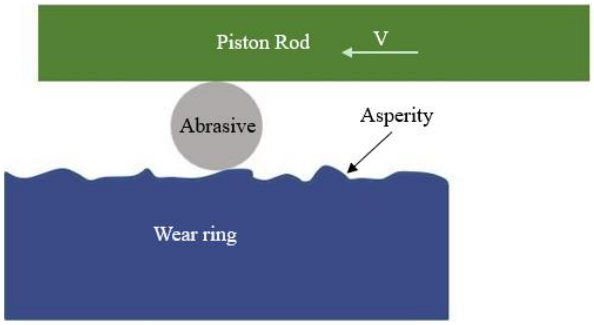


Figure 2: FEA model of abrasive wear

In the context of a contact problem between a smooth, rigid surface and an elastomeric rough surface, the calculation can be performed based on the elastic deformation of the asperities on the surface. The elastic deformation of an asperity can be represented as:

$$\omega = \sigma_a - h \tag{6}$$

where $\sigma_a = z(l/2)$ is the asperity height, h is the separation between the fundamental plane of the rough surface and the rigid plane. According to [16], it is crucial to evaluate whether the asperity deformation is less than the critical deformation associated with elastic deformation. The material properties and contact conditions determine the critical deformation of the asperity, that is

$$\omega_e = \left(\frac{\pi K_h H_m}{2E}\right)^2 R = \left(\frac{K_h \phi}{4}\right)^2 \frac{l^{D-1}}{G^{D-2}\sqrt{\ln \gamma}} \tag{7}$$

where K_h is the hardness coefficient of the material, generally obtained by empirical formula $K_h = 0.454 + 0.41\nu$, ν is the Poisson ratio of the material, H_m and E is the Brinell hardness and elastic modulus, respectively, determined by the material characteristics. The contact pressure generated by the elastic deformation of the asperity is:

$$p_c = \frac{4E}{3\pi}\sqrt{\frac{\omega}{R}} \tag{8}$$

2.3. Percolation theory

The leakage condition in a contact seal occurs when the gaps on the sealing surface interconnect to create a percolation channel across the surface. Since the irregular interaction of asperities on the sealing surface, the separation presents a complex and random shape. The percolation theory describes fluid's random transfer and flow process in random porous media. The formation of a percolation channel is related to the magnification factor of the sealing surface [17]. As the magnification factor ζ increases, more roughness details become evident, and the contact area on the sealing surface correspondingly decreases.

The critical magnification factor ζ_c represents the magnification factor at which a percolation channel running through the sealing surface for the first time appears. **Figure 3** visually represents how the sealing surface's contact state evolves with an increasing magnification factor. The grid on the surface is the resolution, which inversely correlates with the magnification factor: as the magnification factor increases, the resolution decreases, revealing more details of the surface topography. When the magnification continues to increase, the percolation channel also expands. The primary focus of this study is to calculate the leakage rate through the percolation channel, mainly at the critical magnification factor ζ_c. Under the differential pressure effect, the fluid flows through the channels from the high-pressure to the low-pressure side.

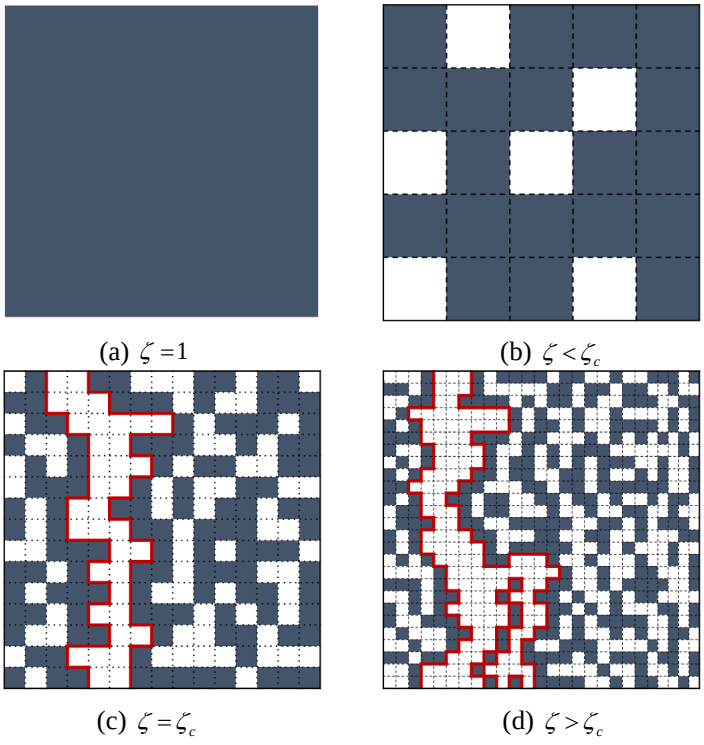


Figure 3: Relationship between separation at sealing surface and magnification factor

In this paper, to reduce the computational load of the model, the effect of differential pressure flow,

which is the leading cause of leakage in reciprocating seals, is mainly considered. The effect of shear flow is much smaller, so it is ignored. The volume flow rate of the leaking oil through the critical constriction of the penetration channel is

$$\dot{Q} = \alpha \frac{u_1^3(\zeta_c)}{12\eta} \Delta p \tag{9}$$

where α is the correction factor for the shape of the critical constriction, u_1 is the separation height at the critical constriction, ζ_c is the magnification factor at which the critical constriction occurs, η is the oil viscosity, Δp is the pressure difference between two sides of the seal. The relative contact area of a surface under a specific magnification factor is defined as the ratio of the actual contact area observed under this magnification factor to the contact area as observed under the original size, which can be expressed as

$$\frac{A(\zeta)}{A_0} = \frac{1}{\sqrt{\pi G_0}} \int_0^{P_0} d\sigma e^{-\sigma^2/(4G_0)} \tag{10}$$

where

$$G_0(\zeta) = \frac{\pi}{4} E'^2 \int_{q_0}^{\zeta q_0} dq q^3 C(q) \tag{11}$$

and the power spectral density function of the rough surface is

$$C(q) = \frac{1}{2\pi q} C^{1-d}(q) \tag{12}$$

According to [18], with the increase of the magnification factor of the rough surface, when the relative contact area decreases to 0.4, the percolation channel through the sealing surface begins to appear. Therefore, $A(\zeta)/A_0 = 0.4$ is the basis for determining the critical magnification ζ_c. The separation height at the critical constriction $u_1(\zeta)$ of the percolation channel is a crucial parameter in calculating the leakage rate through the percolation channel. When the magnification factor is reduced, some areas of the surface will re-contact; at this time, the average separation height of the channel decreases, and the constriction height can be expressed as

$$u_1(\zeta) = \bar{u}(\zeta) + \bar{u}'(\zeta) A(\zeta)/A'(\zeta) \tag{13}$$

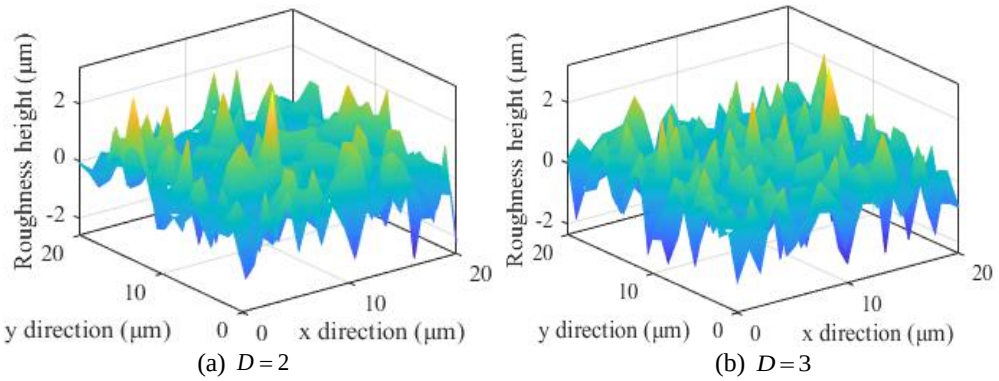
where $\bar{u}(\zeta)$ is the average separation height between the sealing surfaces, $\bar{u}'(\zeta)$ and $A'(\zeta)$ are the derivatives of $\bar{u}(\zeta)$ and $A(\zeta)$ with respect to ζ. The specific solving process of these parameters can be referred to [19].

3. SIMULATION RESULTS AND DISCUSSION

According to fractal theory, different fractal dimensions and scale parameters are selected to investigate the impact on the three-dimensional microscopic characterization of the sealing surface, as shown in **Figure 4** and **Figure 5**.

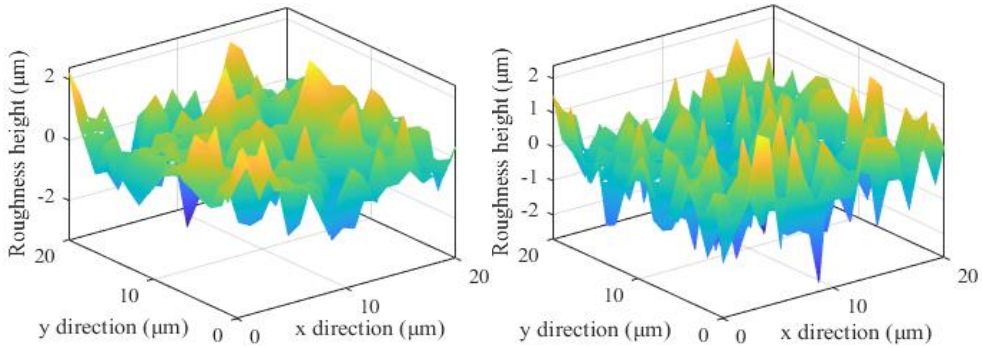
The fractal dimensions significantly affect the surface's form and structure, predominantly in roughness and texture. An increased number of fractal dimensions results in a more complex and rougher surface. The surface with fewer fractal dimensions tends to be smoother. The scale parameter determines the magnification of the rough surface. As this scale parameter decreases, the rough surface reveals more details. The rough surfaces generated have a direct impact on the contact state of the sealing surface. To analyze the wear state of the rough surface, $A(\zeta)/A_0 = 0.5$ is used as the

criterion to determine $D = 2.3$, $G = 10^{-7}$ at the compression rate of 15%.



(a) $D = 2$ (b) $D = 3$

Figure 4: Rough surface topography under different fractal dimensions



(a) $G = 10^{-6}$ (b) $G = 10^{-8}$

Figure 5: Rough surface topography under different scale parameters

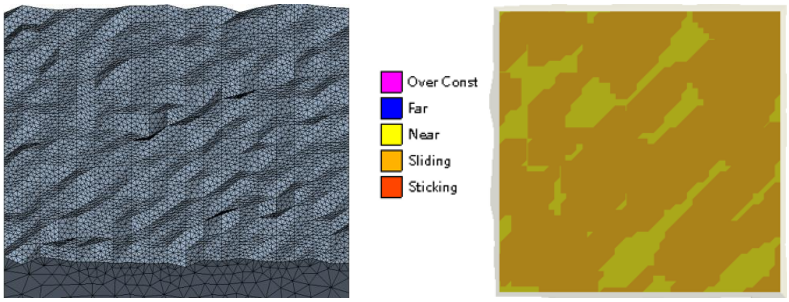
The simulation parameters of the reciprocating seal in FEA are shown in **Table 1**. The piston rod and abrasive particle are made of structural steel, and the wear ring is made of PTFE. The O-ring is made of nitrile butadiene rubber (NBR), and the Mooney-Rivlin model describes its hyperelasticity.

Table 1: Simulation parameters of the reciprocating seal

Simulation parameter	Value	Simulation parameter	Value
Fluid viscosity	$0.0386 Pa \cdot s$	Sealed pressure	$15 MPa$
RMS roughness	$1 \mu m$	Elastic modulus	$550 MPa$
Rod speed	$0.05 m/s$	Stroke length	$200 mm$
Rod diameter	$18 mm$	Wear coefficient	$7.2 \times 10^{-6}\ mm^3/Nm$
Number of strokes	0,75,150,225,300	C_{10}, C_{01} of the O-ring	$0.2 Mpa$, $6.3 Mpa$
Fractal dimension	2.3	Scale parameter	10^{-7}

The rough surfaces generated in MATLAB are imported into ANSYS Workbench. We simulated the compression state of the seal by extruding the hyperelasticity rough surface with a rigid, smooth surface. **Figure 6** presents the FEA model of the created rough topography and the consequent contact state of the compressed rough surface. The axial deformation of the asperities results in most of the surface area being in a sliding state, denoted by the orange colour. The yellow regions indicate the separation between the rough and rigid surfaces, where no physical contact occurs. Most regions are

in contact at this stage, and the percolation channel has not yet formed. The sealing components within the area exhibit a robust resistance to oil leakage.



(a) FEA model (b) Contact state

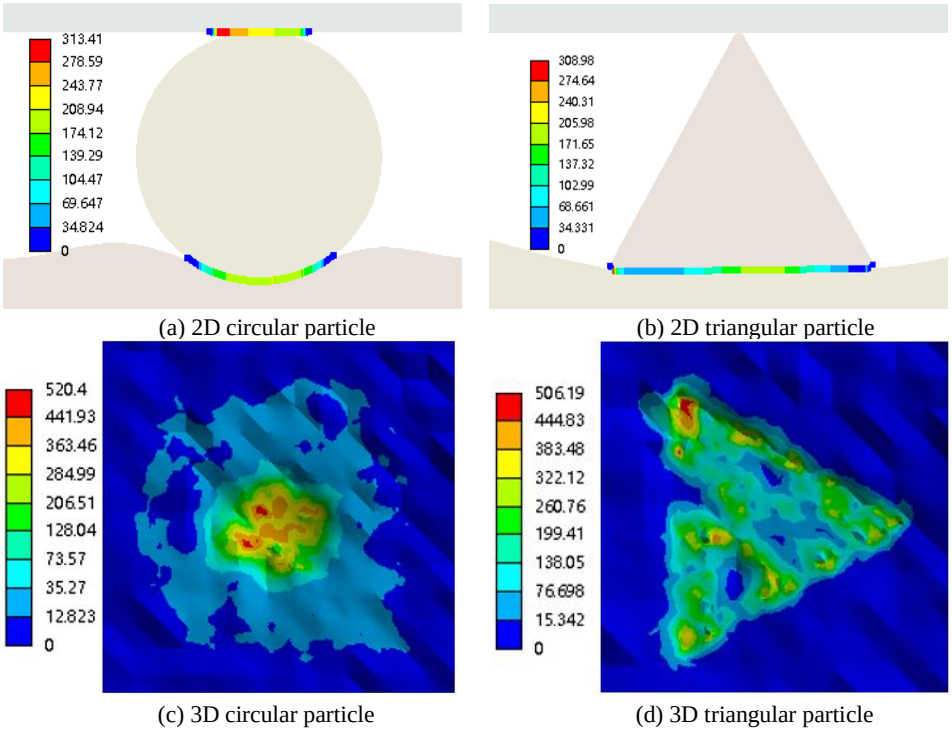
Figure 6: FEA model and contact state of rough surface

The abrasive particles are incorporated into the system based on the direct compression of the rough and rigid surfaces. In hydraulic actuator systems, the size of the impurities mixed into the oil is usually 6-20 μm. This paper selects circular particles with a diameter of 6 μm and equilateral triangular particles with a side length of 10 μm. The particles are placed at the centre of a rough surface with a size of 20μm*20μm to examine the resulting Von Mises stresses following compression, as illustrated in **Figure 7**. Abrasive wear occurs in each local area of the entire sealing surface, and it can be considered that if the percolation channel is formed in the maximum contact pressure area, the channel in other areas has already been formed. Therefore, the selected local area has the highest contact pressure in the sealing length, and there is only one abrasive particle in this area. The total leakage rate can be determined by adding the leakage rate of all local areas within the circumference length of the seal.

After the wear ring is compressed by applying a load on the piston rod, the piston rod begins to move horizontally to the left. The 3D model primarily simulates the state of compression of the rough surface under compression by abrasive particles. Owing to the 3D topography revealing more detailed asperities compared to the 2D topography, the average Von Mises stress of the compressed 3D model is more significant than that of the 2D model.

In the initial compression stages, circular particles primarily interact with the wear ring through rolling contact. As the compression rate increases, the blocking effect of the asperities on the particles is enhanced, which causes the circular wear particles to transition from rolling to sliding relative to the wear ring, eventually becoming embedded in the wear ring surface and generating stress on the sealing surface. Throughout this process, the circular wear particle moves from the rough peak to the valley until the frictional force cannot overcome the asperity resistance to the wear particle, achieving stability. Under the reciprocating motion of the piston rod, the circular particle wears down the asperity on both sides until the roughness height decreases and the wear particle resumes movement following a reduction in contact stress.

The triangular wear particle is mainly sliding contact in the compression process. The upper surface of the wear particle experiences sliding friction, which embeds the corner of the triangular wear particle into the wear ring surface during the movement of the piston rod. After that, the triangular particle remains relatively static with the base surface in the subsequent friction. Given the minimal friction between the wear particle and the piston rod, the triangular wear particle does not roll with the relative movement of the two contact surfaces. Hence, abrasive particles tend to produce scratches on the wear ring surface, characterizing a pattern of cutting wear.

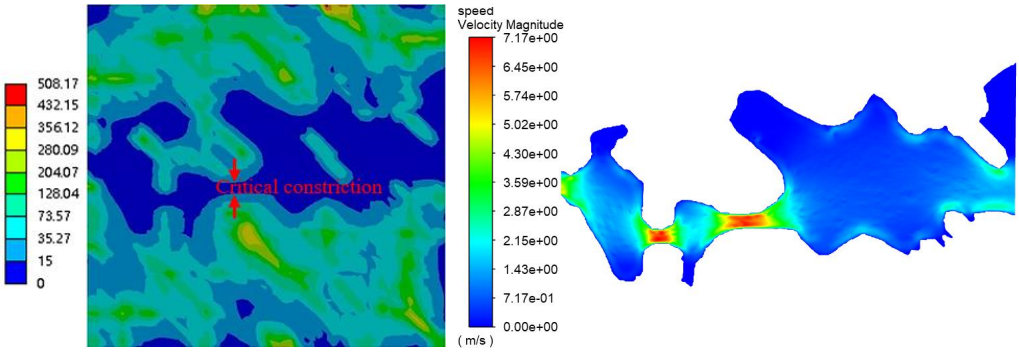


(a) 2D circular particle (b) 2D triangular particle

(c) 3D circular particle (d) 3D triangular particle

Figure 7: 2D and 3D Von Mises stress under different abrasive particles

Through the asperity pressure on the sealing surface and the Archard wear model, we can calculate the wear amount of each asperity to update the rough surface topography. The pressure distribution of the worn surface can be obtained using the worn topography for compression simulation. **Figure 8** shows that the original contact area has been worn to generate a blue percolation channel, the narrowest part of which is the critical constriction. By extracting the leakage channel and conducting a Computational Fluid Dynamics (CFD) simulation analysis of differential pressure flow, the flow rate at the critical constriction can be determined, which is very important for calculating the leakage rate of the sealing surface.

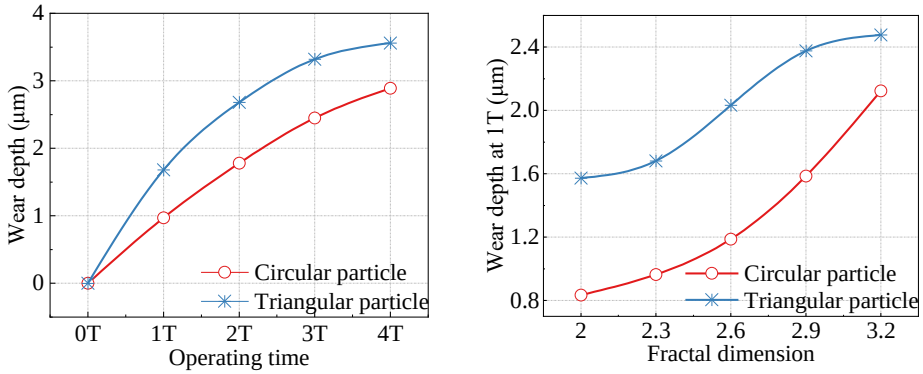


(a) Von Mises stress on sealing surface (b) CFD analysis of percolation channel

Figure 8: Numerical analysis of sealing surface and percolation channel

Operating time and fractal dimensions are two crucial parameters for surface wear. The former reflects the working conditions, and the latter reflects the surface topography characteristics. **Figure**

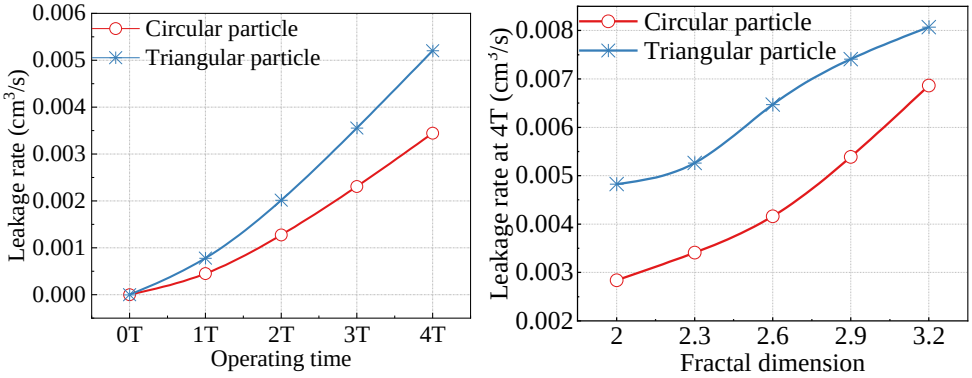
9 shows the average wear depth of the asperity on the rough surface of the wear ring with the operating time $T = 300s$ (corresponding to 75 strokes) and the variation of fractal dimensions under two wear particle conditions. Due to the cutting effect of triangular particles, the wear rate in the total running time is more significant than that of circle particles. With the increase of the fractal dimensions, the surface becomes progressively rougher. However, the impact on the wear induced by triangular particles remains immaterial under larger degrees of roughness. For circular wear particles, the increase in roughness height will significantly increase the hindering effect on particle movement, so the wear rate of circular particles significantly increases under large fractal dimensions.



(a) Time (b) Fractal feature

Figure 9: Reciprocating seal wear depth at different serve times and fractal feature

The flow rate at the critical constriction of the percolation channel can be obtained by the wear amount and the topography evolution law, and the oil film thickness can be obtained by the numerical calculation method. **Figure 10** reflects the mapping relationship of leakage rate with operating time and fractal characteristics. Due to the large wear amount of triangular particles, the leakage rate is also more significant. It is worth noting that in the leakage rate change curve under varying operating times, the leakage rate exhibits minimal alterations at the initial stages of wear because the asperity can still undergo axial deformation under normal load at a certain height, and it is connected with other adjacent asperities to form continuous contact clusters to prevent oil leakage. As the wear process occurs, the asperity height decreases continuously, resulting in the separation in the contact area; the leakage increases rapidly in the late wear period. The influence trend of fractal dimensions on the leakage rate is consistent with the wear depth. Increasing the surface roughness harms the sealing performance, and the leakage rate increases significantly.



(a) Time (b) Fractal feature

Figure 10: Reciprocating seal leakage rate at different serve time and fractal feature

4. CONCLUSION AND FUTURE WORK

This study introduces a numerical approach to estimating the leakage rate of reciprocating seals experiencing abrasive wear under mixed lubrication. The circular and triangular particles are selected as research objects to obtain the Von Mises stress of the sealing surface under normal compression and with the presence of wear particles. The Archard wear model calculates the wear amount of each asperity in the abrasive wear state, and the surface topography is updated until the percolation channel through the sealing surface appears. By integrating percolation theory with CFD analysis, the flow rate and oil film thickness at the critical constriction under the influence of differential pressure flow can be obtained to calculate the leakage rate under various wear conditions.

The simulation results reveal that the triangular particle mainly plays a cutting effect, resulting in a larger wear volume than the circular particle. Moreover, the wear rate of circular wear particles increases significantly with greater surface roughness. The evolution of the leakage rate aligns with the changes in wear volume, and a running-in stage appears in the early operational phase when the difference in leakage volume is relatively indiscernible. As the wear depth intensifies, the leakage rate experiences a substantial increase.

In future work, we aim to conduct wear tests on PTFE specimens using abrasive particles of different shapes in a wide temperature range to explore the microscopic wear mechanism of the sealing surface. The topography evolution of the seal after wear can be verified by comparing the test and simulation results. The percolation theory can be adopted to establish the mapping relationship between the seal's performance degradation and topography evolution. In addition, we are currently developing an adjustable end-seal test bench, which will enable us to conduct percolation tests on worn specimen surfaces at lower sealed pressures. We expect this method to enable us to understand the leakage mechanism, predict faults, and propose optimal design schemes for sealing surfaces.

ACKNOWLEDGEMENTS

This research was supported by the China Scholarship Council (CSC) [Grant No. 202306260259].

NOMENCLATURE

A_0	Nominal contact area	n	Frequency base number
$A(\zeta)$	Contact area at magnification ζ	p_c	Asperity pressure
$C(q)$	Power spectral density function	q	Wavevector
D	Fractal dimension	$\dot{Q}$	Leakage rate
E	Elastic modulus	R	Curvature radius
F_N	Nominal load	s	Sliding distance
G	Scale parameter	u_1	Separation height at the critical constriction
h	Separation height	W_a	Wear volume
H_m	Material hardness	Δp	Pressure difference
k_a	Non-dimensional wear coefficient	σ_a	Asperity height
K_a	Wear coefficient	γ^n	Spatial frequency
K_h	Hardness coefficient	ω	Elastic deformation of an asperity
l	Asperity wavelength	ζ	Magnification factor
L	Sample length	ν	Poisson ratio
M	Number of superimposed wave layers	η	Oil viscosity

REFERENCES

[1] Nikas, G. K. (2010). Eighty years of research on hydraulic reciprocating seals: review of tribological studies and related topics since the 1930s. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 224(1), 1-23.

[2] Salant, R. F., Maser, N., & Yang, B. (2007). Numerical model of a reciprocating hydraulic rod seal.

[3] Lorenz, B. (2012). Contact mechanics and friction of elastic solids on hard and rough substrates (Vol. 37). Forschungszentrum Jülich.

[4] Khruschov, M. M. (1974). Principles of abrasive wear. Wear, 28(1), 69-88.

[5] Wang, J., Li, J., & Ma, C. (2023). A performance degradation analysis method for a reciprocating rod seal in the wear process under mixed lubrication conditions. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 237(3), 681-697.

[6] Peng, C., Fischer, F. J., Schmitz, K., & Murrenhoff, H. (2021). Comparative analysis of leakage calculations for metallic seals of ball-seat valves using the multi-asperity model and the magnification-based model. Tribology International, 163, 107130.

[7] Zheng, W., Sun, J., Ma, C., & Yu, Q. (2022). Percolation interpretation of film pressure forming mechanism of mechanical seal and calculation method of film pressure coefficient. Tribology International, 173, 107664.

[8] Majumdar, A., & Bhushan, B. (1991). Fractal model of elastic-plastic contact between rough surfaces.

[9] Rabinowicz, E. D. L. R. P., Dunn, L. A., & Russell, P. G. (1961). A study of abrasive wear under three-body conditions. wear, 4(5), 345-355.

[10] Koottathape, N., Takahashi, H., Iwasaki, N., Kanehira, M., & Finger, W. J. (2012). Two-and three-body wear of composite resins. Dental Materials, 28(12), 1261-1270.

[11] Liu, D., Wang, S., & Zhang, C. (2018). A multiscale wear simulation method for rotary lip seal under mixed lubricating conditions. Tribology International, 121, 190-203.

[12] Yan, W., & Komvopoulos, K. (1998). Contact analysis of elastic-plastic fractal surfaces. Journal of applied physics, 84(7), 3617-3624.

[13] Li, J. S., Tang, Y., Li, Z. T., Ding, X. R., & Li, Z. (2017). Study on the optical performance of thin-film light-emitting diodes using fractal micro-roughness surface model. Applied Surface Science, 410, 60-69.

[14] Fischer-Cripps, A. C. (1999). The Hertzian contact surface. Journal of materials science, 34(1), 129-137.

[15] Liu, Y., Liskiewicz, T. W., & Beake, B. D. (2019). Dynamic changes of mechanical properties induced by friction in the Archard wear model. Wear, 428, 366-375.

[16] Yuan, Y., Gan, L., Liu, K., & Yang, X. (2017). Elastoplastic contact mechanics model of rough surface based on fractal theory. Chinese Journal of Mechanical Engineering, 30(1), 207-215.

[17] Persson, B. N. J. (2007). Relation between interfacial separation and load: a general theory of contact mechanics. Physical review letters, 99(12), 125502.

[18] Yang, C., & Persson, B. N. J. (2008). Contact mechanics: contact area and interfacial separation from small contact to full contact. Journal of Physics: Condensed Matter, 20(21), 215214.

[19] Persson, B. N. J., & Yang, C. (2008). Theory of the leak-rate of seals. Journal of Physics: Condensed Matter, 20(31), 315011.

TRIBOLOGICAL PROPERTIES OF DIFFERENT SLIPPER DESIGNS OF AN AXIAL PISTON PUMP

peer reviewed

Svenja Horn[1]*, Roman Ivantysyn[1], Torsten Schmidt[2], Jürgen Weber[1], Jörg Schneider[2]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*
[2] *Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik, Reichenhainer Straße 88, 09126 Chemnitz*

* Corresponding author: Tel.: +49 351 463-33701; E-mail address: svenja.horn@tu-dresden.de

ABSTRACT

New application areas such as compact drives and displacement control using variable electric motors require that hydrostatic machines exhibit good performance also at very low relative speeds without being damaged during critical mixed friction conditions. While some pumps currently address this issue using leaded metals in their sliding bearings, European regulations (Reach Regulation) mandate the transition to lead-free materials in the near future. To address these challenges, this paper investigates about surface structures that offer the ability to selectively modify the pressure field within the fluid gap, generating additional hydrodynamic pressure. In this research project, various slipper surface structures of an axial piston pump were developed through extensive simulation studies and tested on a pump test-rig as well as on a dedicated hydrostatic tribometer.

This paper outlines the development, construction, and optimization of a novel hydrostatic tribometer capable of adjusting speed, pressure, acceleration and temperature. Multiple measurements were conducted with different slipper geometries. The results encompass torque, temperature and wear measurements. They are presented along with a comprehensive analysis and validation of the findings in comparison to simulation and pump test results. In this study the tribometer revealed specific constraints not detected in pump bench tests, underscoring the imperative nature of such examinations.

Keywords: Tribometer, Slippers, Surface, Tribology, Wear

1. INTRODUCTION

The increasing electrification of hydraulic systems poses new operational speed challenges for hydrostatic pumps. Specifically, at lower speeds, these pumps need to sustain loads without succumbing to mixed friction and wear. Presently, some pumps can handle these conditions, largely due to their use of leaded metals in their sliding bearings. These metals are proficient at withstanding critical contact scenarios. However, the shift towards lead-free materials is imminent to align with European standards (Reach Regulation). Previous research has shown that surface structures can improve the tribological characteristics of hydrostatic bearings [1]. Meso-structures refer to surface patterns with slight height variations (in the µm range) spread over broader areas (in mm). Such structuring can selectively alter the fluid gap's pressure field, leading to an enhanced performance.

In previous publications, [2], [3] and [4] such structures were designed and tested for a slipper of an axial piston pump. The optimization method and pump measurements of a wave and a step surface structure are explained in detail in these publications. Especially the step structure achieved a notable enhancement in the overall pump efficiency when compared with a non-structured flat slipper. A

remarkable feat was the reduction of the pump's start-up torque by an astonishing 80% by integrating the stepped slipper design. This pump underwent rigorous testing on a test bench under challenging scenarios, from start-stop sequences and extremely low speeds (both loaded and unloaded), to cold starts and a demanding endurance test under an elevated pressure. Impressively, the slippers weathered these tests with negligible wear of less than one micrometer. Nevertheless, to truly gauge the potential constraints of different slipper designs, a more streamlined component test is essential.

In conventional pump measurements on a laboratory system test bench, it is very time consuming to remove the slippers for wear measurements or to completely change them in order to compare the performance of different designs and materials. Furthermore, it is impossible to investigate only the slipper/swashplate tribo-contact without including the losses of the other sealing gaps. Tribometer evaluations stand out for their ability to test the sliding and frictional attributes of various designs in isolation. The aim of the presented work is to investigate on different slipper designs and materials and their performance at different testing conditions, bypassing the lengthy procedures associated with pump test rigs. The difficulty in designing a suitable tribometer is to keep the set-up as simple as possible while still providing realistic results. Such as standard tribometers often renounce the hydrostatic pressure field within the slipper pocket, a pivotal mechanism that is essentially determining the performance of the lubricating interfaces [5]. This disadvantage will be overcome with the herein presented hydrostatic tribometer set-up.

Chapter 3 of this paper shows how a commercially available tribometer can be adapted with hydrostatics to explore the tribological properties of the slipper/swashplate interface. It describes the challenge to choose a good compromise between keeping the test-rig set-up simple while achieving useful results. The hydrodynamic performance of the different meso-structured slipper designs is investigated by torque measurements over a broad speed spectrum and several pressures in chapter 4. The advantages of the capabilities of such a tribometer, as compared to a full pump test rig, were highlighted by the effects of thermal deformation on the performance of the slipper. Those are described in chapter 5, also including ANSYS® simulations to reveal the underlying physical mechanisms.

2. STATE OF THE ART

For the reasons mentioned above, the focus in the next section will be on functional slipper test rigs that allow for easy slipper replacement. Those test rigs can be categorized into four distinct types, as illustrated in **Figure 1**, with complexity escalating from left to right. The initial type (a) features both a stationary slipper and swash plate, a setup employed by Manring to measure the pressure beneath the slipper sealing, as referenced in [6]. Such a configuration facilitates the integration of a wide array of sensors, eliminating the need for telemetry systems or load-resistant bearings, but at the cost of overlooking component movement. Conversely, the second category of tribometer (b) introduces a rotating disc accompanied by one or more hydrostatically supported slippers. This approach is selected for the herein developed tribometer, primarily because it subjects the bearings only to axial stresses, thereby simplifying the rotational assembly. The first empirical investigations on such a test rig was done by Renius in 1973 [7]. These were the first isolated slipper efficiency measurement tests, accomplished by recording both the frictional forces and leakages for a single slipper. It is noteworthy that this set up also included pulsating high pressure using a valve. Next to efficiency the publications also featured the Stribeck curve for slippers for different balancing ratios. Many other researchers used the type (b) set up, as it recreates a quiet realistic test environment to evaluate slipper performance mainly by measuring friction and leakage [8–11]. The stationary slippers allow for interesting sensor placement for example temperature [12, 13], pressure [10] or displacement sensors to measure the micro movement and gap height [14–16].

A few studies have also been performed on a wobble plate type set up, which is the closest set up one can chose before rotating the slippers. Notably Iboshi used such a set up and still measured the displacement of the slippers, however while prohibiting slipper spin [17]. That the slipper spin is important and existing was shown by Bergada [10] and then later measured in a type (d) test rig by Chao [18].

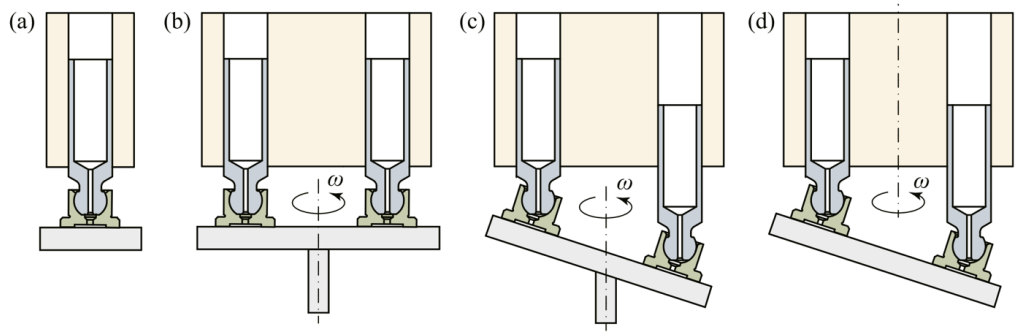


Figure 1: Types of slipper test rigs [19]

This section presents the most relevant research on functional slipper test-rigs. For a complete review of slipper test rigs published in the last 50 years, the work of Chao can be recommended [20]. A previously uncharted territory, however, is the goal to retrofit a commercially available tribometer with hydrostatic capabilities, all the while retaining the original rotating components, block, piston, and slipper, in their unmodified form. While this methodology seems economically favorable for those already equipped with a tribometer, it's not devoid of challenges, as explained in this paper.

3. DEVELOPMENT OF A HYDROSTATIC TRIBOMETER

3.1. Adaption of an existing rotary tribometer to hydrostatics

A concept has been developed to accommodate the original pump components of the swash plate/slipper friction pairing for hydrostatic testing in a hydrostatic tribometer environment. This enables an isolated friction measurement to be conducted, without any influence from other parts of the power unit. The specifications for slipper testing have been set to a maximum load of 300 bar and a rotating speed of up to 2300 rpm. The implementation is based on a TRM 5000 rotary tribometer from Wazau, shown in the photo in **Figure 2**. A new spindle has been added to the tribometer's basic structure to accommodate axial loads of up to 12 kN, in order to realize the test specifications. This axial load limits the number of pistons to one, a compromise that was accepted in order to reach high pressures. However, this in turn means that there is a very high asymmetric load impact on the tribometer frame, producing a tilting of the system. The system has been blocked and stabilized as good as possible reducing the tilting from over 150 µm to 20-40 µm in x and y-direction. To measure force and torque a Kistler Type 9265A 3-component dynamometer was used in combination with a multichannel Kistler charge amplifier Type 5080A1088004. Operating pressures up to 300 bar are provided by a separately controllable "Cytrobox" hydraulic unit from Bosch Rexroth. The original swashplate is substituted by a plate with rotational symmetry made of identical material, possessing identical hardness characteristics and sharing the same finish as the original. While the block and the slippers rotate in an axial piston pump, the presented tribometer setup includes a rotating swashplate. This setup offers the advantage of solely having the slipper / swashplate pairing's sliding interface present, without any influences from the valve plate or the piston sealings. The slipper is employable in its original form in the tribometer.

Figure 2 illustrates the assembly of the components and the adaptation of the pressure supply line. The role of the block holder is to connect the block and hydraulic hose, and channel hydraulic oil from the hose into the cylinder chamber of the block where the piston is located. The hose is connected using a quick release coupling. To seal the interface between the block holder and the block, a copper seal with an aperture for filling a cylinder chamber is used. The seal is secured tightly with eight M6 screws and one M14 screw.

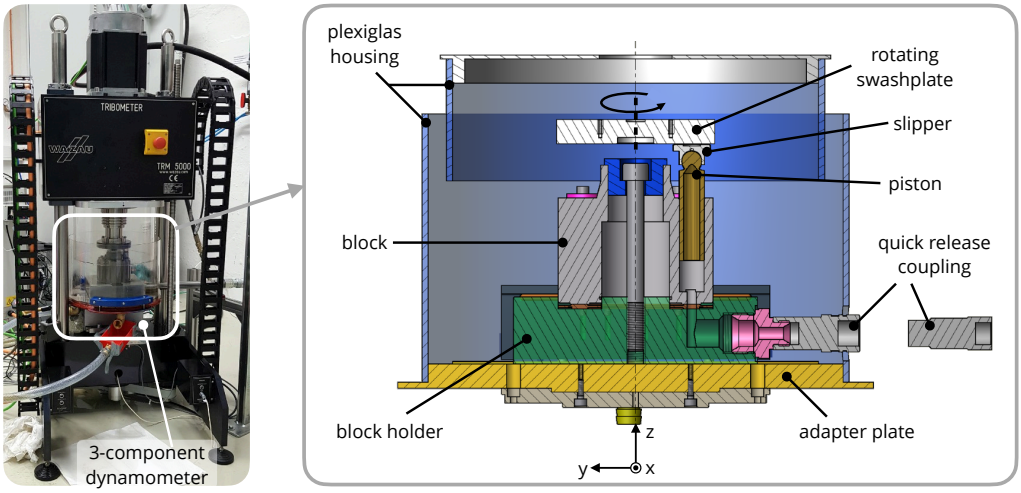


Figure 2: Tribometer set-up (left) and components of the new hydrostatic adaption (right)

The optimal hose routing was established through preliminary tests to minimize the impact of hydraulic hose supply lines on measured torque. The weight, position, and bending radius of the hose can affect the measuring system. The most effective hose guidance was determined in static tests without rotating the swashplate. To prevent oil leakage, a Plexiglas housing has been installed. Leakage oil can be drained from two oil drain holes and pumped back to the CytroBox oil tank through a 10 µm filter using an oil level-controlled pump. Additionally, the leakage oil drain is separated from the return lines, preventing any torque influence. To avoid a sudden impact of the slipper surface on the swashplate surface during pressure build-up, a compression spring has been inserted into the block bore. This indicates that the surfaces of the test pair are in contact with a very low preload of approximately 50 N before the pressure builds up and impact is avoided. The tribometer is PC-controlled using the "TriboControl V11" software. The CytroBox is controlled using "IndraWorks DS". The measurement of the forces and moments using the 3-component dynamometer are recorded with "imc Studio".

3.2. Measurement set-up

One of the greatest difficulties of the measurement setup is to measure very low friction values in the x-y plane as accurately as possible, despite the extremely high forces due to the piston pressures in the z-plane in the range of several kN. Ideally, the torque should be measured directly at the disc, as close as possible to the friction point. However, this in turn means that there must be a suitable bearing of the high axial forces that does not cause any additional frictional forces. Since a corresponding set-up was not available, a force measurement platform from Kistler was used, which is mounted under the block (see **Figure 2**). The high-precision quartz force sensors installed in it can be adjusted in their measuring range between 0-1.5 kN or 0-3 kN, whereby the smaller measuring range is selected for the x-direction of the measured force that is relevant for the torque calculation.

After numerous optimization loops, a pressure-dependent offset of the forces to be measured remains,

which can be traced back to the tilting of the tribometer frame mentioned in the previous paragraph and a sensor-specific crosstalk. To account for this, an initial right/left-wise rotation combination at 100 or 500 rpm is carried out at the beginning of each pressure level. The mean value of the respective moments is then used as the offset. In order to achieve a similar lubrication and temperature behaviour to that of the pump, an oil and air cooling system is installed as shown in **Figure 3** (left). It sprays oil in room temperature onto the swash plate and slipper and can also be switched off to overheat the slipper. A pyrometer measures the temperature through a hole in the plexiglass (**Figure 3** middle) on one spot on the disc (**Figure 3** right) before or after the slipper passes that position, depending on the rotating direction.

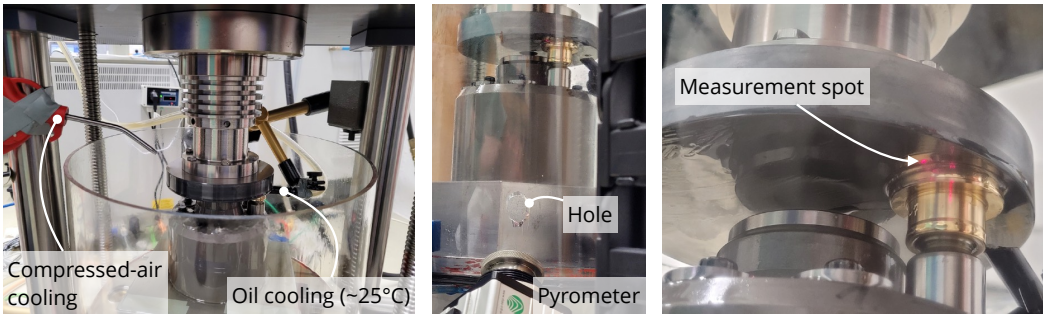


Figure 3: Set-up of the compressed-air and oil cooling of the slipper / disc interface

3.3. Investigated slipper designs

Figure 4 shows the slippers studied in this investigation. They were optimized and developed through simulation, explained in detail in [3]. All designs have the same outer radius of 29.6 mm but different inner radii and thus a different hydrostatic lift capacity. The reference design (left) is a flat, single land slipper design with an optimized balance factor of 99 %. The balance factor is an indicator of the slipper's pocket size and is given in (1). It is calculated by dividing the hydrostatic force F_{fz}, that is estimated as logarithmic pressure drop over the sealing land, by the force of the piston on the slipper F_{AK}, multiplied by the cosine of the swashplate angle β:

$$B = \frac{F_{fz}}{F_{AK}} \cdot cos\beta \tag{1}$$



Figure 4: investigated slipper designs and their optimized micro structure

The wave design (middle) contains a sine-wave micro structure on the sealing land with an amplitude of ±3 µm and a balance of 102.3 %. The paper presents results on wave design only for pump

measurements. Tribometer results with wave slipper can be found in [21]. The design of the step slipper (right) features a step down of 6 µm towards the external radius, located approximately midway along the sealing land and has the highest balance value of 105 %. The location and magnitude of this step was carefully engineered using advanced numerical simulations as described in [3]. For the manufacturing of the slippers different production processes have been investigated and optimized, documented in [21], whereas a high accuracy drilling has been finally adapted.

Additionally, a slipper with multiple lands has been investigated as it is the current state of the art in industry. Multi-land slippers, also known as slippers with several lands separated by millimeter-deep grooves, provide hydrodynamic lift and stabilize the slipper. The middle land defines the hydrostatic pressure field [22].

3.4. Differences between the hydrostatic tribometer and the pump system

Due to the high complexity of the sealing gaps in the pump, it is not possible to directly compare the measurement results of the tribometer and the pump. Each test scenario has its own advantages and disadvantages and can be used for specific investigations. The following section explains the main differences between the new hydrostatic tribometer set-up and a conventional axial piston pump bench test.

- The tribometer setup uses a rotating swashplate rather than a rotating block, which also requires a second seal on the valveplate side. In this way it is possible to measure only the slipper seal interface. A disadvantage is the lack of centrifugal force that prevents the slippers from tilting. Especially for high speeds, the hydrodynamic lift should be changed compared to the pump system.
- The current tribometer set-up does not include a change between high and low pressure, which can cause the slipper to lift off or tilt due to the sudden change in load. This effect, also known as "breathing", allows the slipper to lift off and reset its tilt, which could also play an important role with regard to the thermal load on the surfaces. For future work an additional control valve could simulate the change in pressure.
- In the pump, the case is filled with oil hence forcing the slippers to run through oil. On the tribometer, an air and oil cooling were installed in order to adapt the lubrication behavior and temperature in the gap between slipper and swashplate. This simplified design allows the temperature to be measured on the swashplate using a pyrometer and the use of a Plexiglass housing to observe the experiment. This set-up allows for an easy replacement of the slipper as well as the usage of different oil types. However, the movement through the oil, as in the pump, could influence the inclination of the slipper and have an enormous cooling effect, which has been underestimated (see chapter 5).
- The swashplate angle on the tribometer is 0° instead of 17° as in the pump for full stroke at 100%. Since the piston does not move axially in the bore, the slipper can be inspected isolated. However, the force balance in formula (1) is changed to a higher value, leading to the assumption that higher gaps occur and thus only a highly reduced friction can be measured between slipper and swashplate. To account for this effect, a “tribometer slipper design” has been developed for each meso-structure through simulation wherein the slipper’s inner diameter has been adjusted to reduce the lift force and balance value. As the tribometer design unexpectedly was subject to excessive wear, the same design as in the pump with a higher balance was selected for most the measurements herein.

4. COMPARISON OF THE SLIPPER DESIGNS USING TORQUE MEASUREMENTS

This section compares the torque losses of the different slipper designs first from measurements of the tribometer and second from results of the full pump test-rig. On the one hand, critical areas are to

be identified and on the other hand, the proportion of the total hydromechanical losses accounted for by the slippers is to be analyzed.

4.1. Tribometer torque measurement results

The aim of the torque measurements was to record the Stribeck curve for the entire speed range from 0-2300rpm for the various slipper designs, being cooled by the external oil and air cooling. This was done by gradually increasing the speed and measuring the torque at steady state. The left plot of **Figure 5** provides the speed and force curves. The left y-axis gives the measured force sensor results in x-direction while the right y-axis shows the speed values, plotted over the samples. At the start and end a +/-500rpm combination for offset calculation has been conducted. The first two curves accelerate from 0 to 2300 rpm with small speed steps in the lower range while the last curve accelerates directly from 0 to 2300 rpm. For each design 0, 50 and 100 bar have been measured twice.

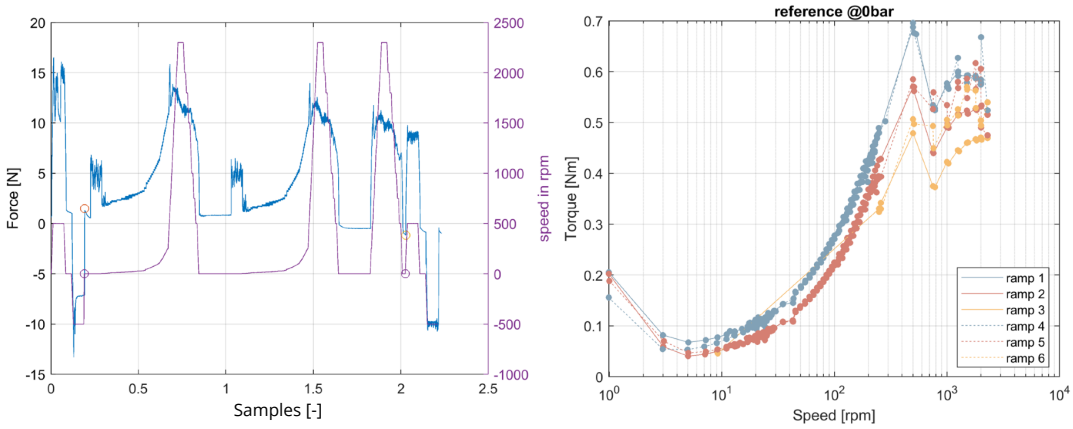


Figure 5: Reference slipper force (left) and torque measurements (right) at 0 bar

As an example for the reference design at 0 bar, the right graph in **Figure 5** shows the derived torque values, based on the measured force for the extracted speed ramps in logarithmic representation. All six speed ramps are plotted for the two trials, the second trial with dashed lines. It is visible that the curves are similar for each ramp especially for speeds below 600 rpm, confirming repeatability of the measurements even at these low forces. As the temperature rises slightly for each ramp the torque losses decrease for this unloaded scenario. This measurement procedure was repeated with the multi-land and step slipper.

Figure 6 shows the torque losses at 0 bar, 50 bar and 100 bar over the speed that is displayed logarithmic and reaching from 0.1 rpm to 2300 rpm for the three different slipper designs reference, multi-land and step. To avoid dynamic effects, the speeds have been measured at steady state and connected in the graph only for reason of clarity. Photos of the surface of each design after testing confirm the results of the torque measurements. For 0 bar, all curves adhere to the Stribeck curve in a qualitative manner. The design with the least friction losses is the step-design and demonstrates significantly better values than alternative designs across all speeds. The torque loss of all designs is at minimum between 5 and 20 rpm and then rise with different slopes. A plateau is reached at approximately 500 rpm, with no notable increase in torque observed for higher speeds for the multi-land and the reference design. As there is no such plateau in the pump measurements in **Figure 7**, the slippers probably are the main contributor to the friction at low pressures, while the other sliding pairings, such as the piston/bushing or block/valve plate, become more significant power loss sources at higher pressures. This is supported by the fact that the replacement of the slipper designs in the

pump significantly changed the low speed performance, while having little effect on the high speed readings, as can be seen in more detail in [4].

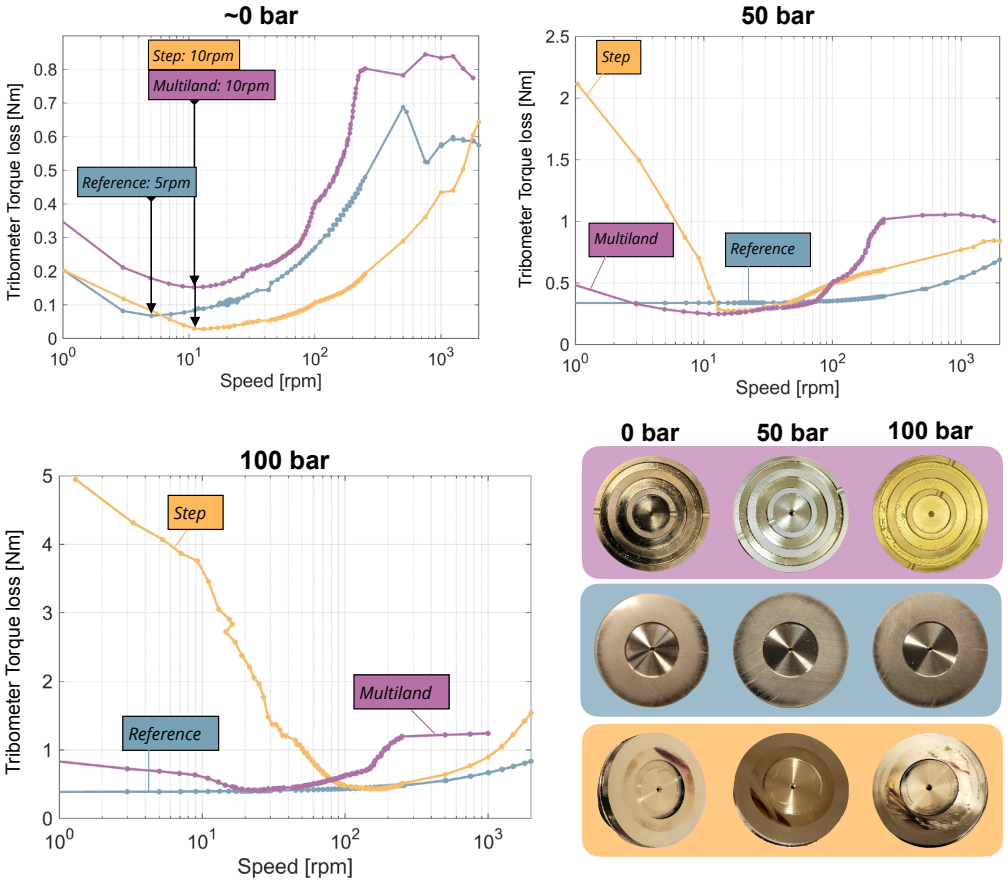


Figure 6: Comparison of tribometer torque loss measurements for 0 bar, 50 bar and 100 bar and a qualitative overview of the wear patterns

Once hydrostatic pressure is built up as for the 50 and 100 bar measurements, the pressure-dependent offset determination is not always reliable. Therefore, the absolute values may vary due to the deformations of the test bench as explained in paragraph 3.4. The qualitative curves of the three designs are very similar for 50 and 100 bar. Quantitatively, the step slipper has twice the friction losses at 100 than at 50 bar, while the reference and the multi-land slipper remain in the same range. The curve of the reference slipper is remarkably flat. At extremely low speeds in the mixed friction range, there is no inclination observed. As anticipated, the slippers are more hydrostatically balanced and "swim" due to the 0° swashplate angle and hence lower piston forces during the tribometer experiment, which leads to lower friction values, especially for low speeds. It is not clear, why the step slipper is not exhibiting the same trend in the low speed region, where it experiences high torques and thus abrasion, even though it has the largest pocket size. One possible explanation is that the step structure is damaged or deformed during the high speed portion at 0 bar, which then has a negative effect on the performance at low speed for 50 and 100 bar. The thermal deformation could play a significant role. The step slipper was shown to have very low leakage and small gap heights, if the temperature is significantly higher at high speed. Consequently, the slipper could also be affected at

low speed, as these measurements are chronologically after the high speed section. That there is serious wear can be seen in the wear patterns of the surface photos. A dark stripe is visible on the inner ring of the surface after 0 bar measurements and smearing of the slipper material after 100 bar measurements.

4.2. Torque measurements of the full pump test bench

Figure 7 shows the torque measurements of the full pump measurement at 0 (left) and 50 bar (right) in a similar presentation to the tribometer Stribeck curves, in order to categorise the torque ranges in which the pump behaves in comparison to the tribometer. The multi-land slipper was not measured in the pump at all speed ranges but the wave slipper and the lead-free step slipper were. While the pump measurements include the torque losses of the entire rotating kit, including the slipper, piston, block and shaft bearings, on the tribometer only one slipper is measured individually. To give a reasonable comparison between the two measurements, the torque losses of the pump were divided by the number of rotating slippers (nine) in order to assess how much just one slipper contributes.

As has been stated earlier, for this pump the slippers seem to be the main contributor of friction at low speed: At 0 bar, the friction losses in the lower speed ranges between 1 - 200 rpm are most likely dominated by the friction losses of the slippers, both ranging from 0 to 0.6 Nm. For higher speeds above 100 rpm, the friction values in the pump increase significantly more for all designs, which could be caused by the losses of the pistons and the valve plate. This trend is even more pronounced at 50 and 100 bar and is in agreement with the simulated analyses in [23], using the same pump. At 50 bar the torque values range from 0.3 Nm to 2.8 Nm exhibited a similar range for both the pump and the tribometer. The qualitative progression of the step curve is quite similar for both tests with a maximum value of just over 2 Nm in the mixed friction area and a minimum friction value of around 0.3 Nm at 10-50 rpm.

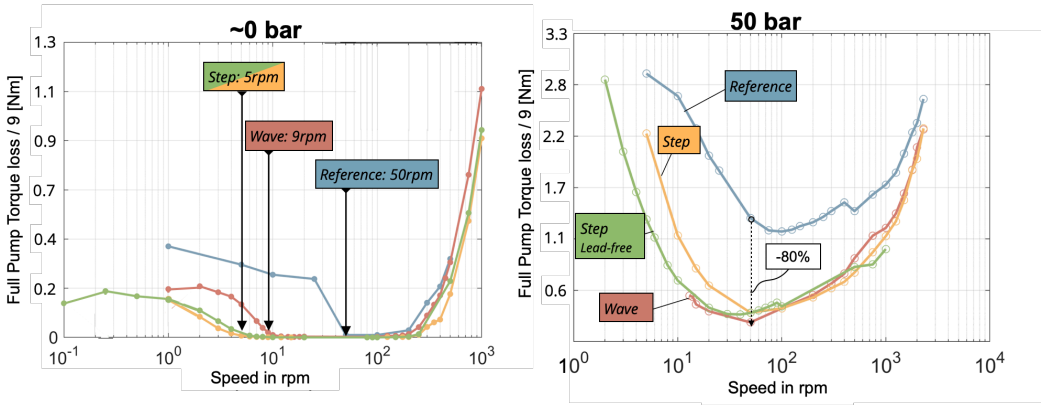


Figure 7: Full pump torque loss divided by 9 for 0 bar (left) and 50 bar (right) over logarithmic speed

While it is clear that the pump and tribometer measurements cannot be directly compared due to higher number of rotating parts in the pump, the overall trends are helpful for an evaluation of the tribometer measurements.

5. EFFECT OF TEMPERATURE ON DIFFERENT SLIPPER DESIGNS

The previous section showed that on the tribometer, the temperature can have a significant effect on slipper performance and wear. Since the pump measurements did not reveal any such trends, the following chapter provides temperature studies from the tribometer and the pump. Paragraph 5.1

presents an overheat test on the tribometer, followed by pump measurements with an increased inlet temperature. Subsequently, the thermal deformation of the slipper is simulated in ANSYS® and the resulting findings are discussed.

5.1. Overheat test using the tribometer set-up

In the following section, the temperature behavior of the step slipper is examined in comparison to the multi-land slipper. The slipper is accelerated three times from 0 to 2300 rpm at 200 bar, in order to introduce a lot of energy into the system. Each maximum speed is held for 5 minutes. A pause of 15 minutes between the ramps was chosen to allow for a cooling-down effect before again introducing energy to the system. **Figure 8** shows the comparison between the step slipper trials with (pink) and without (blue) cooling. While on the left y-axis the temperature is plotted, on the right y-axis the speed is given (orange). For both cases, it is to be expected that the temperature rises during acceleration and the constant speed plateau at 2300 rpm, as these conditions exhibit a great amount of energy into the sliding pair.

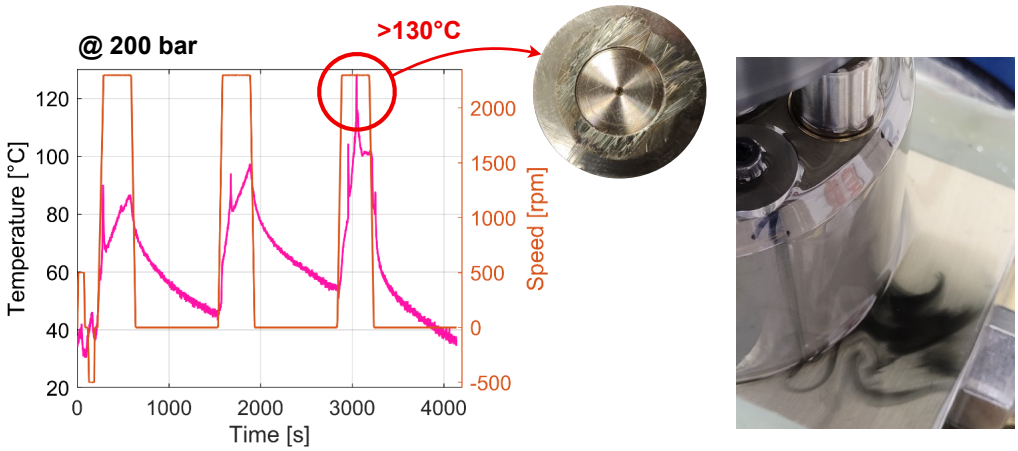


Figure 8: Slipper surface reaches extremely high temperatures and experiences severe wear

The temperatures of the cooled slippers reach less than half the value of the uncooled slipper with maximum values of 65°C. The uncooled experiments reveal distinct needle-like peaks that can be related to severe wear. During the last speed plateau, temperatures exceed 130°C. It should also be noted that these high temperatures could affect low speed performance, as already mentioned above. The photos of the surfaces of the slippers after the measurements are shown in the center of Figure 8. The wear pattern reveals, that there is not only particle abrasion (3rd body) but also a smearing effect of the material due to the high temperatures during contact. This mixed friction occurred after the surface temperature of the disc exceeded 100˚C, indicating that the oil performance was degraded to a point where insufficient load holding took place. This oil degradation is evident by the pitch-black oil leaking out of the slipper gap, as shown in the photo on the right side in Figure 8.

While the slipper with the step and waved design experienced excessive wear, the multi-land slipper, which is directly taken from the standard series pump, experienced little to no wear in the same test set up. This design is characterized by multiple lands, adding one inner and one outer stabilizing land to the center sealing land. It is widely used as industrial standard and described in more detail f.e. in [24]. The results can be seen in the next figure. Additional to the temperature of the multi-land design without cooling, **Figure 9** also shows the step slipper with and without cooling. It is very interesting to see, that the un-cooled multi-land slipper exhibits a similar behavior to the cooled step slipper.

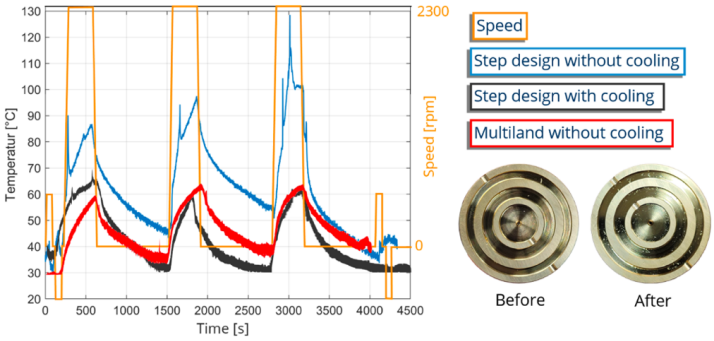


Figure 9: Temperature development for step (with and without cooling) vs. multi-land slipper (without cooling)

5.2. Pump measurements at increased inlet temperature

To investigate the influence of higher temperatures on the pump set-up, steady state measurements at increased inlet temperature as well as an elevated temperature endurance test has been carried out. The former is presented in **Figure 10**, showing the influence of the temperature on the pump losses. The left plot shows the measured torque losses for four different operating conditions at different pressures and speeds over the rising inlet temperature from 35 to 55 °C. All curves show a global reduction in friction losses. The right plot provides the detailed efficiencies of 500 rpm and 350 bar (low speed, high pressure). The hydromechanics efficiency is staying relatively constant while the volumetric efficiency is dropping significantly because of the lower viscosity at higher temperatures.

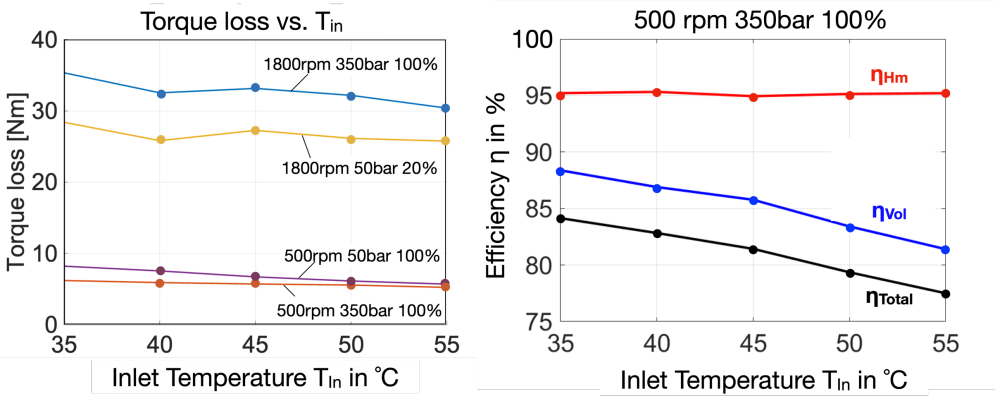


Figure 10: Influence of temperature on the pump losses. Left: Measured torque losses for 4 different operating conditions. Right: Measurement vs. simulation for low-speed high pressure.

Furthermore, an endurance test lasting 565 hours can be provided. Different materials and designs are subjected to a cyclic elevated pressure test in a pump at nominal speed. The combination of increased pressure of over 350 bar and cyclical speed changes results in an estimated temperature increase of over 70°C throughout the system. As shown in **Figure 11** (right), different materials and designs are used for the nine slippers. The profilometric diagrams in Figure 11 (left) allow a quantitative analysis. On average, the lead-free slipper achieves the lowest wear with a maximum of 1 µm on the inner radius. The diamond hardened brass slipper also performs well, but the inner edge wears more. The brass slipper shows the most wear, up to 1.5 µm on the inner radius. It is also interesting to note that the inner ring of the lead-free slipper runs in flat, whereas the brass slipper has

a run-in plane that slopes down towards the inner edge. This may suggest a variation in pressure deformation caused by differences in elasticity moduli, which will be further examined.

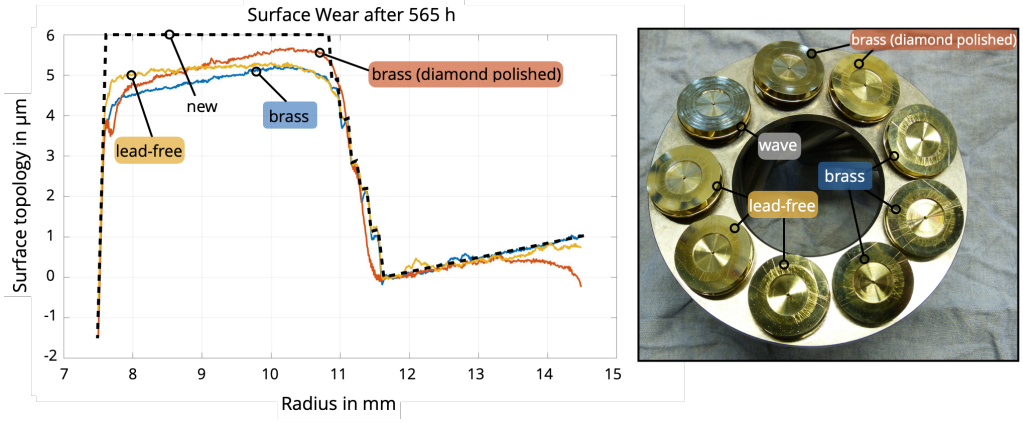


Figure 11: Surface topology of different materials of the step design after endurance testing over 565 hours (left), photo of slipper surfaces after endurance testing (right)

These trends differ to the tribometer results. Therefore, another reason apart from the maximum temperature must be responsible for the different behavior of the slippers. In the following, the temperature deformations for different maximum temperatures and temperature *gradients* in the slippers are analyzed by FEM simulations.

5.3. Simulation of the slipper deformation

To study the deformations on a basic level a simplified ANSYS® model was created that allows an independent investigation of the temperature and pressure deformation and their effect on slipper wear. As the pressure field of the step slipper is quite complex and only the deformation trends are of main interest, the reference design of the slipper was chosen to model using hexahedra elements of quadratic order, compare **Figure 12** (a). It is connected with a spherical joint to the piston that has a fixed support on the bottom as illustrated in Figure 12 (b). The boundary conditions to calculate the pressure and temperature deformation are given in Figure 12 (c).

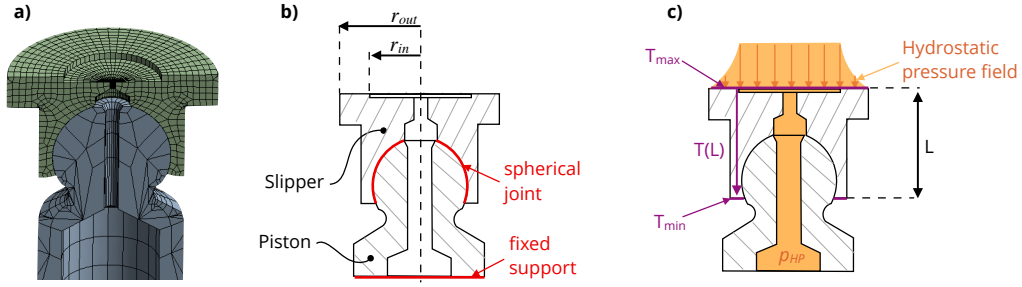


Figure 12: slipper and piston mesh (a), boundary conditions (b) and pressure and temperature loads (c)

For the reference slipper a logarithmic pressure drop can be assumed between the slipper and disc, which decreases over the sealing land from r_{in} to r_{out} up to the atmosphere pressure shown in (1) and derived in [25]. The piston as well as the orifice in the slipper and the slipper pocket are pressurized with p_{HP}.

$$p_{hydrostatic} = p_{HP} \frac{\ln\left(\frac{r}{r_{out}}\right)}{\ln\left(\frac{r_{in}}{r_{out}}\right)} \tag{2}$$

The heat causing slipper deformation originates in the gap between slipper and swashplate. Especially if the gap is very small, mixed friction arises and produces additional heat. Therefore, the maximum temperature is defined on the sealing surface of the slipper. The minimum temperature is assumed to be on the slipper's bottom and is estimated or derived from the simulation results of the pump simulation with Caspar FSTI. The heat transfer is described by the Fourier law and can be used to calculate the temperature distribution T(L) in the slipper.

The critical operating condition of the tribometer is simulated with 200 bar pressure and 130 °C on the slipper surface at a temperature of approx. 50 °C on the slipper bottom. **Figure 13** presents the simulation results with a magnification of the deformation.

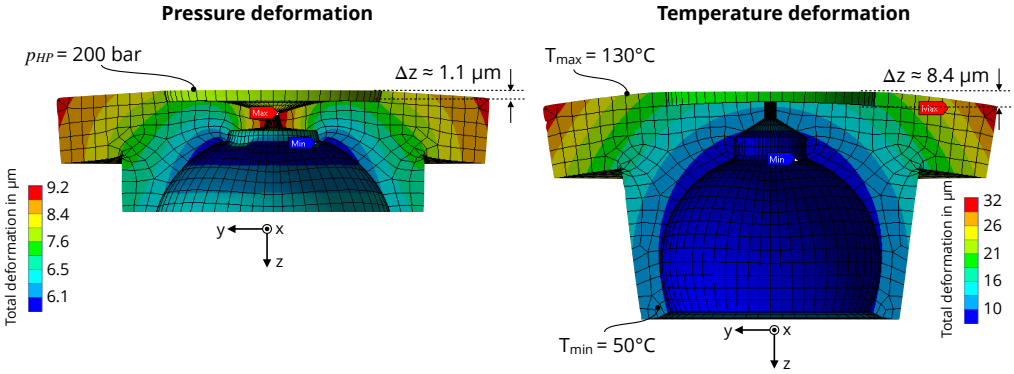


Figure 13: pressure (left) and temperature (right) deformation of the slipper

In the left side of Figure 13 simulation results of the pressure deformation for 200 bar are shown. The colors describe the total deformation of the slipper in all axis. The slipper surface deforms convex with a total radial change in z-direction of about 1.1 µm. Due to the contact conditions between piston and slipper the simulation results are not valid in the lower area of the slipper opening. In reality the edge of the slipper widens and the gap opens locally.

In the right side of Figure 13 the simulation results of the temperature deformation for 130 °C at the slipper surface are shown. The deformation is also of convex shape with a total radial change in z-direction of about 8.4 µm, which is significantly greater compared to the pressure influence. **Figure 14** shows the deformation in z-direction over ΔT (T_{max} -Tmin across the slipper) for different maximum temperatures 70, 100 and 130 °C. The surface deformation is greatly influenced by the temperature gradient, while the maximum surface temperature has no effect. This mathematical relationship can be derived from the linear thermal expansion equation, with ΔL as change in length, L as total length and α as thermal expansion coefficient.

$$\Delta L = \alpha \cdot \mathrm{L} \cdot \Delta \mathrm{T} \tag{3}$$

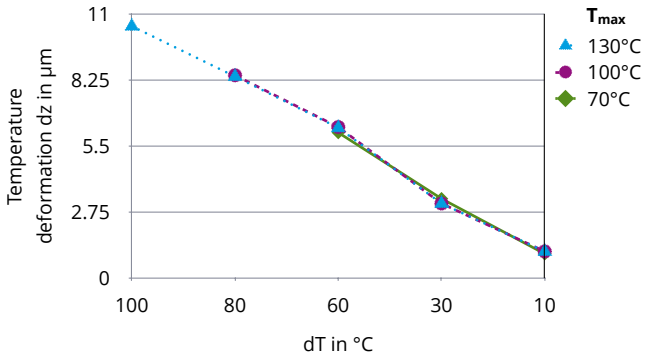


Figure 14: temperature deformation in z-direction over the temperature gradient for different maximum temperatures

5.4. Conclusion of the thermal behaviour of the step slipper

The experiments described were used to test the thermal limits of the step design, which appears to behave unfavourably at high temperature gradients, e.g. very high temperatures in the gap while surrounded by cooler oil and air. This is due to the convex temperature deformation of the slipper surface, which is quantified in the simulation section. The additional material of the step in the inner ring is counterproductive during such deformation.

This theory was confirmed by the tribometer and pump measurements. The differences of the thermal boundary conditions between the set-ups seem to be crucial. The tribometer set-up allows for different cooling arrangement. Without any cooling the slipper moves in an oil-free surrounding, providing little-to no convective cooling, as the slipper itself does not move. In the pump the slipper rotates through the case oil, which is typically warmer than the inlet and outlet temperature of the pump. Any heat generation in gaps will be dissipated into the case, while simultaneously preheating all parts to a homogenous temperature. In the tribometer the lack of case oil allows the individual parts to have significantly different temperatures, as the air surrounding the parts thermally insulates them from each other. In addition, the parts have a much higher thermal gradient, as the surround air is at room temperature always, while warm oil enters the slipper and increases significantly in the gap.

The multi-land slipper appears to be more tolerant to uneven thermal loading compared to the step design. This can be mainly explained by the fact that the amount of leakage for the multi-land slipper is clearly higher than for the step slipper. This is shown by the efficiency graphs in **Figure 15** in Appendix A, providing pump measurement results over a large set of pressure and speeds values. Additionally, it is assumed that the grooves of the multi-land slipper work as a kind of cooling channels. As the leakage oil has no direct path but has to exit through the dedicated openings on the outer sealing land. This also means that the leakage oil stays in contact with the slipper for a longer time as compared to the single land slipper, where it has a more direct leakage path. This suggests that the multi-land slipper may be a more suitable option than the step slipper for certain applications like dry-case operation, in which the case oil is drained to reduce churning losses.

6. CONCLUSION AND OUTLOOK

This paper presents a hydrostatic tribometer designed to evaluate slipper performance in various scenarios, which are challenging to test in full pump tests. Unlike earlier hydrostatic tribometers, this model uses unaltered original pump parts integrated into a commercial disc-on-disc tribometer. It features an external pressure source and operates with stationary components (block, piston, and

slipper) and a moving swash plate. While it reduces the physical representation of a pump, its focused study on the behaviour of interchangeable slipper designs under specific conditions is beneficial.

Previous results of the innovative step-design on the pump test rig were promising [2], yet some measurement results were difficult to explain even using the advanced simulation tool Caspar FSTI. Particularly notable were the differences in wear levels during high-speed, high-pressure, and high-temperature endurance tests compared to steady-state measurements, which included low speeds and varying conditions. The increased wear in the latter, despite significantly reduced torque losses with the step design, suggested transient conditions as a cause, particularly thermal transients not captured by the simulation tool, which is limited to thermal steady states. The tribometer's low-speed performance evaluations of slipper designs showed trends and magnitudes comparable to pump test rig results. However, increasing pressure created a misleading impression of the step design's inferiority, contrary to pump environment observations. This discrepancy underscores the limitations of tribometers, such as the absence of centripetal force and pressure permutation, which can yield differing results under pressurized conditions. Nonetheless, the findings offered valuable insights into slippers' complex behaviours, particularly in low-pressure scenarios at low speeds.

The thermal boundary conditions in the tribometer are significantly different compared to the regular pump. The slipper and swash plate, not submerged in oil and surrounded by room temperature, exhibit a reduced convection coefficient, especially due to the lack of rotational movement of the slipper. To allow for a better heat dissipation an active cooling system was installed. It blows cold air and sprays oil at room temperature onto the swash plate and slipper. Turning off or adjusting this system allows to test in different thermal boundary conditions. Pump tests and simulations indicated improved slipper torque losses with rising inlet oil temperature, contrasting the presented tribometer results. When the cooling system was turned off there was a significant thermal strain on the slipper with temperature exceeding 130˚C, which lead to black leakage oil and significant slipper wear. However, even with the cooling system turned on, the wear patterns were not matching the pump testing. It was concluded that the high temperature gradient from >100˚C in the gap to surfaces sprayed with cold oil and air caused a higher than usual thermal deformation leading to wear and increased torque readings. This behavior was confirmed by FEM simulations, illustrating that high thermal gradients increase the deformation significantly. Unfortunately, this thermal wear is a self-amplifying process, where higher temperature led to more wear, which lead to even higher local temperatures. The results indicate the need for slippers to operate in a more uniform temperature field to minimize wear. This is achieved when submerging the slipper with leakage oil, which will lead to a smaller temperature gradient even at elevated temperatures.

The shown findings emphasize the complexity of pumps and their components. Changing just one boundary condition could alter the performance drastically. The hydrostatic-tribometer has shown its merits when testing different designs in defined ramps and conditions. However, the resulting trends cannot be projected into the pump environment, as too many variables have changed. Especially the different thermal conditions have shown to be a major factor for this discrepancy. One step further to a comparable pump set-up would be to allow the slipper to run through oil and to establish a homogeneous temperature throughout the sliding parts. Further investigations will involve material studies including a lead-free and a steel slipper.

In conclusion, it can be said that the results give a deeper insight into the complex physical phenomena that slippers exhibit and can be used to test for limits of operation and thermal behaviour. These insights are vital for understanding and optimizing pumps, especially in the context of electrification and the need to test critical speed levels, maximum allowed acceleration, and thermal limits.

7. ACKNOWLEDGEMENTS

The presented research activities are part of the project “Tribologieoptimierung von Pumpensystemen durch fertigungsgerechte Einbringung von Mesostrukturen” (Ref. No. 20757 BR/1). The authors would like to thank the Fluid Power Research Fund of the VDMA for the funding and support.

NOMENCLATURE

α	Thermal expansion coefficient	1/K
B	Balance	%
β	swashplate angle	°
F_{fz}	Hydrostatic force	N
F_{AK}	Piston force	N
L	length of the slipper	mm
p	pressure	bar
r	radius	mm
T	Temperature	°C
ΔT	Temperature difference	K
Δz	Deformation in z-axis	µm

APPENDIX A

Figure 15 displays the volumetric (left) and hydromechanics (right) efficiency graphs of pump measurement results over several pressure blocks from 50 to 300 bar within each speeds from 500 to 1500 rpm are measured. The single measured efficiency values are connected by lines to better visualize the differences of the designs. The purple curve represents the multi-land efficiencies in both graphs, which are lower than the step and the reference efficiencies for all operating conditions concerning the volumetric losses. This supports the theory that the multi-land design produces more leakage that is helping to cool the slipper, reducing the deformation due to temperature and thus preventing contact and wear. As the hydromechanic efficiency of the multi-land slipper is quite good especially for higher pressures, the overall performance of the multi-land slipper is as good as the step slipper for 300 bar and 1-2 percentage points worse for lower pressures. The pump measurements also have shown, that the efficiency advantage of the step slipper over the multi-land slipper even increases for lower swashplate angles due to stabilizing effects of the micro-structure described in more detail in [3].

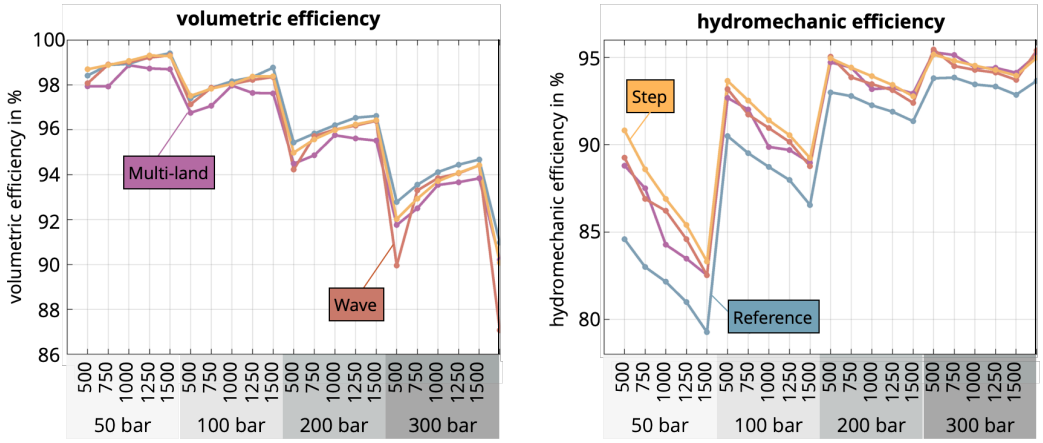


Figure 15: Volumetric and hydromechanic efficiency for different slipper designs and OCs

REFERENCES

1. Ivantysynova M, Baker J (2009) Power loss in the lubricating gap between cylinder block and valve plate of swash plate type axial piston machines. Int J Fluid Power 10:29–43. https://doi.org/10.1080/14399776.2009.10780976
2. Horn S, Ivantysyn R, Weber J (2022) Validated efficiency improvements by implementation of structures on the slipper surface of an axial piston pump. IEEE GFPS 2022
3. Horn S, Ivantysyn R, Weber J (2022) Tribo-optimized lubricating interfaces in hydrostatic pumps with surface shaped slippers. IFK2022
4. Ivantysyn R, Horn S, Weber J (2022) Design of a Lead-Free Slipper Bearing for Low Speed Axial Piston Pump Applications. IEEE GFPS 2022
5. Haidarschin G, Hesebeck M, Su E, Diesselberg M (2018) Benchmarking of potential substituents for leaded bronze in axial sliding bearings for mobile hydraulic applications. Fluid power networks Proc 19th - 21th March 2018 11th Int Fluid Power Conf 54–67
6. Manring ND, Wray CL, Dong Z (2004) Experimental studies on the performance of slipper bearings within axial-piston pumps. J Tribol 126:511–518. https://doi.org/10.1115/1.1698936
7. Renius KT (1973) Experimentelle Untersuchung an Gleitschuhen von Axialkolbenmaschinen. O + P Zeitschrift für Fluidtechnik 17:75–80
8. Canbulut F, Sinanoğlu C, Koç E (2009) Experimental analysis of frictional power loss of hydrostatic slipper bearings. Ind Lubr Tribol 61:123–131. https://doi.org/10.1108/00368790910953631
9. Hooke CJ, Kakoullis YP (1978) The Lubrication of Slippers in Axial Piston Pumps. In: 5th Fluid Power Symposium
10. Bergada JM, Haynes JM, Watton J (2008) Leakage and Groove Pressure of an Axial Piston Pump Slipper with Multiple Lands. Tribol Trans 51:469–482. https://doi.org/10.1080/10402000802044332
11. Rizzo G, Massarotti GP, Bonanno A, et al (2015) Axial piston pumps slippers with nanocoated surfaces to reduce friction. Int J Fluid Power 16:1–10. https://doi.org/10.1080/14399776.2015.1006979
12. KAZAMA T, TSURUNO T, SASAKI H (2008) Temperature Measurement of Tribological

Parts in Swash-Plate Type Axial Piston Pumps. Proc JFPS Int Symp Fluid Power 2008:341–346. https://doi.org/10.5739/isfp.2008.341

13. Tang H sheng, Yin Y bao, Zhang Y, Li J (2016) Parametric analysis of thermal effect on hydrostatic slipper bearing capacity of axial piston pump. J Cent South Univ 23:333–343. https://doi.org/10.1007/s11771-016-3078-0
14. Lin S, Hu J (2015) Tribo-dynamic model of slipper bearings. Appl Math Model 39:548–558. https://doi.org/10.1016/j.apm.2014.06.009
15. Hooke CJ (1983) The effects of non-flatness on the performance of slippers in axial piston pumps. Proc Inst Mech Eng Part C J Mech Eng Sci 197:239–247. https://doi.org/10.1243/PIME_PROC_1983_197_104_02
16. Wang X, Yamaguchi A (2002) Characteristics of hydrostatic bearing/seal parts for water hydraulic pumps and motors. Part 1: Experiment and theory. Tribol Int 35:425–433. https://doi.org/10.1016/S0301-679X(02)00023-3
17. Iboshi N, Yamaguchi A (1982) Characterisitics of a Slipper Bearing for Swash Plate Type Axial Piston Pumps and Motors. Bull JSME 2091
18. Chao Q, Wang Q, Xu B, Chen Y (2017) Experimental verification of slipper spinning motion in axial piston pumps. In: Proceedings of the ASME/BATH 2017 Symposium on Fluid Power and Motion Control FPMC2017. pp 1–7
19. Chao Q, Zhang J, Xu B, et al (2019) Test rigs and experimental studies of the slipper bearing in axial piston pumps: A review. Meas J Int Meas Confed 132:135–149. https://doi.org/10.1016/j.measurement.2018.09.027
20. Chao Q, Zhang J, Xu B, et al (2019) Test rigs and experimental studies of the slipper bearing in axial piston pumps: A review. Meas J Int Meas Confed 132:135–149. https://doi.org/10.1016/j.measurement.2018.09.027
21. Horn S, Ivantysyn R, Schneider J, et al (2022) Tribologieoptimierung von Pumpensystemen durch fertigungsgerechte Einbringung von Mesostrukturen. Dresden, Germany
22. Renius KT (1972) Zum Entwicklungsstand der Gleitschuhe in Axialkolbenmaschinen 1. O + P Zeitschrift für Fluidtechnik 16
23. Shorbagy A, Ivantysyn R, Berthold F, Weber J (2022) Holistic analysis of the tribological interfaces of an axial piston pump - Focusing on pump ' s efficiency. In: IFK2022. Aachen
24. Renius KT (1972) Zum Entwicklungstand der Gleitschuhe in Axialkolbenmaschinen. Ölhydraulik und Pneum 16:494–497
25. Hamrock BJ, Schmid SR (1991) Fundamental of Fluid Film Lubrication

FAST COMPUTATION OF LUBRICATED CONTACTS: A PHYSICS-INFORMED DEEP LEARNING APPROACH

peer reviewed

Faras Brumand-Poor[1]*, Niklas Bauer[1], Nils Plückhahn[1], Katharina Schmitz[1]

[1] *RWTH Aachen University – Institute for Fluid Power Drives and Systems (ifas)*

* Corresponding author: Tel.: +49 241 80 47743; E-mail address: faras.brumand@ifas.rwth-aachen.de

ABSTRACT

The frictional behavior of pneumatic seals determines functionality in various fluid power switching applications. Understanding the complex relationship between friction and component properties is challenging. Experimental descriptions are often infeasible due to time and cost constraints. Therefore, an elastohydrodynamic lubrication (EHL) simulation model, the ifas-DDS, was implemented to calculate the friction in translational pneumatic seals [1]. While the EHL simulation provides an accurate solution to the underlying partial differential equations, it is computationally expensive. An approach to obtain an accelerated solution is the use of neural networks (NN). However, their main disadvantage is that they do not necessarily embed the physical mechanism underlying a particular dataset. In recent years, a new form of NN has emerged, the physics-informed neural network (PINN), which is imposed with physical laws, increasing the algorithm's accuracy. This approach offers several possibilities, e.g., extrapolation and more robust training. In this paper, a variation of the Reynolds equation, implemented in the ifas-DDS, is solved with a PINN. The complete hydrodynamic PINN (HD-PINN) framework is presented and compared to the ifas-DDS afterward. The results show the possibility of the HD-PINN framework for modeling the EHL with high speed, less to no accuracy loss, and minimal tuning effort.

Keywords: Tribology, Physics-informed neural networks; Elastohydrodynamic simulation; Pneumatic sealing; Physics-informed machine learning

1. INTRODUCTION

Lubricated contacts are part of many fluid power components. They occur e.g., in seals of hydraulic cylinders and pneumatic valves or functionally critical contacts to contacts in pumps. Since they often have a crucial influence on the performance and efficiency of the respective components, it is important to comprehensively understand all phenomena within these contacts by physically motivated models. However, obtaining an analytical solution for the complex phenomena in lubricated contacts is often impossible without making several simplifications. Furthermore, characterizing this through experiments is often unfeasible due to their time-consuming and expensive nature.

Consequently, the behavior of lubricated contacts is modeled using an elastohydrodynamic lubrication (EHL) simulation, representing the acting phenomena through the Reynolds equation, and defining the pressure distribution within the tribological contact. One example of a validated EHL simulation model is the ifas-DDS, which was implemented at ifas to calculate the frictional behavior of reciprocating pneumatic seals. This is achieved by solving the Reynolds equation to obtain the hydrodynamic pressure within the contact and a FEM calculation for the seal's deformation. The EHL yields good agreement with experimental data, shown in **Figure 1**, but requires a long computation time for a rather short-term process, due to the complexity of the partial differential equations and their numerical solution [2]. A straightforward method to reduce

computation time is to enhance computational power. However, this may not always be feasible and may even become more challenging as the simulation grows in complexity. This is especially true when a component with multiple contacts or even a system consisting of multiple of these components is to be considered. In that case, the computation time of current EHL simulation models makes these simulations almost impossible to apply.

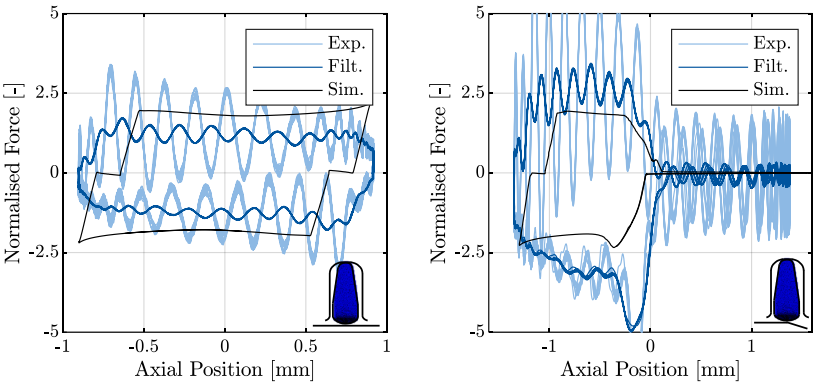


Figure 1: Comparison of measured (filtered and unfiltered) and simulated friction force. Left: Without moving over the control edge. Right: Moving over the control edge. [2]

A novel approach to accelerate the computational process of simulations has emerged in recent years: the physics-informed neural network (PINN). Traditional neural networks require a shorter computation time regarding most iterative solver of distributed parameter simulations. However, they lack an understanding of the physical laws governing a specific dataset. In typical regression tasks, the objective of a neural network is to minimize the deviation between the predicted output and the true output. However, tuning the network's parameters to achieve this goal may only be valid within the provided data range. In contrast, PINNs incorporate the underlying physical laws of the investigated simulation, providing a more universal and accelerated solution for complex simulation models.

This contribution introduces a framework for developing and parametrizing a PINN to speed up the computation of pressure distribution in EHL models. For this, the framework's various components are described and the advantages regarding a successful implementation of PINNs are highlighted. Firstly, the hydrodynamic lubrication (HL) is presented in section 2. Succeeding, an overview of the literature regarding PINNs, their application for HL, and their main drawbacks is provided. The fourth section of this paper deals with the whole framework. It describes the implemented solutions for the common issues encountered in PINNs described in the previous section. In section 5 the accuracy of the PINN is compared to the rigid ifas-DDS. In this contribution the deformation of the seal is not considered, therefore resulting in a purely HL model. Eventually, the results are summarized in section 6.

2. HYDRODYNAMIC LUBRICATION

EHL simulations are employed to calculate friction, leakage, and wear characteristics within lubricated contacts. These simulations account for the interplay between the lubricant and contacting surfaces, involving the computation of surface deformations within the contact area and the subsequent build-up of hydrodynamic pressure.
The ifas-DDS is a simulation model used to describe the interaction between a seal and the adjacent

counterface. It takes into consideration the presence of lubricating fluid between the seal and counterface. The deformation of the seal is determined through the utilization of the finite element software Abaqus. The HL phenomena are described by the Reynolds equation and integrated in Abaqus through user subroutines.

In the scope of this study, the primary emphasis is on solving the Reynolds equation and developing the PINN framework. Consequently, the deformation of the two contact partners is not taken into consideration. Typically, a seal undergoes deformation under load, leading to modifications in the HL behavior. To validate the presented PINN, the simulation model ifas-DDS without deformation is employed, herein referred to as rigid ifas-DDS. The ifas-DDS in its original form encompasses an extended Reynolds equation using flow factors Φ^{τ} and Φ^{p}, as described by Patir and Cheng, which allows for the consideration of surface topography [1], and an implementation of the Jakobsson-Floberg-Olsson cavitation model [3] introducing the parameter θ.

The complete Reynolds equation used in the ifas-DDS extending the originally derived equation from Osbourne Reynolds in 1886 is given as [4]:

$$\frac{v}{2}\frac{\partial}{\partial x}\Big((1-\theta)\rho h + (1-\theta)\rho R_q \Phi^{\tau}\Big) - \frac{1}{12\eta}\frac{\partial}{\partial x}\left(\Phi^{p}\rho h^3(1-\theta)\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial t}\big((1-\theta)\rho h\big) = 0 \quad (1)$$

3. PHYSICS-INFORMED NEURAL NETWORKS

As mentioned in the preceding section, the Reynolds equation is typically solved using finite volume, element, or difference methods. Recently, machine learning models have found success in the field of tribology [5,6]. Notable examples include the application of autoencoders and convolutional neural networks for fault detection in tribological systems. Particularly noteworthy is the work of Hess and Shang, who developed a convolutional neural network to determine the elastohydrodynamic pressure distribution in journal bearings [7]. Machine learning, especially neural networks as universal function approximators, has demonstrated significant potential, not only for data modeling but also for solving partial differential equations.

The mathematical foundation, demonstrating that a feed-forward neural network with at least one hidden layer can approximate any continuous function with arbitrary accuracy, was laid by Cybenko in 1989 [8]. In the same year, Hornik extended this to Borel measurable functions [9]. A year later, Hyuk, one of the pioneers in the field of physics-informed neural networks, investigated the ability to solve differential equations with neural networks [10]. While Hyuk did not use the term "physics-informed," the idea and motivation behind his work align with the principles of PINNs. The loss function for the neural network under investigation was expanded to incorporate the underlying differential equation.

However, this approach was relatively overlooked but saw a resurgence in interest due to progress in available computational resources, more advanced ML models, and efficient gradient calculation techniques by automatic differentiation.

This revival of physics-inspired machine learning was led by Owhadi, as one of the first in 2014, embedding prior knowledge about the desired solution. He suggested reformulating the PDE solution problem as a Bayesian inference problem and presented the potential of enhancing the algorithm with prior knowledge [11]. Raissi et al. developed a probabilistic ML algorithm for

general linear equations using the Gaussian process [12,13]. The Gaussian process is tailored to the investigated integro-differential or partial differential equation. This approach was extended by Raissi et al. for nonlinear partial differential equations [14,15].
A significant milestone for the physics-informed machine learning community was the introduction of physics-informed neural networks [16], which Cuomo et al. describe as a mesh-free technique for solving PDEs by transforming the problem of directly solving the underlying equation into a loss function optimization problem [17].
Raissi introduced PINNs as a novel data-driven solver class in a two-part publication in 2017 and later in a merged article in 2019, PINNs aim to concurrently minimize the deviation of known values such as boundary and initial conditions (supervised losses), and the PDE residual (unsupervised), within the investigated computational space [18,19,20].

The first published work on PINNs solving the Reynolds equation was by Almqvist in 2021, focusing on the fundamental application of PINNs for the Reynolds equation [21]. More advanced algorithms were developed by Zhao et al. and Li et al., dealing with the 2D Reynolds equation for linear sliders and gas bearings, respectively [22, 23]. Most recently, Rom became the first to develop PINNs for solving the stationary Reynolds equation with the Jakobsson-Floberg-Olsson (JFO) cavitation model [24]. Building on this success, Cheng et al. implemented a PINN to solve the Reynolds equation with either the JFO or the Swift-Stieber (SS) cavitation model, applying three different multi-task learning methods to balance the various loss terms [25]. It's worth noting that most of these publications primarily focus on the PINN itself rather than on a framework for developing PINNs to solve HL tasks. Therefore, a significant amount of manual work is still required.

In **Figure 2** an exemplary hybrid (physics-informed and data-based) PINN is shown [26]. The four key elements are the neural network, the automatic differentiation, the losses, and the optimizer. As already described, the loss contains more than the difference between actual and predicted output. The residual loss term of the PDE is determined by the computation of partial derivatives using automatic differentiation [27]. Using the chain rule, the derivatives of the outputs are computed with respect to the network inputs, weights, and biases. Typically, the training of PINNs is based on the L2-norm (mean squared error, MSE) computed on uniformly sampled collocation points within the computational space [26]. In this work, the residual loss of the Reynolds equation is computed for uniformly distributed positions of the seal and the adjacent counterface.

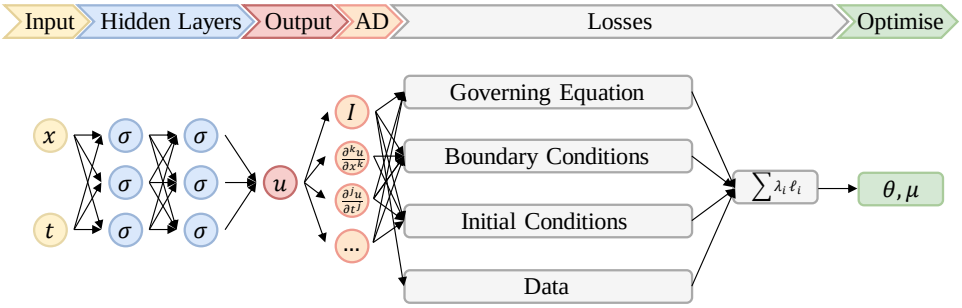


Figure 2: Schematic illustration of a physics-informed neural network. Adapted from [26].

In **Figure 2** the four different types of loss terms are illustrated: The loss based on the governing equation (e.g., Reynolds equation), on the boundary conditions (e.g., pressure at the boundaries of the investigated geometry), on the initial conditions (e.g., the initial pressure distribution over the whole geometry) and eventually on the available data (provided by the e.g., ifas-DDS).

The traditional PINN consists of the three first loss terms, adding the data loss results in a hybrid PINN, which is not considered in this work, to investigate the potential of the physics-informed idea. These losses are minimized jointly and therefore represent a multi-objective optimization. This class of optimization deals with concurrently minimizing a set of more than one, potentially conflicting objective [28]. Optimization problems concerning engineering tasks generally require trade-offs to synchronously fulfill all terms to the desired degree [29]. The desired solution for multi-objective optimization problems can be represented as a set of Pareto optima, which defines a feasible solution, where one objective cannot be minimized without worsening another term [30]. A multi-objective optimization is independent of balancing the different terms [31]. However, the non-convex nature of the physical solution space in multi-objective optimization poses challenges for finding the globally optimal solution through gradient-based optimization [26]. Moreover, the gradients in multi-objective optimization of physical systems can vary greatly. Most optimizer adjusts the network parameters using an explicit scheme (the updated parameter is based on the current parameter) and can face instability if the optimization steps are too high. For classical neural networks, the problem of different magnitudes in the loss gradient occurred rarely, leading to the predominant use of explicit schemes instead of implicit ones (considering the updated parameter in the update itself) to speed up the training procedure. However, this problem occurs in PINNs and generally requires a reduction of the loss update in each step. Unfortunately, a small magnitude can trap the optimizer in a local minimum. Therefore, a loss balancing algorithm is implemented for the hydrodynamic PINN (HD-PINN) framework and presented in section 4.3.

The training of PINNs involves a great number of design options (hyperparameters), e.g., network width, network depth, and learning rate. As mentioned, PINNs solve multi-objective optimization, which introduces additional hyperparameters to the whole set due to the need for loss balancing. The high number of dimensions of these parameters makes it challenging to find the optimal set. Therefore, finding the optimal hyperparameter set, yielding the best performance, becomes an iterative process. Depending on the specific problem, a single training routine can be time-consuming, and infeasible for a high number of iterations. Basic approaches such as manual or grid search suffer from the curse of dimensionality and might be inefficient [32]. This problem of PINNs is addressed by enhancing the HD-PINN framework with a sophisticated and automated hyperparameter search and presented in section 4.2

In summary, PINNs offer a promising method for solving HL simulation models by optimizing their parameters based on several physics-informed objectives. The following section introduces the HD-PINN framework, which includes the investigation of the differential equation, automated hyperparameter search, and loss balancing.

4. METHODOLOGY

In this section, the HD-PINN framework is presented. First, the investigated Reynolds equation is derived, and the PINN input and output are described. Then, the loss calculations for the residual loss term of the Reynolds equation and the pressure boundaries are explained briefly. The subsequent part deals with the efficient and robust parameter tuning of the framework. The third part of this section presents an algorithm for the improvement of the training due to loss balancing.

4.1. PINN for Solving the Reynolds Equation

As stated in the previous section, a handful of researchers have investigated the PINN for solving the Reynolds equation. The researchers focused primarily on the PINN itself and therefore still had

to manually tune a great number of design options e.g., layer width or depth.
In this work, the emphasis is on creating an automated HD-PINN framework designed to address the common issues and limitations associated with PINNs. It is initially validated using a simplified version of the Reynolds equation but can be easily extended to the complete mathematical expression with the transient term and the JFO cavitation model. The underlying assumptions are as follows:

- Time dependencies are neglected ($\frac{\partial}{\partial t} = 0$)
- Cavitation is not considered ($\theta = 1$)
- The surface is ideally smooth ($\Phi^{\tau} = 0, \Phi^{p} = 1$)

Therefore, the investigated Reynolds equation holds:

$$\frac{v}{2}\frac{\partial}{\partial x}\rho h - \frac{1}{12\eta}\frac{\partial}{\partial x}\left(\rho h^3 \frac{\partial p}{\partial x}\right) = 0 \tag{2}$$

The PINN obtains the dynamic viscosity η and density of the fluid ρ, the height h, and the velocity of the counter surface v as input and determines the pressure p as output. The equation above contains partial differentials, which are required to obtain the physics-informed loss term for the residual of the Reynolds equation. These values are obtained by automatic differentiation, which allows for accurate and efficient function derivation [33].
Furthermore, the boundary loss terms, which are Dirichlet boundary conditions for the Reynolds equation, consist of two values for the left and right boundary respectively. These are compared to the boundary pressure predicted by the PINN and embedded as an MSE term. For PDEs including Neumann boundary conditions, automatic differentiation can be used to efficiently determine the desired loss function. The physics-informed Reynolds loss is implemented as the following:

$$\ell_{Rey} = \mathrm{MSE}\left(\frac{v}{2}\frac{\partial}{\partial x}\rho h - \frac{1}{12\eta}\frac{\partial}{\partial x}\left(\rho h^3 \frac{\partial p}{\partial x}\right), 0\right) \tag{3}$$

The MSE of the Reynolds equation compared to zero is calculated as the loss and strongly resembles the residual term of common EHD simulations. For the boundary loss the following function is implemented:

$$\ell_{\mathrm{BC},l,r} = \mathrm{MSE}\left(p_{HD-PINN,l,r}, p_{b,l,r}\right) \tag{4}$$

In summary, the PINN loss comprises three terms that do not depend on any specific data provided by simulation of experiments. The rigid ifas-DDS primarily serves to acquire a pressure distribution for validating the PINN. In the next sections, two main features of the HD-PINN framework are presented. These solve the previously explained common drawbacks occurring in the application of PINNs. Eventually, the whole training process is described.

4.2. Hyperparameter Optimization

Compared to traditional neural networks, PINNs exhibit even more hyperparameters, especially due to the already-mentioned loss balancing. Basic approaches for finding optimal parameters often

suffer from the curse of dimensionality and are therefore not feasible for the HD-PINN framework.

Bayesian optimization [34] offers a solution to this dilemma and has already been successfully integrated into PINNs [32]. Bayesian optimization aims to obtain an optimal set of hyperparameters with as few training procedures as possible by approximating an unknown loss function (PINN losses) with a probabilistic surrogate model (Gaussian Process). After evaluating one set of hyperparameters, the next set is chosen based on an acquisition function (expected improvement), which aims to achieve a trade-off between exploration and exploitation by comparing the current best set to the next chosen set.
The HD-PINN framework incorporates Bayesian optimization to determine the optimal hyperparameters, including the number of layers, layer width, and four parameters relevant to the loss balancing task: learning rate, decay rate, temperature factor, and saudade value. These parameters will be introduced in the following section.

4.3. Loss Balancing

The Bayesian optimizer aims to find the optimal set of hyperparameters for the given task. For a given set of hyperparameters, their performance is analyzed by training the PINN for a defined number of epochs and eventually evaluating the final sum of losses. During one epoch, the PINN determines the pressure distribution based on the provided inputs for a pre-defined number of collocation points (ten in this work). With the pressure distribution, the residual loss of the Reynolds equation and the loss for the boundary values are determined.
The loss is further utilized and differentiated with automatic differentiation to calculate the changes in the network parameters, including weights and biases. The optimization of these network parameters is achieved using the adaptive moment estimation (ADAM) algorithm, a first-order gradient-based explicit optimization method for stochastic objective functions [35]. ADAM is widely recognized as one of the state-of-the-art optimizers for deep neural networks [36,37]. It is computationally efficient with little memory requirements, capable of solving problems of large scale, and has been successfully implemented in PINN applications [38].

As stated in section 3, explicit optimizers might face problems due to the heavily varying loss term magnitudes. Therefore, a loss balancing scheme according to [26] is added to the HD-PINN framework. Bischof proposed an algorithm based on combining three already existing methods and adding one novel feature. The four features of the Relative Loss Balancing with Random Lookback (ReLoBRaLo) are as follows:

1. The sum of scalings is bound by a softmax function, which is adjusted by the so-called temperature factor $\mathcal{T}$ [39].
2. The learning progress is considered by dividing the loss at the current iteration by the loss of the last iteration [40].
3. Inspired by the Learning Rate Annealing, the scalings are set to higher values initially and decay exponentially over the iterations with the decay factor α. This leads to the utilization of loss statistics from more than one training step [41].
4. A random lookback (defined by the saudade value ρ_s) is embedded in the exponential decay and randomly decides the consider the previous steps' loss statistics or look back to the beginning of the training to compute the scalings [26].

This novel balancing algorithm adds a significant performance enhancement to the HD-PINN framework by tackling the loss scaling problems. **Figure 3** illustrates the training

progress of the complete HD-PINN framework. The Bayesian optimizer selects parameters, which are used for the initialisation of the PINN and the network parameter optimizer. Afterward, the actual pressure prediction is done by the PINN and the loss is computed. If the training is not finished, the loss is used to update the PINN's weights and biases, and a new pressure distribution is predicted. This loop continues until a certain number of iterations is reached. Afterward, the parameter determined by the Bayesian optimizer is updated.
In the next section, the whole framework is validated for two different scenarios and an illustration of the training progress is provided.

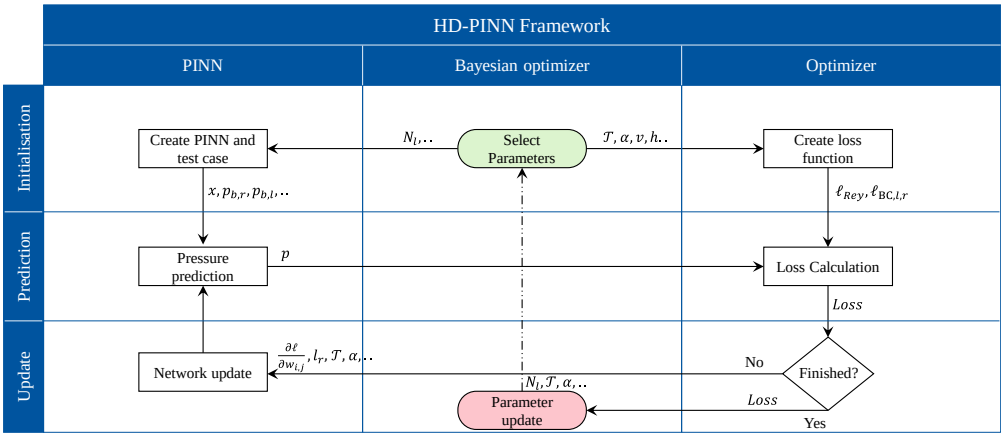


Figure 3: The HD-PINN framework and its training progress.

5. RESULTS AND DISCUSSION

The HD-PINN is tested for two test scenarios, shown in **Figure 4**, a convergent ($h = [1, 0.5]$), and a divergent ($h = [0.5, 1]$), height profile. The convergent case resembles Almqvist's work for a linear slider with dimensionless film thickness [26]. However, in comparison to that work, the pressure boundary is not set to zero (left pressure for the convergent gap $p_{b,l} = 0.2$, for the divergent gap $p_{b,l} = 0.7$ and right pressure for the convergent gap $p_{b,r} = 0.7$, for the divergent gap $p_{b,r} = 0.4$).

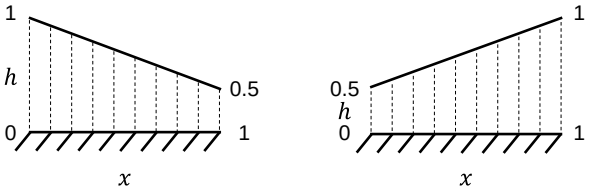


Figure 4: Left: Convergent height profile Right: Divergent height profile.

The PINN is tuned by the Bayesian optimizer for 30 trials and eventually, the best-estimated parameters are used for the actual training process (200 training iterations). The validation data is obtained by the rigid ifas-DDS. It is noteworthy that the convergent height profile required the maximum number of iterations (1000) to obtain the residues set in the ifas-DDS.

First, the convergent height profile is presented. In **Figure 5** the pressure distribution and the loss progression are illustrated. The HD-PINN predicts the pressure with high accuracy, with a maximum deviation of around 1 %. The residual loss of the Reynolds equation and the boundary loss decrease drastically over 200 epochs, with minor fluctuations in the boundary term and none in

the Reynolds term.

For the divergent case, an even higher accuracy of the pressure distribution (deviation of less than 1 %) can be observed in **Figure 6**. Similar to the first case, the Reynolds-based loss remains stable without any fluctuations, while the boundary loss term exhibits more volatility.

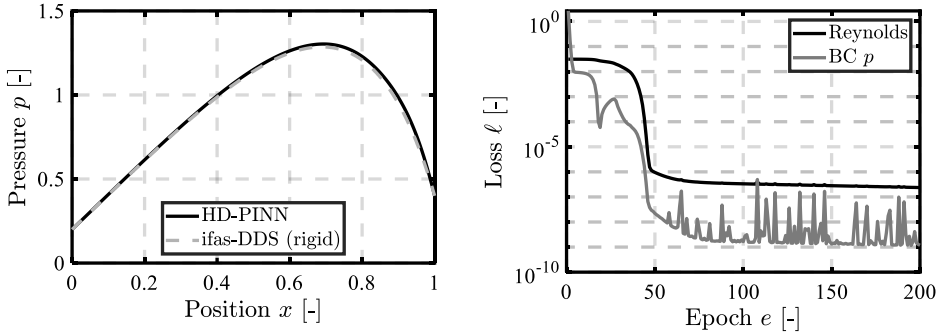


Figure 5: Left: Pressure distribution of HD-PINN vs. rigid ifas-DDS for the convergent gap. Right: The loss trajectories for the convergent gap.

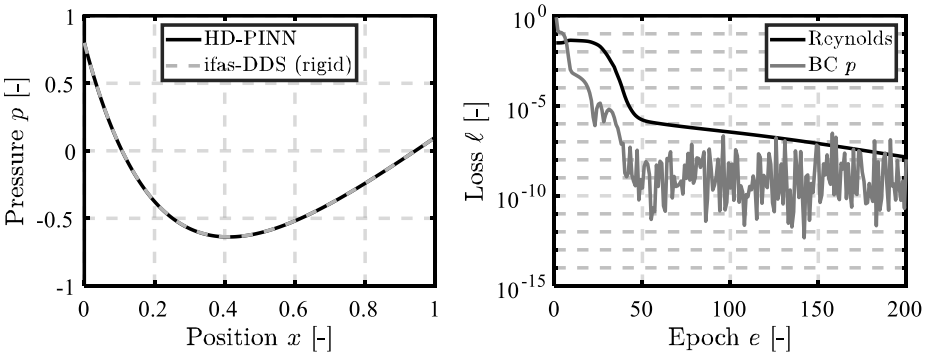


Figure 6: Left: Pressure distribution of HD-PINN vs. rigid ifas-DDS for the divergent gap. Right: The loss trajectories for the divergent gap.

To obtain a visual representation of the learning process of the PINN four different pressure distributions after different epochs (1, 10, 30, 40) are illustrated in **Figure 7**. Like the boundary loss trajectories, the pressure boundary values are met first and already precisely predicted after ten epochs. An accurate prediction is achieved after 40 epochs, which illustrates the fast-learning ability of the PINNs and suggests decreasing the whole 200 epochs to reduce the training time in future research.

6. SUMMARY AND CONCLUSION

This contribution introduces an automated HD-PINN framework for predicting pressure build-up in sealing contacts within a housing, described by the Reynolds equation. It starts by presenting the concept of physics-informed neural networks and their advantages. Then, it addresses the common drawbacks of PINNs and describes various components implemented in this work to mitigate these issues.

The incorporation of the Bayesian optimizer and the loss balancing algorithm reduces the amount of manual tuning of the hyperparameters and scaling the loss terms.

The validation of the framework, utilizing a variant of the Reynolds boundary equation with a fixed height profile demonstrates the PINNs' ability to model lubricated contacts with high speed and, notably, with minimal accuracy loss compared to traditional HL simulations.
Future steps involve integrating the missing terms of the Reynolds equation to obtain an accelerated solution for the complete rigid ifas-DDS, which encompasses transient behavior and cavitation. Additionally, an extended set of inputs will be defined to characterize the investigated geometry, broadening the PINN's range of validity to accommodate various geometries rather than being specific to just one. Furthermore, a second model will be developed to account for sealing deformation and thereby address the entire EHL in the ifas-DDS.

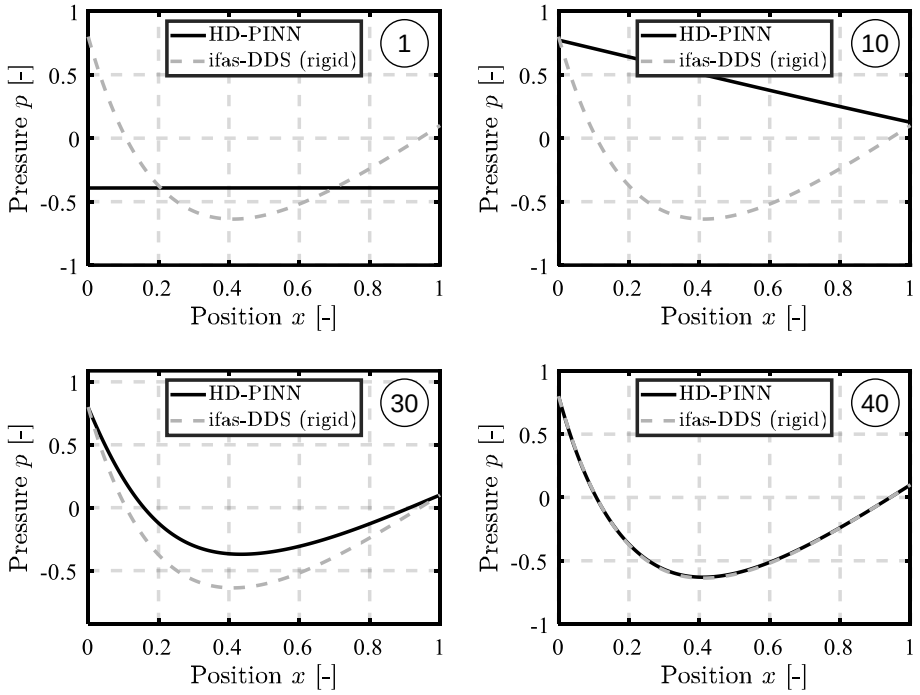


Figure 7: Pressure distribution of the HD-PINN after 1, 10, 30, and 40 epochs for the divergent gap.

ACKNOWLEDGMENTS

The authors thank the Research Association for Fluid Power of the German Engineering Federation VDMA for its financial support. Special gratitude is expressed to the participating companies and their representatives in the accompanying industrial committee for their advisory and technical support.

NOMENCLATURE

e	Epoch	-
h	Dimensionless height profile	-
ℓ	Loss	-
l_r	Learning Rate	-
N_l	Number of hidden network layers	-

t	Dimensionless time	-
p	Dimensionless pressure	-
$p_{b,l,r}$	Dimensionless left and right boundary pressure	-
R_q	Root of the mean squared roughness	-
v	Velocity of the counter surface	-
x	Dimensionless position	-
α	Decay	-
η	Dimensionless dynamic viscosity	-
θ	Dimensionless cavitation	-
ρ	Dimensionless density	-
ρ_s	Saudade value	-
Φ^p	Flow factor	-
Φ^τ	Shear flow factor	-
$\mathcal{T}$	Temperature factor	-
$\frac{\partial}{\partial x}$	Partial derivative regarding the position and time	-
$\frac{\partial}{\partial t}$	Partial derivative regarding the time	-

REFERENCES

[1] Bauer, N., et. al. (2023) Elastohydrodynamic Simulation of Pneumatic Sealing Friction Considering 3D Surface Topography. Chem Eng & Technol 46 (1).

[2] Bauer, N., et. al. (2022) Experimental determination and EHL simulation of transient friction of pneumatic seals in spool valves. International Sealing Conference: 21st ISC.

[3] Angerhausen, J., et al. (2019) Simulation and experimental validation of translational hydraulic seal wear. Tribology International 134 296-307

[4] Bauer, N., et. al. (2021) Strategies for Implementing the Jakobsson-Floberg-Olsson Cavitation Model in EHL Simulations of Translational Seals. Journal of Fluid Power, 22(2), 199-232

[5] Marian M., Tremmel S. (2021) Current Trends and Applications of Machine Learning in Tribology—A Review. Lubricants 9(9), 86

[6] Paturi. U., Palakurthy S., Reddy N. (2023) The Role of Machine Learning in Tribology: A Systematic Review. Archives of Computational Methods in Engineering 30, 1345-1397

[7] Hess N., Shang L. (2022) Development of a Machine Learning Model for Elastohydrodynamic Pressure Prediction in Journal Bearings. J. Tribol 144(8)

[8] Cybenko G. (1989) Approximation by Superpositions of a Sigmoidal Function. Math. Control Signals Systems 2, 303-314

[9] Hornik K., Stinchcombe M., White H. (1989) Multilayer feedforward networks are universal approximators. Neural Networks, 2(5) 359-366

[10] Lee H., Kang S. (1990) Neural algorithm for solving differential equations. Journal of Computational Physics 91(1) 110-131

[11] Owhadi H. (2014) Bayesian Numerical Homogenization. arXiv: 1406.6668v2

[12] Raissi M., Perdikaris P., Karniadakis G. (2016) Inferring solutions of differential equations using noisy multi-fidelity data. arXiv: 1607.04805v1

[13] Raissi M., Karniadakis G. (2017) Machine Learning of Linear Differential Equations using Gaussian Processes. arXiv: 1701.02440v1

[14] Raissi M., Perdikaris P., Karniadakis G. (2017) Numerical Gaussian Processes for Time-dependent and Non-linear Partial Differential Equations. arXiv: 1703.10230v1

[15] Raissi M., Karniadakis G. (2017) Hidden Physics Models: Machine Learning of Nonlinear Partial Differential Equations. arXiv: 1708.00588v2

[16] Blechschmidt J., Ernst O. (2021) Three Ways to Solve Partial Differential Equations with Neural Networks -- A Review. arXiv:2102.11802v2
[17] Cuomo S., et al. (2022) Scientific Machine Learning Through Physics-Informed Neural Networks: Where we are and What's Next. Journal of Scientific Computing 92(3)
[18] Raissi M., Perdikaris P., Karniadakis G. (2017) Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. arXiv: 1711.10561v1
[19] Raissi M., Perdikaris P., Karniadakis G. (2017) Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations. arXiv: 1711.10566v1
[20] Raissi M., Perdikaris P., Karniadakis G. (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378(1), 686-707
[21] Almqvist A. (2021) Fundamentals of Physics-Informed Neural Networks Applied to Solve the Reynolds Boundary Value Problem. Lubricants 9(8)
[22] Zhao Y., Guo L., Wong P. (2022) Application of physics-informed neural network in the analysis of hydrodynamic lubrication. Friction, 11, 1253–1264
[23] Li L. et. al. (2022) ReF-nets: Physics-informed neural network for Reynolds equation of gas bearing. Computer Methods in Applied Mechanics and Engineering, 391
[24] Rom M. (2023) Physics-informed neural networks for the Reynolds equation with cavitation modeling. Tribology International 179: https://doi.org/10.1016/j.triboint.2022.108141
[25] Cheng Y., et. al. (2023) HL-nets: Physics-informed neural networks for hydrodynamic lubrication with cavitation. Tribology International 188: https://doi.org/10.1016/j.triboint.2023.108871
[26] Bischof R., Kraus M. (2022) Multi-Objective Loss Balancing for Physics-Informed Deep Learning. arXiv: 2110.09813v2
[27] Cai S., et. al. (2021) Physics-informed neural networks (PINNs) for fluid mechanics: A review. arXiv:2105.09506v1
[28] Caruana R. (1997) Multitask Learning. Machine Learning 28, 41-75
[29] Chang K.-H. (2015) Chapter 17 – design optimzation. E-Design Ed. Academic Press, 907-1000
[30] Sener O., Koltun V. (2019) Multi-Task Learning as Multi-Objective Optimization. arXiv:1810.04650v2
[31] Heydari A., Thompson C., Mehmood A. (2019) SoftAdapt: Techniques for Adaptive Loss Weighting of Neural Networks with Multi-Part Loss Functions. arXiv:1909.04630v1
[32] Escapil-Inchauspé P., Ruz G. (2023) Hyper-parameter tuning of physics-informed neural networks: Application to Helmholtz problems. arXiv:2205.06704v2
[33] Baydin A., et. al. (2018) Automatic Differentiation in Machine Learning: a Survey. Journal of Machine Learning Research 18(153):1−43
[34] Močkus J. (2023) On bayesian methods for seeking the extremum. Optimization Techniques IFIP Technical Conference Novosibirsk: https://doi.org/10.1007/3-540-07165-2_55
[35] Kingma D., Ba J. (2014) Adam: A Method for Stochastic Optimization. arXiv:1412.6980v9
[36] Schmidt R., Schneider F., Henning P. (2021) Descending through a Crowded Valley - Benchmarking Deep Learning Optimizers. arXiv:2007.01547v6
[37] Reyad M., Sarhan A., Arafa M. (2023) A modified Adam algorithm for deep neural network optimization. Neural Comput & Applic 35, 17095–17112
[38] Singh S., et. al. (2023) Adam Optimization of Burger's Equation Using Physics-Informed Neural Networks. International Conference on Advancement in Computation & Computer Technologies
[39] Rajeswaran A., et. al. (2019) Meta-Learning with Implicit Gradients. arXiv:1912.12355v1
[40] Chen Z. et. al. (2018) GradNorm: Gradient Normalization for Adaptive Loss Balancing in Deep Multitask Networks. arXiv:1711.02257v4
[41] Wang S., Teng Y., Perdikaris P. (2020) Understanding and mitigating gradient pathologies in physics-informed neural networks. arXiv:2001.04536v1

Chapter 6

Materials

ADDITIVE MANUFACTURING OF HYDRAULIC COMPONENTS – PRESSURE LOSS COMPARISON OF DIFFERENT SELF-SUPPORTING CHANNEL GEOMETRIES

peer reviewed

Zita Tappeiner[1]*, Michiel Donners[1], Matthias Schmid[1], Katharina Schmitz[1]

[1]*Institute for Fluid Power Drives and Systems, RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen*

* Corresponding author: Tel.: +49 241 8047729; E-mail address: zita.tappeiner@ifas.rwth-aachen.de

ABSTRACT

Additive manufacturing (AM) and in particular laser powder bed fusion (LPBF) are increasingly being used as manufacturing technology in hydraulics for flow optimization, function integration and weight reduction. These advantages can especially be exploited in hydraulic manifolds. Conventional manifold intersections are created by crossing two vertical bores. The turbulence resulting from the sharp edges and the deflection leads to undesired flow losses. These can be avoided with the design freedom of LPBF, which allows flow optimization in hydraulic channels. However, the development of new channel geometries is limited by design guidelines. Starting from a straight, round channel geometry, this paper presents the steps to design self-supporting channel geometries for horizontal build up. Therefore, different cross-sectional shapes are tested, and critical design details are explained. In addition, this paper examines the influence of post-processing methods on AM components. A comparison of the different geometries is shown with a CFD simulation as well as FEM simulation for strength investigation. For experimental investigation and simulation validation, selected test specimens were printed and post-processed. With a new designed test rig, the pressure losses of the different geometries and post-processing methods were measured and a comparison with the simulative results is shown. Overall, this paper provides an overview of the necessary steps in the design of hydraulic AM components for flow optimization.

Keywords: Additive Manufacturing, Laser Powder Bed Fusion, Flow Optimization, Design, 316L

1. INTRODUCTION

Components for hydraulic applications are mostly manufactured using conventional manufacturing processes, such as drilling and milling. Accordingly, manifold intersections are usually arranged at 90° to each other and dead volumes occur. Optimized positioning of the channels is possible, but the channel course can only be changed to a limited extent due to the tool geometry and is therefore mostly straight. [1–3]

The use of additive manufacturing (AM) for hydraulic components opens up many new advantages. First of all, structurally optimized components can be designed. This leads to an increase in efficiency due to the weight reduction in moving applications and enables a general material saving. Material resources and component weight can thus be saved. In addition, components can be designed with optimized flow. Flow-optimized channels can reduce pressure losses in the lines and thus save energy. An additional advantage is the functional integration that can be achieved through AM. This means that additional components can be saved, or the design can be adapted to desired functions. Due to these advantages, AM is of increasing importance in the production of hydraulic components, in particular in the production of manifolds [4–8]. Especially the targeted optimization of manifolds in

terms of flow and weight, while maintaining strength, is the subject of numerous research projects [9–13]. The aim of the latest studies is to automate the design process. After entering boundary conditions, entire manifolds can be automatically optimized and designed for AM [14]. For this new way of developing components, new design guidelines will be made available to serve as a template for the design of AM components. [15] Some of these guidelines are even specifically adapted to manifolds. [16] However, neither these guidelines nor the "Design for additive Manufacturing" (DfAM) [17] provide concrete rules for individual case decisions [18]. For example, if a circular cross-sectional area is not horizontally printable, various self-supporting channel geometries are available. In many publications, however, the decision process for a specific shape is not comprehensible or can only be applied to a few applications. This paper therefore aims to provide basic design recommendations for the design of hydraulic channels. As an example, the material 316L (1.4404) is treated in this study. Therefore, statements on material properties refer specifically to this material, but are often transferable to other materials.

1.1. Additive Manufacturing and post-processing

The manufacturing process studied in this paper is Laser Powder Bed Fusion (LPBF). The Fraunhofer Institute for Laser Technology [19] in Aachen played a key role in developing this process, which was patented back in 1996 [20]. LPBF is based on the layer-by-layer melting and subsequent solidification of components from metal powder. Components can thus be generated directly from a digital model. In this way, components with complicated geometries can be manufactured that cannot be realized using conventional methods [21].

Conventional manufacturing processes usually result in uniform microstructures, whereas additive manufacturing (AM) parts typically exhibit noticeable porosity and increased surface roughness [22]. However, the results depend largely on the specific printing parameters. Key factors such as printing direction, laser parameters, porosity, defects, surface finish, and subsequent heat treatment affect the mechanical properties of the material.

Likewise, the post-processing influences on material parameters. For example, the thermal post-processing method hot isostatic pressing (HIP), leads to increased density by eliminating porosity [23]. However, the thermally post-treated components only achieve lower yield strength compared to the as-printed samples. In contrast, a significant increase is observed in the maximum elongation and in ductility [24]. High temperatures can relieve residual stresses in the component and eliminate the direction-dependent properties from applying the material layer by layer. The microstructure is homogenized by the high temperature so that there are no longer melt pool boundaries [25]. Abrasive Flow Machining (AFM), a process for precise surface enhancement, can be used as a post-processing step after LPBF to improve surface quality. In the AFM process, viscoelastic polymers loaded with abrasive particles are passed through voids or channels to mitigate the inherently rough surface finish of untreated components [26]. For hydraulic applications, the target roughness values are typically around 0,4 µm [27]. In fluid dynamics, increased surface roughness results in pressure losses and therefore increased energy losses, reducing overall system efficiency [28]. When AFM is combined with HIP, further enhancements can be expected [29]. Rösler's comprehensive study [30] compared AFM with alternative methods, including vibratory finishing, sandblasting, chemically assisted vibratory finishing, and wet blasting. The investigation revealed that chemically assisted vibratory finishing (AFM) effectively reduced the R_a value from 10 µm to 0,7 µm, offering a cost- and time-efficient approach. In addition, AFM can improve the mechanical properties and lifetime of channels manufactured by the LPBF process. Increased internal surface roughness leads to the development of cracks, which negatively affects fatigue strength [31].

1.2. Comparison concept for channel geometries

A radius of a pipe or channel is often used to calculate flow parameters in a hydraulic system. This radius is known if a channel is circular but can also be calculated if the channel-shape differs from this. In this case, the hydraulic diameter can be described according to (1). It is calculated by multiplying the surface of the channel by 4 and dividing it by the circumference. This formula can be applied in turbulent flows (Re>2300) but is practically used in laminar flows as well, to analytically calculate and describe non-circular channels, although the invalidity of the formula in this state of flow. This formula offers a possible calculation to describe non-circular channels, which are mainly being investigated in this paper.

$$D_{hyd} = \frac{4A}{U} \tag{1}$$

Because of the invalidity of this formula in laminar flows, two other concepts for comparing different geometries are introduced in **Figure 1**. One approach for comparability is to keep the cross-sectional area constant. With constant volume flow, this would lead to constant average velocities. However, considering the transition to the round fitting geometry, it is difficult to keep the cross-sectional area constant over the entire length of the structure. The last option, which will be further pursued in this work, is a constant inner circle of the geometries. This method leads to as little influence on the run-in flow profile as possible.

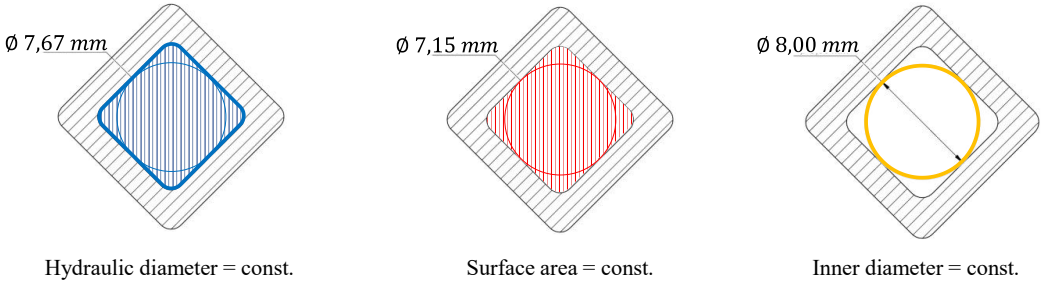


Figure 1: Different comparison criteria for the non-circular geometries

1.3. Channel Geometries for AM

Circular channels exceeding an inner diameter of 8 mm cannot be produced by LPBF without using inner supporting structures [32]. To solve this problem, the shape of the channel is being changed because greater diameter is often being used in hydraulic applications to reduce average flow speeds. Schmelzle et al. applied a diamond-shaped channel-section and were able to reduce weight by 60%. They compared a droplet shape to a diamond shaped channel and found that symmetrical channel geometries are the most material efficient compared to unsymmetrical shapes (like a drop) [10]. Alshare et al. reduced pressure losses by 21% and weight by 84% [11]. DIN EN ISO/ASTM 52911-1 suggests a droplet shape as an alternative to circular but does not standardize this shape for hydraulic applications [32]. In this proposal, a 90° rectangular roof is positioned on three fourths of a circle, where the overhang angle exceeds 45°, as can be seen in **Figure 2**.

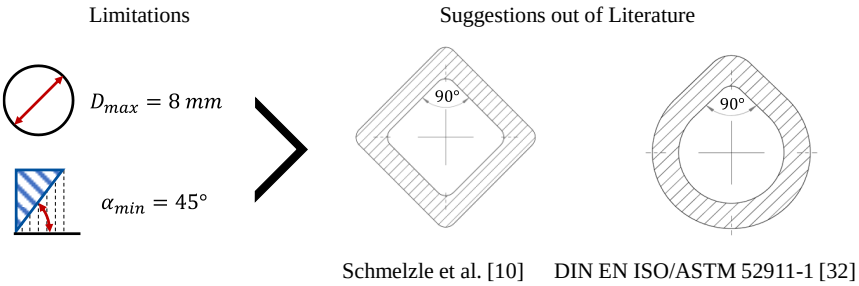


Figure 2: Limitations and suggestions for channel design

1.4. Topology of additively manufactured components

In addition to the advantages regarding flow optimization, the AM offers great potential regarding weight savings, especially in the application of hydraulic circuits. After designing the inner contour, the outer contour must be determined for the specimens. In the simplest case, a geometry with estimated constant wall thickness is evaluation by FEM to meet load requirements. However, the aim is also to optimize the wall thickness depending on the load. For the manufacturing process as such, the wall thickness required by the design is particularly relevant for the duct geometries. In contrast to the case of a round pipe, the modified duct geometries have an inhomogeneous stress distribution at constant pressure load. The stress distribution is the starting point for the design measures to avoid failure. In the process of topology optimization, the reduction approach is followed. The initial geometry of the component is therefore over dimensioned. The stress distribution is determined in the FEM with the meshed model. Iteratively, model elements with uncritical stress values are eliminated. This results in successively reduced nominal geometry [33]. For the special application case regarding additively manufactured components, the anisotropy of the material is also of great importance. The direction of manufacturing affects the subsequent expansion behavior of the channel geometries [34]. However, due to the rectified component alignment in the manufacturing process, this aspect is not initially considered here and compensated for by the safety factor.

2. METHODS

To compare different cross-sectional areas, four different shapes are selected and first optimized simulatively in terms of flow and strength. Three test specimens were manufactured from each geometry using LPBF with 316L. Thus, a total of twelve test specimens were printed, of which four remained untreated, four were post-treated by the HIP method, and four were post-treated by the AFM method. In subsequent bench tests, the pressure difference generated by the test specimens was measured and compared with the simulated results.

2.1. Design Process

In this work, three alternatives to a circular shape were investigated. A triangular and a quadratic shape are investigated because of their symmetrical shape, which are the most material efficient due to a homogeneous distribution of stress [10]. A symmetrical pentagonal or hexagonal shape cannot be printed in a horizontal build orientation because of the overhang angle at the top exceeding the maximal printable angle of 45° [35, 36]. As a third concept, a droplet shape, is compared to a circular shape because the DIN EN ISO/ASTM 52911-1 suggests this form as an alternative to a circular channel geometry [32]. The sharp edges occurring in the alternative shapes are rounded off, to prevent tension maxima.

A round cross-section with a diameter of 8 mm is chosen as the object of comparison. A diameter of 8 mm is the largest possible round cross-section for additive manufacturing in the horizontal mounting direction. The constant inner diameter enables a connection between the investigated and the circular shape without cross-section narrowing for the run-in flow profile, which also enables easier machining. The second concept is based on a constant hydraulic diameter of 8 mm, which is calculated out of the inner surface and circumference. The inner diameter decreases compared to concept one, by equalizing a hydraulic diameter of 7,67 mm. This concept is most logical from a theoretical perspective because this formula is most used to calculate divergent shapes of circular. In the third and last concept, the average speed of flow is constant, by equalizing the inner surface to a circular channel. Hereby, the inner diameter further reduces in size compared to concept one and two. To influence the run-in flow as little as possible, the first concept with a constant inner diameter is chosen as the comparison criterion in this work.

The connection between the investigated channel geometries and the circular connection to the hydraulic system is being constructed by a straight phase at an angle of 45 degrees perpendicular to the building platform. The occurring sharp edges are smoothened out with a radius of 1 mm to reduce stress maxima and improve flow efficiency. These concepts of dimensioning are combined with the connection and are constructed into test parts, which are being investigated in a test rig, to validate the simulated results of the CFD- and FEM-simulations. These results should bring insight into an optimal channel geometry for hydraulic applications, considering mechanical and flow parameters. The geometry of the transition and the shape of the square specimen are shown in **Figure 3**.

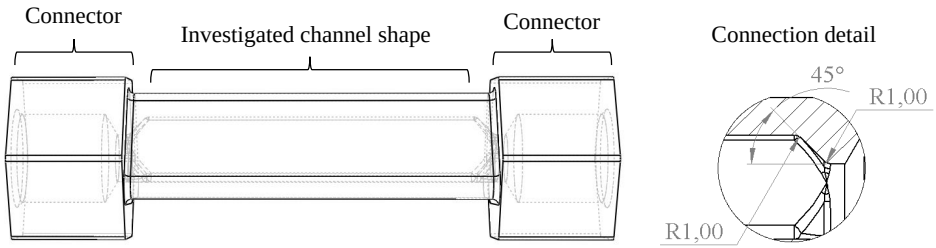


Figure 3: Transition from round cross-section to self-supporting channel geometry

Four test parts were produced three times repetitively to post-process differently (no post-processing, HIP, AFM). The four test parts consisted of the same connector at both ends and the connection between the circular and investigated channel geometry. In the middle of the test part, a circular, triangular, rectangular, and droplet-shaped channel geometry was constructed into the final test parts. CFD- and FEM-results were taken into review to select an optimal concept for dimensioning the inner diameter of the middle section.

2.2. CFD-Simulation

Computational Fluid Dynamics simulations (CFD) were applied to bring insight into the flow characteristics of the test parts. The pressure loss and streamlines were mainly investigated and optimized. In this work, Ansys Fluent software was used to bring these insights. The mesh was constructed with an element size of 0,48 mm, an inflation layer of 0,4 mm and a growth rate of 1,2 divided into 16 layers. These parameters are based on experiences and on Zardin et al. [2]. They found that a ratio of element size to diameter should not exceed 0,06 to ensure realistic results. An ISO VG 46 mineral oil was used in the test setup and modelled in the simulations, which is often used in hydraulic applications. To simulate the oil flowing through the test part, an inlet, an outlet, and a wall are defined. A fully developed flow is specified at the inlet with a velocity inlet. The simulated

corresponds with the analytical profile. At the outlet, a pressure outlet is specified at a static pressure of 150 bar, which corresponds to the operating point. A turbulent intensity of 6.78% was specified at the in- and outlet, as can be calculated out of formula (2). A no-slip condition and an average roughness height of $15 \cdot 10^{-6}$ m were specified to model the rough surface of the as-built test parts. The numerical coupled scheme with a Green-Gauss cell-based gradient and a second order discretization were used to ensure realistic results, with a residuum of 10^{-6} that acted as the criterion of convergence. The SST-k-omega model combined with the Low Reynolds Number Correction was used, after comparing the values of pressure losses of different models (like k-epsilon and laminar) to the analytical pressure loss of a pipe. These settings were used to simulate all concepts and gain insight into the flow characteristics of the test parts [37].

$$I = 0{,}16 \cdot Re^{-1/8} \tag{2}$$

The influence of the post-processing methods is not considered in the simulation, but the wall roughness is only adapted to literature values for the surfaces of additively manufactured components. Since the flow condition in the components is laminar at the operating point of 150 bar and 16,5 l/min (operating point for the comparison), the roughness theoretically does not have a large influence on the pressure drop. The consideration of the generated surface values will be part of further research.

2.3. FEM-Simulation

A Finite Element Method simulation (FEM) is used to determine the wall thicknesses of the channel geometries. In this publication, the outer diameter and corner radius were varied in 80 combinations. An optimization, for findings the lightest combination, that also exceeds the safety margins was conducted. The safety factor of 1,7 is based on DIN 2413 [38], which is applied to calculate wall thicknesses of pipes in hydraulic applications. The results of the design study have been validated in Ansys Mechanical, which confirmed the choice of varying the outer diameter and corner radius. **Figure 4** shows that maximum tensions according to Von Mises occur in these spots. The wall thickness increases from the droplet to the squared and triangle shaped sections. This, of course, increases the weight and outer circumference of the channel sections and increases building time and costs, which represents a disadvantage.

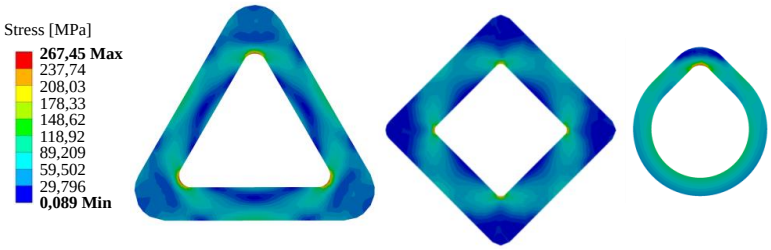


Figure 4: Stress within the components

Based on the results of the FEM, a topology optimization is carried out. Since the FEM already provides a resilient geometry with constant wall thickness from manually run through steps, this is taken as the starting point. An estimation on a round channel using Barlow's formula gives an alternative reference value to reasonably choose the initial value of the wall thickness. Especially when new geometries are to be developed that have not yet passed through any elaborate FEM loops. This ensures that the solution converges reliably. Due to the inhomogeneous stress distribution in the drop-shaped channel geometry, the geometry is divided into several sections, for each of them the meshing is done independently. Especially in the area of the tip of the drop geometry, a considerable mesh refinement is carried out. A study carried out here shows that an element size of 20% of the

initial wall thickness should not be exceeded and at the same time, the limit of the absolute size of maximum 0,1 mm edge length of the elements should be kept. Preferably, a homogeneous hexahedral meshing within the individual sectors is used. This leads to valid results and reflects the discretization of the wall thickness in a comprehensible way. The target value is the inner contour of the pipe geometry. Therefore, it is fixed as a conservation variable in the bearing. Consequently, the front side is also fixed as a fixed bearing. The original wall thickness is preserved and leads to a load-bearing structure.

2.4. Post-processing parameters

After fabrication, the specimens were sawed off the build plate and all external support structures were removed. Four of the test specimens were left as-built and only machined on the functional surfaces so that they could be bolted to the test rig.

One third of the specimens were reworked using the HIP process. In this process, the specimens were subjected to a temperature of 1150°C and a pressure of 1000 bar for 4 hours. During the process, the specimens are surrounded by argon. The final operating point was ramped and cooling and pressure reduction occurred slowly over a period of 4 hours.

Another third of the specimens, each with one of the different geometries, was subjected to grinding post-treatment. The AFM method was used for this purpose. The grain material of the medium used was aluminum zirconia with a grain size of mesh 24/36. The viscosity was classified to 2227,72 Pa·s, measured at 1 rad/s. Flow grinding was performed for 24 minutes at 65 bar.

2.5. Test Rig

On the test rig used for the experimental investigation, the test specimens can be screwed in between two adapters. **Figure 5** shows the hydraulic circuit diagram of the test rig. The pressure is adjusted to 150 bar via the proportional valve (1). The cooler (2) is temperature controlled and regulates the operating temperature to 40°C. The flow rate is set to 16,5 l/min via the pump (3). A differential pressure sensor (4) measures the pressure drop across the specimen. After reaching steady state, for all specimens the pressure loss is recorded.

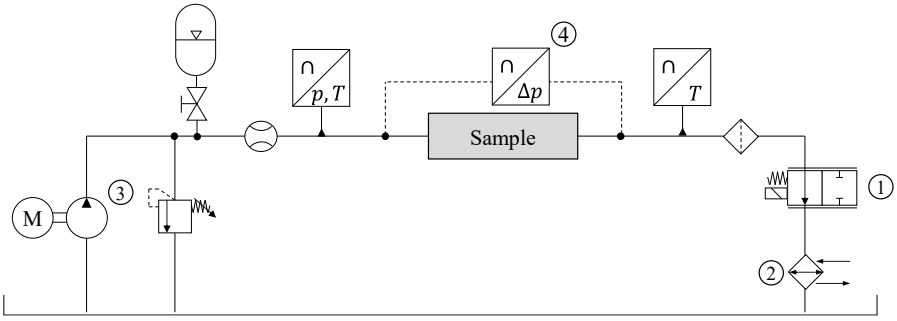


Figure 5: Hydraulic circuit diagram of the test bench

3. RESULTS

With the help of the CFD and FEM results, the final shapes of the components have been designed. The results of the simulations were particularly decisive for the selection of the shape, the course of the cross-section transition and the choice of wall thickness. The strength of the components was confirmed on a test rig by a static compression test at 165 bar (10% above the operating pressure). All components exhibited the necessary strength.

3.1. CFD-Results

The simulated pressure losses of the different geometries are listed in the second column of **Table 1**. In addition, **Figure 6** shows the flow lines within the channels for the triangle, square, and drop.

Table 1: Simulation results of the specimens

Form	Pressure loss Δp [bar]	Cross-sectional area A [mm²]	Wall thickness t [mm]
Circle	0,17741	50,625	0,3
Square	0,15076	63,142	1,5
Triangle	0,14643	81,084	2,75
Drop	0,16920	53,479	1,25

The flares at the inlet and outlet of the components are caused by the thread runout of the fittings. Since only the inside is shown, the 8 mm inner diameter of the screw-in fitting can be seen at the inlet. However, since the hydraulic fitting is not flush with the end of the bore, the cross-sectional expansion occurs. Dead volumes are created at these points. Additional dead volumes occur at the transition from the round 8 mm cross-section to the self-supporting channel contour. Due to the larger cross-sectional area, this effect is particularly strong in the triangle.

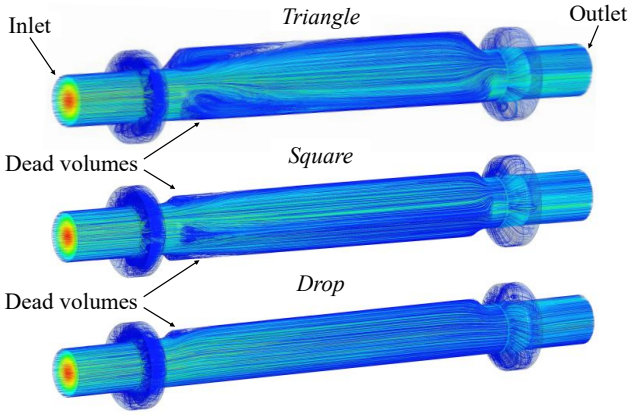


Figure 6: Simulation of the flow lines in the non-circular cross-sectional shapes

3.2. FEM-Results

With the aid of FEM simulation, a suitable wall thickness was developed for each geometry, at which the maximum stress in the component is not exceeded, considering the safety factor. These wall thicknesses, also listed in Table 1, differ for the various shapes. The highest wall thickness is required for the triangle since significant stress peaks form at the corners. The most suitable shape in terms of stress is the circle, which requires a wall thickness of only 0,3 mm. However, such thin walls cannot be printed with the selected manufacturing process, which is why the wall thickness was increased to 2 mm for this component. This corresponds to the wall thickness of a standard 12S tube. In order to make a prediction about possible weight savings, a topology optimization was carried out in this work. The resulting geometries will also be used for future static compression tests. The right side in **Figure 7** shows an example of the reduced geometry of a drop shape because of topology optimization without a safety factor. The stress state in the detail of the tip is shown opposite. From this, the material accumulation in this geometry section is derived.

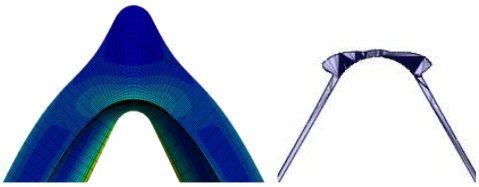


Figure 7: Topology optimization upon of drop shape

Due to the inhomogeneous wall thickness over the circumference, only a rudimentary estimation is possible at first. As expected, there is an accumulation of material at the tip of the drop. In order to be able to give a holistic assessment, a comparative value to the original geometry is used. The cross-sectional area of the resulting wall is determined and compared with the result determined in the same way for a channel with a round cross-section. This results in a new average wall thickness. Compared to the round duct, this results in a weight reduction of 16%.

3.3. Experimental results

The results of the experimental (filled markers) and simulative (unfilled markers) investigation are shown in **Figure 8**. For each component geometry, the pressure drop is plotted for the untreated, the thermally post treated and the AFM-treated component. Differences between the various geometries can be seen. The post-treatment processes also lead to a change (mostly a reduction) in the pressure difference for all components except the triangle.

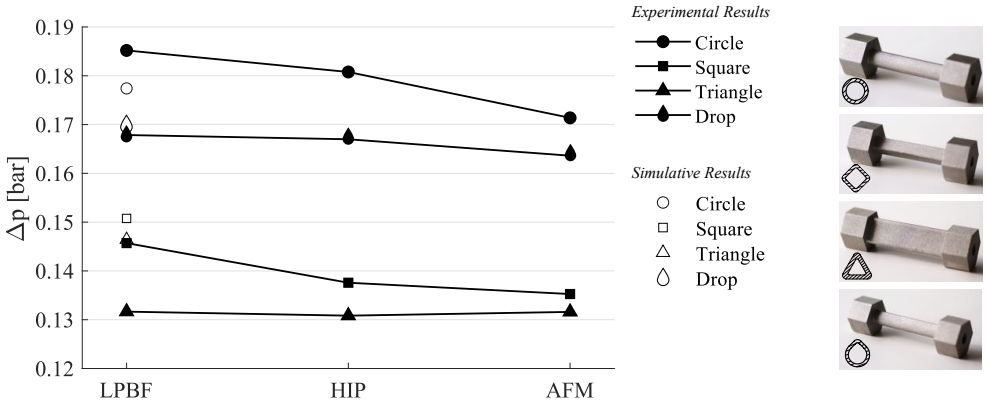


Figure 8: Measured pressure loss of different geometries and post-processing methods

The designed components could all be built horizontally without support structures in the channels and no leakage occurred in the hydraulic application. A static pressure test proofed, that the specimens can withstand a pressure of 150 bar and even 165 bar did not cause any leakage. In the round shape, a clearly discernible difference in quality occurred on the downward-facing surface. Here, there was a collapse of the shape and a significantly increased roughness. This confirms that the use of the round cross-section is not recommended for diameters above 8 mm.

The HIP process only has an influence on the pressure difference for the round and square shapes. The AFM method shows more effect on pressure loss reduction. Except for triangle shapes, the pressure difference generated by the component decreases for all shapes compared to the unprocessed parts. A connection to the flow lines in Figure 6 can be recognized. In the case of the triangle, a particularly large dead volume occurs and, due to the larger cross-sectional area, the velocity inside the component decreases under constant volume flow. Accordingly, the grind effect of the abrasive

medium is lower here. In the round shape, the flow is fastest and without dead volume, so that the greatest effect can be achieved. It should be noted that the AFM process was applied before the machining post-processing, so that the entry channel into the component was still at 8 mm and there was no widening as in Figure 6 at the transition to the round fitting-geometry. In the case of the non-circular geometries, this resulted in the highest material removal occurring precisely in this area, which was later partially widened.

4. DISCUSSION

The experimental results confirm the simulation because the rank order of the geometries was predicted correctly. Comparing the different geometries, a round geometry causes the highest pressure loss. However, this must be evaluated considering the comparison criterion of constant inner diameter. Consequently, the cross-sectional area also has to be taken into account. The constant inner diameter leads to a higher cross-sectional area for non-circular shapes. The pressure loss corresponds with the cross-sectional area. However, the relationship is not proportional. Accordingly, a constant cross-sectional area would not have led to constant pressure differences. The constant inner circle as a comparison criterion is more suitable for obtaining a run-in flow profile and is therefore appropriate for comparison here. In terms of post-processing, the trends are the same for almost all components. Since the HIP process is not usually used for external post-processing, but to change the material properties, there is no significant change in the pressure loss in the samples. However, since very high temperatures occur during the process, it is possible that individual splashes or powder residues are melted and change their surface properties.

5. CONCLUSION AND OUTLOOK

In this paper, the fundamentals for the use of LPBF in hydraulic applications were determined. Starting with an iterative CFD simulation, a transition from the connection geometry to a self-supporting shape was designed. FEM was then used to determine a suitable wall thickness for the various geometries. In bench tests, all components have proven their strength and leak proofness. In addition, the measured values showed that the CFD simulation was correctly designed, and the predicted sequence of shapes was correct. Thus, the test rig and simulation settings can be used for further studies on additively manufactured components. Whether the constant inner diameter is the correct comparison criterion depends on the application. If the flow is to be kept as constant as possible without turbulence, then it is a suitable criterion. If the installation space is limited, a constant cross-section would be more suitable. Another approach would be to use a constant amount of material, but this would result in much smaller cross-sections for the non-circular shapes, since these require greater wall thicknesses. The HIP and AFM methods have demonstrated their effectiveness in reducing pressure loss. However, both have only been tested in a single parameter combination and can be further refined in targeted experiments. In particular, widening of the channel should be avoided for grinding.

The aim of further research is to investigate more complex channel courses. With the knowledge gained from the linear test specimens, further investigations of three-dimensional channel courses can be carried out. In this context, a general methodology should be developed that offers solutions for various requirements. For non-round shapes, for example, the orientation of the shape in curves will become relevant. In addition, the roughness of the components should be determined and integrated into the simulation.

ACKNOWLEDGEMENTS

This research and development project is funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) within the Technologietransfer-Programm Leichtbau (TTP LB) and supervised by Projektträger Jülich.

NOMENCLATURE

A	Cross-sectional area	[m²]
AFM	Abrasive Flow Machining	
AM	Additive Manufacturing	
CFD	Computational Fluid Dynamics	
DfAM	Design for Additive Manufacturing	
D_{hyd}	Hydraulic diameter	[m]
FEM	Finite element Method	
HIP	Hot Isostatic Pressing	
LPBF	Laser Powder Bed Fudion	
Δp	Pressure difference	[bar]
U	Area perimeter	[m]

REFERENCES

[1] Chambon, R., Tollenaere, M. (1991) Automated Al-based mechanical design of hydraulic manifold blocks. Computer-Aided Design:213–222

[2] Zardin B, Cillo G, Rinaldini C et al. (2017) Pressure Losses in Hydraulic Manifolds. Energies 10:310. https://doi.org/10.3390/en10030310

[3] Zhang J, Liu G, Ding R et al. (2019) 3D Printing for Energy-Saving: Evidence from Hydraulic Manifolds Design. Energies 12:2462. https://doi.org/10.3390/en12132462

[4] Semini C, Goldsmith J, Manfredi D et al. (2015) Additive manufacturing for agile legged robots with hydraulic actuation. In: 2015 International Conference on Advanced Robotics (ICAR), pp 123–129

[5] Pietropaoli M, Ahlfeld R, Montomoli F et al. (2017) Design for Additive Manufacturing: Internal Channel Optimization. Journal of Engineering for Gas Turbines and Power 139. https://doi.org/10.1115/1.4036358

[6] Diegel O, Schutte J, Ferreira A et al. (2020) Design for additive manufacturing process for a lightweight hydraulic manifold. Additive Manufacturing 36:101446. https://doi.org/10.1016/j.addma.2020.101446

[7] Rekadze P, Rodionov L, Satsyuk I (2018) Analysis of hydraulic units manufactured by powder bed fusion. In: 2018 Global Fluid Power Society PhD Symposium (GFPS)

[8] Zhang C, Wang S, Li J et al. (2020) Additive manufacturing of products with functional fluid channels: A review. Additive Manufacturing 36:101490. https://doi.org/10.1016/j.addma.2020.101490

[9] Chekurov S, Lantela T (2017) Selective Laser Melted Digital Hydraulic Valve System. 3D Printing and Additive Manufacturing 4:215–221. https://doi.org/10.1089/3dp.2017.0014

[10] Schmelzle J, Kline EV, Dickman CJ et al. (2015) (Re)Designing for Part Consolidation: Understanding the Challenges of Metal Additive Manufacturing. Journal of Mechanical Design 137. https://doi.org/10.1115/1.4031156

[11] Alshare AA, Calzone F, Muzzupappa M (2019) Hydraulic manifold design via additive manufacturing optimized with CFD and fluid-structure interaction simulations. RPJ 25:1516–1524. https://doi.org/10.1108/RPJ-03-2018-0064

[12] Geating JT, Wiese MC, Osborn MF (2017) Design, fabrication, and qualification of a 3d printed metal quadruped body: combination hydraulic manifold, structure and mechanical interface. 2017 International Solid Freeform Fabrication Symposium

[13] Zhu Y, Wang S, Zhang C et al. (2020) AM-driven design of hydraulic manifolds: enhancing fluid flow and reducing weight. In: 12th International Fluid Power Conference, pp 155–159

[14] Biedermann M, Beutler P, Meboldt M (2021) Automated design of additive manufactured flow components with consideration of overhang constraint. Additive Manufacturing 46:102119. https://doi.org/10.1016/j.addma.2021.102119

[15] Kumke M (2018) Methodisches Konstruieren von additiv gefertigten Bauteilen. Dissertation, Technische Universität Braunschweig

[16] Rolinck N, Schmitt M, Schneck M et al. (2021) Development Workflow for Manifolds and Fluid Components Based on Laser Powder Bed Fusion. Applied Sciences 11:7335. https://doi.org/10.3390/app11167335

[17] Godec D, Gonzalez-Gutierrez J, Nordin A et al. (2022) A Guide to Additive Manufacturing. Springer International Publishing, Cham

[18] Laverne F, Segonds F, Anwer N et al. (2015) Assembly Based Methods to Support Product Innovation in Design for Additive Manufacturing: An Exploratory Case Study. Journal of Mechanical Design 137. https://doi.org/10.1115/1.4031589

[19] Alfaify A, Saleh M, Abdullah FM et al. (2020) Design for Additive Manufacturing: A Systematic Review. Sustainability 12:7936. https://doi.org/10.3390/su12197936

[20] Meiners W, Wissenbach K, Gasser A (1996) Verfahren zur Herstellung eines Formkörpers

[21] Kumar S (2022) Additive Manufacturing Solutions. Springer International Publishing AG, Cham

[22] Haghdadi N, Laleh M, Moyle M et al. (2021) Additive manufacturing of steels: a review of achievements and challenges. J Mater Sci 56:64–107. https://doi.org/10.1007/s10853-020-05109-0

[23] Maier HJ, Niendorf T, Bürgel R (2019) Handbuch Hochtemperatur-Werkstofftechnik: Grundlagen, Werkstoffbeanspruchungen, Hochtemperaturlegierungen und -beschichtungen, 6., überarbeitete und erweiterte Auflage. Springer Vieweg, Wiesbaden, Heidelberg

[24] Liverani E, Lutey AHA, Ascari A et al. (2020) The effects of hot isostatic pressing (HIP) and solubilization heat treatment on the density, mechanical properties, and microstructure of austenitic stainless steel parts produced by selective laser melting (SLM). Int J Adv Manuf Technol 107:109–122. https://doi.org/10.1007/s00170-020-05072-9

[25] Grech IS, Sullivan JH, Lancaster RJ et al. (2022) The optimisation of hot isostatic pressing treatments for enhanced mechanical and corrosion performance of stainless steel 316L produced by laser powder bed fusion. Additive Manufacturing 58:103072. https://doi.org/10.1016/j.addma.2022.103072

[26] Gillespie LK (1999) Deburring and edge finishing handbook. Society of Manufacturing Engineers; American Society of Mechanical Engineers, Dearborn, Mich., Fairfield, NJ

[27] Matthiesen G, Merget D, Pietrzyk T et al. (2020) Design and experimental investigation of an additive manufactured compact drive. In: 12th International Fluid Power Conference, pp 137–147

[28] Kumar SS, Hiremath SS (2016) A Review on Abrasive Flow Machining (AFM). Procedia Technology 25:1297–1304. https://doi.org/10.1016/j.protcy.2016.08.224

[29] Duval-Chaneac MS, Han S, Claudin C et al. (2018) Characterization of maraging steel 300 internal surface created by selective laser melting (SLM) after abrasive flow machining (AFM). Procedia CIRP 77:359–362. https://doi.org/10.1016/j.procir.2018.09.035

[30] Rösler Oberflächentechnik (2020) Innenkanäle effizient und automatisiert bearbeiten. J Oberfl Techn 60:54–57. https://doi.org/10.1007/s35144-020-0999-z

[31] Yu C-H, Leicht A, Peng RL et al. (2021) Low cycle fatigue of additively manufactured thin-walled stainless steel 316L. Materials Science and Engineering: A 821:141598. https://doi.org/10.1016/j.msea.2021.141598

[32] DIN Deutsches Institut für Normung e.V. (2019) Additive Fertigung - Konstruktion - Teil 1: Laserbasierte Pulverbettfusion von Metallen: DE_307153562(DIN EN ISO/ASTM 52911-1)

[33] Klein B (2003) FEM: Grundlagen und Anwendungen der Finite-Elemente-Methode, 5. Auflage. Vieweg+Teubner Verlag, Wiesbaden

[34] Qu S, Ding J, Fu J et al. (2022) Anisotropic material properties of pure copper with fine-grained microstructure fabricated by laser powder bed fusion process. Additive Manufacturing 59:103082. https://doi.org/10.1016/j.addma.2022.103082

[35] Thomas D (2009) The Development of Design Rules for Selective Laser Melting. Dissertation, University of Wales Institute, Cardiff

[36] Gibson I, Rosen D, Stucker B et al. (2021) Additive Manufacturing Technologies: Third Edition. SPRINGER NATURE, [S.l.]

[37] ANSYS (2023) Ansys Resource Center | Webinars, White Papers and Articles. https://www.ansys.com/resource-center#t=ResourceCenterTab&sort=relevancy&numberOfResults=50. Accessed 24 Apr 2023

[38] DIN Deutsches Institut für Normung e.V. (2020) Nahtlose Stahlrohre für öl- und wasserhydraulische Anlagen - Berechnungsgrundlage für Rohre und Rohrbögen bei schwellender Beanspruchung. Accessed 2020

BRONZE CLADDING FOR BIMETAL PARTS PRODUCED BY LASER DEPOSITION BRAZING

Hannes Freisse[1]*, Emilie Isère[1], Timo Schudeleit[2], Dominik Keller[2], Francesco Rippa[2]

[1] *Kugler Bimetal SA, Chemin du Château Bloch 17, CH-1219 Le Lignon / Geneva (Switzerland)*
[2] *inspire AG, Technoparkstrasse 1, CH-8005 Zurich (Switzerland)*

* Corresponding author: Tel.: +41 78 230 55 68; E-mail address: freisse@bimetal.ch

ABSTRACT

Bronze has emerged as a highly functional material for various tribological applications, owing to its diverse alloys that offer adaptable material properties for a wide range of loads. Bronze can be used as a solid material or applied as a functional layer in bi-metal components, offering a combination of the superior mechanical properties of the steel base body with the excellent tribological properties of the bronze alloy. Conventionally, bi-metal bearings are casted, a labour-intensive and material-inefficient process. Laser deposition brazing offers a potential solution to these disadvantages. However, it faces the challenge of ensuring sufficiently high layer bonding strength between the steel base body and the bronze alloy layer. This paper presents the production of bi-metal bearings through laser deposition brazing and material analysis, focusing on the development of a reliable process that ensures high-quality bonding between the steel and bronze layers.

Keywords: Laser Cladding, Bronze

1. STATE OF THE ART

There are seven different Laser Surface Processing (LSP) techniques, divided into three primary categories depending on the intensity (the quantity of energy per unit area applied to the metal surface by the laser beam) and the modification of the metal's aggregate state: heating, melting, and vaporization (refer to **Figure 1**). Low intensity leads to surface heating, exemplified by heat treatment for hardening purposes. With increasing intensity, metal melting occurs, as seen in welding, while higher intensity can directly vaporize metal, as evidenced in the cutting of metals [1].

Further categorization is based on whether the goal is to influence the micro-geometry of the surface or alter the material properties. These procedures are carried out by the addition of material, with distinctions drawn regarding whether the added material and/or substrate are melted. The added material can take various forms, such as wire, tape, paste, or powder. The following section specifically elaborates on processes involving powdered added material.

When aiming to alter material properties without introducing additional substances, two distinct procedures come into play: laser hardening and laser remelting. Laser hardening involves heating the surface to a temperature below the melting point [2], whereas laser remelting elevates the surface temperature beyond the melting point [3].

The micro-geometry of the surface can undergo modification through processes such as melting the roughness peaks, referred to as laser polishing [4], or by vaporizing and removing material through a technique known as laser ablation [5]. Both laser polishing and laser remelting are material-free processes involving surface melting. However, they differ in that laser polishing selectively melts

only the roughness peaks to alter micro-geometry, whereas laser remelting profoundly remelts the substrate surface, aiming to fundamentally influence material properties.

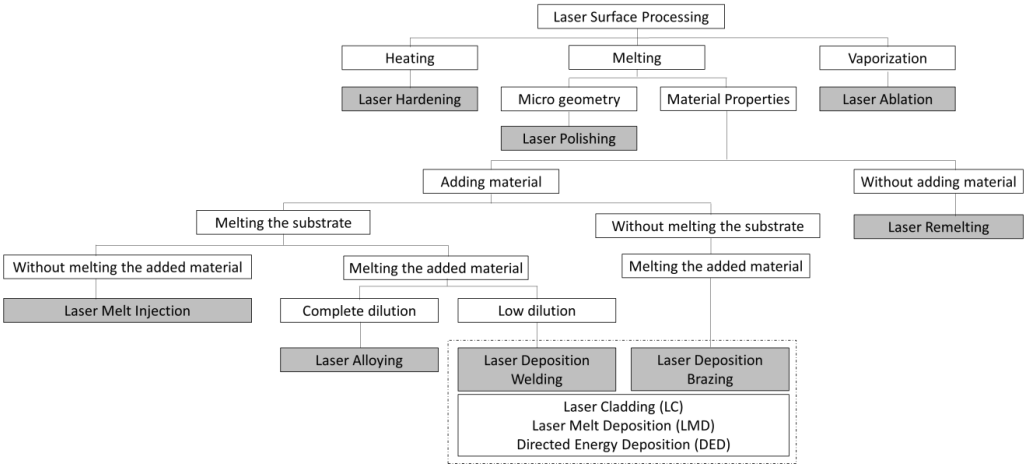


Figure 1: Overview and classification of Laser Surface Processes

In LSP involving the addition of material, four distinct procedures are identified: Laser Melt Injection (LMI), Laser Alloying (LA), Laser Deposition Welding (LDW), and Laser Deposition Brazing (LDB). LMI entails melting the surface without affecting the added material. Typically, ceramic particles are injected into the molten surface to create a wear-resistant metal-matrix-composite (MMC) [6]. LA takes place when metallic powder is used, completely melting, and blending with the molten substrate, resulting in the formation of a new alloy at the part surface with specific properties [7] (refer to **Figure 2**). In LDW, more energy is directed into the powder and less into the substrate compared to the LA process [8]. In LDW, both the powder and substrate have comparable melting temperatures, causing the substrate to melt along with the powder. If the powder material has a significantly lower melting temperature, more laser power is directed into the powder material and less into the substrate material, known as LDB [9]. There is no dilution zone anymore, as seen in LDW.

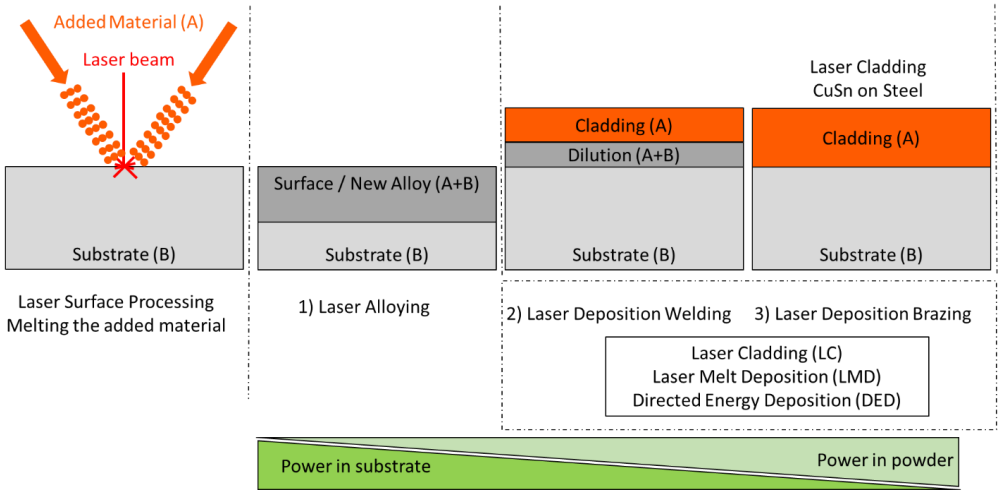


Figure 2: Difference between Laser Alloying, Laser Deposition Welding and Laser Deposition Brazing

The distinction between cladding by welding (LDW) and cladding by brazing (LDB) is highlighted in the established German standard for the classification of industrial production technologies [10]. Despite these differences, both LDW and LDB are commonly associated with terms like Laser Cladding (LC), Laser Metal Deposition (LMD), or Directed Energy Deposition (DED) [11]. All four processes - LMI, LA, LDW, and LDB - can be executed using the same technical setup for delivering powder material into the process zone and injecting it into the laser beam and/or melt pool generated by the laser beam. This paper focuses on LDB, and the subsequent chapter offers a more detailed explanation based on this specific process.

LDB processes can be classified into soft and hard brazing categories [12]. When the melting temperature is below 450°C, the process is referred to as soft brazing. This also includes cladding with white metal, a term denoting tin-based alloys with additions of antimony and lead, commonly employed for various tribological applications [13]. On the contrary, if the brazing alloy has a melting temperature above 450°C, the processes are labelled as hard brazing. Instances of hard brazing include coatings with alloys like copper-tin (CuSn).

Tin bronze typically contains tin (Sn) within the range of 5% to 12%, and a higher tin content is associated with increased hardness. This allows for the customization of material properties to meet specific requirements [14]. CuSn alloys are utilized as a functional surface in tribology applications such as turbochargers, hydraulic pumps, and plain bearings (e.g., in gearbox systems) [15].

The sliding wear behaviour of a material is closely dependent on its microstructural properties. To enhance tribological properties, it is vital for a bearing material to possess three key microstructural constituents: a ductile phase, a load-bearing constituent, and a lubricating element. Bronze with lead (Pb) satisfies these essential conditions for being an effective bearing material. The presence of a ductile phase, specifically in the form of the α-phase (a copper-rich solid solution of tin), provides support to the load-bearing constituent, represented by the Cu-Sn intermetallic phase. Additionally, the effective inclusion of solid lubricant, dispersed as insoluble lead (Pb) particles in the microstructure, contributes to reduced wear and friction coefficients, particularly in dry sliding conditions [16].

However, given that lead is identified as a heavy toxic metal, several regulations, including directives like REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) by the European Union, have been enforced to limit its usage in various mechanical components. These measures aim to prevent the exposure of Pb-containing components to both human and ecological environments [17]. Consequently, there is an urgent requirement to develop and evaluate new Pb-free alternatives, such as bismuth bronze alloys, for bearing applications [18].

2. EXPERIMENTAL PROCEDURE

The study employed an industrial multifunctional machine with multiple axes, allowing for the cladding of both smaller components and larger parts with dimensions of up to 1000mm in diameter, 1100mm in length, and a weight of up to 1000kg. A Yaskawa robot guided the laser head. The multifunctional machine featured two distinct laser heads: one for externally cladding parts and another for internally cladding tubes with a diameter >68 mm. A Laserline GmbH diode laser of the LDM 8000-100 model, boasting a maximum power of 8kW, was utilized. The laser beam's wavelength ranged from 900nm to 1080nm, and it exhibited a laser beam quality of 100mm*mrad.

Two copper-tin alloys, namely CuSn11Bi3 (Tokat300) and CuSn12Ni2 (Tokat325), were employed in powder form, with particle diameters smaller than 150µm. A powder feeding machine of the

PF22H type from GTV Verschleissschutz GmbH delivered the powder to the laser head. Substrate materials included the steel alloys C45 and 42CrMo4. The investigations focused on disk-shaped specimens measuring 114mm in diameter and 18mm in thickness. Each test was repeated three times for both materials and substrate materials, using identical parameters to facilitate statistical analysis of values and standard deviations. Deposition onto the substrates occurred in a spiral pattern, moving from the inside to the outside. Subsequently, specimens were machined to a thickness of 1.3mm using a turning machine.

On the planar machined surface, a penetration test was conducted following ISO 4386-3. The visible inspection penetrant, type VP-30 from Met-L-Chek, was applied, followed by removal after a 10-minute waiting period. Subsequently, a penetrant inspection developer, D-70 from Met-L-Chek, was applied, and after a 2-minute waiting period, the parts were subjected to visual inspection. To examine the bonding between the coating and the substrate across the entire coating area, an ultrasonic test was carried out following ISO 4386-1. The ultrasonic test utilized the AMS 2145E device from Olympus.

3. RESULTS AND DISCUSSION

Figure 3a) depicts a C45 steel sample laser cladded with Tokat300 bronze (CuSn11Bi3), emphasizing the spiral coating path. After laser processing, the bronze displays slight oxidation, leading to a matte appearance on the surface. Subsequently, the surface underwent machining to remove the oxidized layer, resulting in a smooth functional surface. In **Figure 3b**), a representation of the dye penetrant test result is provided. The test reveals no indications, confirming the absence of any open surface imperfections.

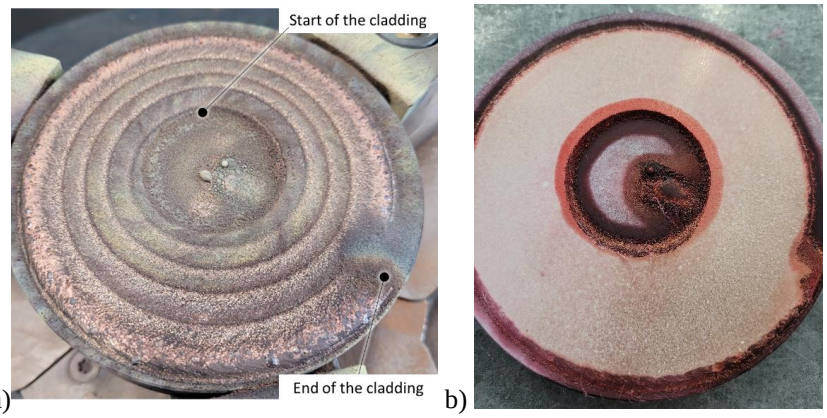


Figure 3: Top view of the specimen a) after the laser cladding and b) after machining and penetration test

The results obtained from the repetition of all four material combinations three times demonstrated consistently high quality and repeatability. All ultrasonic tests consistently revealed no indications, providing conclusive evidence of a defect-free bonding between the bronze cladding and the steel. Dye penetration tests further affirmed the absence of surface open imperfections, underscoring the overall high quality of the cladding. This quality was further validated through metallographic analysis, as shown in **Figure 4**.

The metallographic image illustrates a remarkably straight interface between the Tokat300 bronze coating and the C45 steel substrate, indicating minimal impact on the substrate during the LDB process. Only a few imperfections in the form of small pores are observed in the bronze cladding. An

assessment of the pore diameters across all detected pores in all twelve specimens revealed a mean value of 30µm ± 22 µm with standard deviation, further emphasizing the precision and quality of the cladding.



Figure 4: Metallographic cross-section of laser cladded CuSn11Bi3 (Tokat300) on C45 steel

Porosity analysis values for all four material combinations are provided in Table 1, including mean values and standard deviations. Notably, all mean values are below 1%. It's important to note that there isn't a specific standard for recommended porosity values in cladding. As an alternative, standards for joint welding (excluding beam welding) can be considered. According to such standards, porosity values between 1% and 5% are recommended, depending on the intended applications [19]. Alternatively, for comparison, the standard for electron and laser beam welded joints provides requirements and recommendations for evaluation groups concerning irregularities, suggesting higher values ranging from 2% to 6% [20].

Table 1: Parameters of the cylinder

	Tokat300 (CuSn11Bi3)	**Tokat325 (CnSn12Ni2)**
Steel 42CrMo4	0.2% ± 0.3%	0.6% ± 0.5%
Steel C45	0.8% ± 0.6%	0.2% ± 0.08%

Figure 5 displays the mean and standard deviations of hardness values measured in the bronze cladding of CuSn11Bi3 (Tokat300) and CuSn12Ni2 (Tokat325) on the C45 steel substrate. The standard deviations are minimal, accounting for less than 5% of the absolute values. The elevated hardness values observed in the steel indicate the thermal influence of the LDB process to a depth of 1mm, identified as the heat-affected zone (HAZ). Notably, the hardness values of CuSn11Bi3 are slightly lower than those of CuSn12Ni2, attributed to the lower tin content in CuSn11Bi3.

When C45 steel undergoes a standard hardening process, it is anticipated to achieve a hardness level of around 600HV. Nevertheless, in the Heat Affected Zone (HAZ) of the steel coated with CuSn, the hardness values remain under 400HV. Comparing this to existing literature on steel clad with a different bronze (CuAl - aluminium bronze), it consistently shows a 1mm deep HAZ. In contrast, CuAl laser cladding results in a more substantial hardening of the steel, reaching HAZ hardness values

of 650HV [8]. This dissimilarity is attributed to the higher melting temperature of aluminium bronze CuAl in comparison to tin bronze CuSn, leading to a more pronounced thermal impact on the steel.

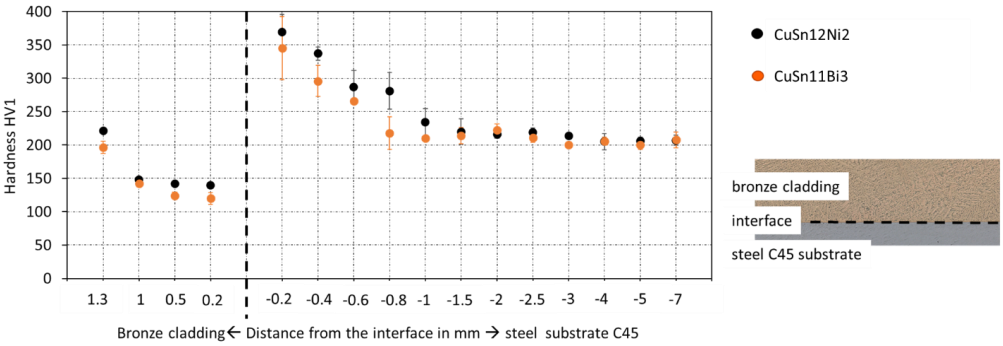


Figure 5: Hardness curve of laser cladded CuSn11Bi3 (Tokat300) on C45 steel

Figure 6 depicts two finalized industrial components ready for use after the machining process. **Figure 6a** depicts a component specifically designed for a planetary gearbox application. This component is made of 42CrMo4 steel and is coated with CuSn11Bi3 (Tokat300). It has dimensions of 220mm in diameter and 260mm in length. **Figure 6b** showcases a component tailored for a power generation application, with a diameter of 345mm and a length of 156mm. The inner surface of this power generation component has undergone coating. To achieve this, the laser head was slightly tilted, enabling the deposition of bronze onto the inner surface. The resulting bronze layer demonstrates a consistent thickness across the entire width of the plain bearing.

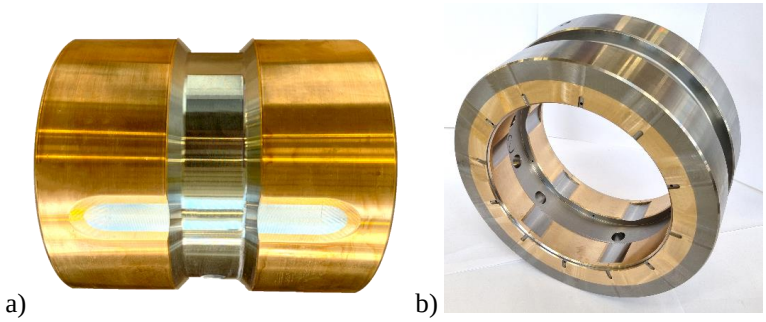


Figure 6: a) Gear box application; b) power generation application

4. SUMMARY

The article delves into the production technology of Laser Deposition Brazing (LDB), as flexible, resource efficient and versatile in alternative to casting for bimetal part production. The process qualification involved testing two distinct bronze alloys on small samples to evaluate and confirm their quality by penetration testing according to ISO 4386-3, a low porosity as well as the absence of cracks in the micrographs. The process was transferred from small samples to production of larger industrial components, confirming the technological suitability of LDB for bimetal components intended for tribological application like in gear boxes or hydraulic pumps. The hardness curves show a typical increase in hardness in the HAZ of the steel substrate. The study shows that Laser Deposition Brazing is a suitable production technology for manufacturing high-quality bimetal components.

REFERENCES

[1] Schaaf, P., 2010, Laser Processing of Materials, 1st ed., Springer, Berlin-Heidelberg (Germany), Chap. 2.

[2] Ruetering, M.A., 2016, „Laser Hardening the known but unknown application“, Laser Technik Journal, Vol. 13, pp. 30-33

[3] Brytan, Z., Bonek, M., Dobrzanski, L.A., Ugues, D., Grande, M.A.G., 2010, “The Laser Surface Re-melting of Austenitic Stainless Steel”, Vol. 654-656, pp. 2511-2514

[4] Richter, B. Blanke, N., Werner, C., Vollertsen, F., 2018, „Effect of Initial Surface Features on Laser Polishing of Co-Cr-Mo Alloy Made by Powder-Bed Fusion”, The Journal of The Minerals (JOM), Metals & Materials Society (TMS), Vol. 71, pp. 912-919

[5] Bliedtner, J., Schindler, Ch., Seiler, M., Wächter, S., Friedrich, M., Giesecke, J., 2016, „Ultrashort Pulse Laser Material Processing“, Laser Technik Journal, Vol. 5, pp. 46-50

[6] Freisse, H., Bohlen, A., Seefeld, T., 2018, „Determination of the particle content in laser melt injected tracks”, Journal of Materials Processing Technology, Vol. 267, pp. 177-185

[7] Yiming, C., Guochao, G., Juijin, Y., Chuanzhong, C., 2018, “Laser surface alloying on aluminum and its alloys: A review”, Optics and Lasers in Engineering, Vol. 100, pp. 23-37

[8] Freisse, H., Langebeck, A., Koehler, H., Seefeld, T., Vollertsen, F., 2016, ”Investigations on dry sliding of laser cladded aluminum bronze“, Vol. 3, pp. 1-10

[9] Ng et al: Laser cladding of copper with molybdenum for wear resistance enhancement in electrical contacts, Applied Surface Science, Volume 253, 2007

[10] DIN 8580:2003-09: Manufacturing processes - Terms and definitions, division

[11] ASTM F3187-16: Standard guide for Directed Energy Deposition of Metals

[12] ISO 857-2:2005: Welding and allied processes – Vocabulary – Part 2: Soldering and brazing processes and related terms

[13] Jeong, J.I., Kim, J.H., Choi, S.G., Cho, Y.T., Kim, C.K., Lee, H., 2021, “Mechanical Properties of White Metal on SCM440 Alloy Steel by Laser Cladding Treatment”, Applied Science, Vol. 11, pp. 1-12

[14] Prasad, B., Patwardhan, A., Yegneswaran, A., 1996, “Factors controlling dry sliding wear behaviour of a leaded tin bronze,” Materials science and technology, Vol. 12, pp. 427–435

[15] Kugler Bimetal, characteristic of alloys, n.d., from “https://www.bimetal.ch/projects/”bronze,”

[16] Prasad, B., 1997, “Dry sliding wear response of some bearing alloys as influenced by the nature of microconstituents and sliding conditions,” Metallurgical and Materials Transactions A, Vol. 28, No. 13, pp. 809–815

[17] European Chemicals Agency, n.d., from “https://echa.europa.eu/de/regulations/reach/understanding-reach”

[18] Oksanen, V.T., Lehtovaara, A.J., Kallio, M.H., 2017, Load capacity of lubricated bismuth bronze bimet-al bearing under elliptical sliding motion, Vol. 388-389, pp. 72-80

[19] ISO5817:2014: Welding - Fusion-welded joints in steel, nickel, titanium and their alloys (beam welding excluded) - Quality levels for imperfections

[20] ISO 13919-1:2019: Electron and laser-beam welded joints - Requirements and recommendations on

quality levels for imperfections - Part 1: Steel, nickel, titanium and their alloys

ON POLYOXYMETHYLENE COMPOSITE FOR SUSTAINABLE HYDRAULIC VALVES

peer reviewed

Ana Trajkovski[1]*, Nejc Novak[1], Mitjan Kalin[1], Franc Majdič[1]

[1]*Laboratory for Fluid Power and Control, Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva cesta 6, 1000 Ljubljana, Slovenia*

* Corresponding author: Tel.: +386 1 4771-413; E-mail address: ana.trajkovski@fs.uni-lj.si

ABSTRACT

The aim of this study was to investigate the tribological properties of polyoxymethylene reinforced with carbon fibres which has proven to give comparable results to high performance polyetheretherketone, regardless their price difference. The results obtained in glycerol-water mixture were compared with the results of tests using conventional hydraulic oil and demineralised water as lubricant. The tests were performed at room and elevated temperature, as expected in hydraulic applications.

The results showed very low coefficient of friction and specific wear rate in glycerol-water mixture, comparable to the values measured in standard hydraulic oil (COF ~ 0.027 - 0.033, specific wear rate ~ 10^{-7} mm^3/Nm), and lower than measured in water (COF ~ 0.14, specific wear rate ~ 10^{-6} mm^3/Nm), at room temperature. At elevated temperature, coefficient of friction slightly increased in hydraulic oil and glycerol-water mixture, but decreased in water, and values become comparable for all lubricants. Specific wear rate significantly increased in hydraulic oil and glycerol-water mixture at elevated temperature, and became comparable to the specific wear rate measured in water. Based on results, both polyoxymethylene composite and glycerol-water mixture can be good alternative for standard hydraulic oil and steel tribo-pairs, leading towards excellent tribological properties.

Keywords: polymer composites, green lubricant, water, glycerol, friction, wear, hydraulics

1. INTRODUCTION

Hydraulic systems are widely used technologies in industry as high-power density systems that enable large forces, rigidity and endurance. The general challenge in improving the efficiency of a variety of sliding contacts is to reduce the coefficient of friction and control the wear mechanism for individual applications. This has often led to the development of new components or new materials and lubrication technologies [1]. Hydraulic systems are usually powered by standard mineral oil-based hydraulic oils. Today, we are aware that we have limited supplies of mineral oils, which are also harmful to nature and people, and require expensive disposal/storage after use [2]. Green tribology combines the classic challenges of tribology with an ecological approach that protects the environment and reduces health risks. One of the challenges on the way to sustainable hydraulics is the use of environmentally friendly green lubricants in combination with lightweight high-performance polymers, especially in the food industry, marine, automotive, aerospace, forestry and mobile machinery. These efforts and strategies for better wear protection and friction reduction in various equipment could potentially reduce such energy losses up to 40 % in the long term [3].

In the last decade, numerous studies have been conducted on the tribological properties of various high-performance polymer-based composites [3]. They have proven to be an excellent alternative to the classic elements made of steel, in various industrial applications for gears, pumps, cylinders, seals,

bearings etc. [4–8]. These polymers have considerably higher strength-to-weight ratio compared to steel alloys, which implies potentially significantly lighter components. High-performance polymer composites based on polyetheretherketone (PEEK), polyphenylene sulphide (PPS), ultra-high molecular weight polyethylene (UHMWPE) or polytetrafluoroethylene (PTFE) have shown good tribological properties under dry conditions or in water [9]. Their tribological properties can be improved with various functional fillers that could extend their service life, or further reduce friction and wear, as well as improve the response at higher loads, sliding speeds, pressure or temperatures [10]. Polyoxymethylene (POM) is a commercially acceptable substitute, with up to 10 times lower price compared to ultra-high performance polymers mentioned above [11]. It shows low coefficient of friction (COF) and specific wear rates in water, especially when reinforced with various fillers (PTFE fibres or particles [12], hexagonal boron nitride nanoparticles (h-BN) [13], especially for higher loads). In our recent study POM reinforced with CF30 gave comparable results in pure glycerol and water to four different high performance polymer composites [14].

The type of lubricant used is another aspect in the analysis of sliding contacts. Green lubricants are becoming mandatory in maritime and hydropower units. They are also increasingly desirable in forestry and mobile machinery [2] . Environmentally adaptable lubricants (EALs), such as vegetable oils, synthetic esters, polyalphaoleins, polyglycols and water, show low toxicity, excellent biodegradability and can be further improved tribologically by chemical modifications or by different anti-wear additives, antioxidants, etc.[15]. Glycerol is another good example of an environmentally friendly, non-toxic, biocompatible lubricant, which is main by-product of biodiesel production [16] . From an environmental point of view, glycerol could therefore be a better option for use in fluid power systems. Glycerol as a lubricant can potentially provide a very low coefficient of friction, enable low wear and reduce corrosion, although it has a very high viscosity, almost 20 times higher than traditional mineral based oils [17, 18]. Such high viscosity is not particularly desired, as it results in greater energy losses, more energy is required to overcome a thicker lubricating film, leading to lubricant breakdown and likely earlier components or system failures. However, since glycerol dissolves in water, the freezing point [19, 20], film thickness and viscosity of glycerol can be reduced by adding an appropriate amount of water. In this way, even a so-called liquid superlubricity can be obtained, with a friction coefficient of less than 0.01 reported [21]. Aqueous solutions of glycerol show good results in steel sliding contacts of rolling and sliding bearings [22, 23], especially when the water content is below 20 %, under different tribo-test conditions (different nominal stress and sliding speeds) [21]. Aqueous solutions of glycerol show good properties under boundary, mixed and elastohydrodynamic conditions [18, 22], although they are very sensitive to the water content under boundary lubrication conditions. When glycerol is used as a lubricant, there is data on a very low coefficient of friction even at higher contact pressures or sliding speeds between steel tribological couples [24]. So far, studies have mainly focused on the use of glycerol-aqueous solutions in conventional steel/steel contacts. Few studies address the simultaneous use of polymer composites and glycerol or similar green lubricants [14, 25].

Nowadays, it is very important to find an alternative solution for both innovative tribo-materials and for hydraulic fluids that are easier to obtain and that are safe for human health both in case of possible leakage and storage. The aim of this study is to analyse the tribological properties of commercially available pure POM and POM reinforced with 30 % carbon fibres (CF) in glycerol-water mixture. To evaluate such a green lubricant, we will compare the tribological properties of the selected polymer-steel sliding contact with wear and friction parameters in standard ISO VG46 hydraulic oil and the oldest environmentally adaptable lubricant, demineralized water. We will also compare and analyse their tribological properties at different working temperatures (room and elevated) and at two different loading speeds.

2. METHODS

2.1. Materials – samples and lubricants

For the tribological tests we prepared pure POM and POM reinforced with 30 % carbon fibers (POM CF30) samples. The polymer discs were cut from commercially available 30 mm diameter rod to a thickness of 5 mm. The polymer samples were polished in several steps using a RotoForce-3 automatic sample polishing and preparation device, to a final roughness of 0.1 µm. The ball was a commercially available, standard hardened bearing ball, 25 mm in diameter, made of AISI 440-C stainless steel. Before each test, the samples, clamps and ball were cleaned in ethanol and dried in air.

The first lubricant was commercially available, most commonly used ISO VG46 hydraulic oil. The second lubricant was a glycerol-water mixture. Commercially available redistilled glycerol, with a glycerol content ≥ 99.5 %, was used for preparation of 40 % glycerol and 60 % water (G+W) mixture lubricant. Based on our preliminary studies, by adding up to 40 % of water in the mixture, the good lubricating properties of pure glycerol remain, for a different polymer composites tested at room temperature. Demineralized water (W) was used as a third reference lubricant. The properties of selected lubricants were determined using an automatic viscometer SVM 3001 (Anton-Paar), which are in presented in the Table 1.

Table 1: Lubricants properties

Lubricant	Kin. Viscosity at 25 ◦C [mm^2/s]	Kin. Viscosity at 80 ◦C [mm^2/s]	Density at 25 ◦C [g/cm^3]	Density at 80 ◦C [g/cm^3]
ISO VG46	100	9	0.86	0.86
G+W	11.59	2.19	1.17	1.12
W	0.89	0.36	0.99	0.97

2.2. Tribological tests – experimental procedure

Tribological tests were performed on a Cameron-Plint TE 77 high-frequency tribometer (Figure 1). During the test, the steel bearing ball slides in a reciprocal mode on the polymer disk-like plate. The average sliding velocity was set to 0.2 and 0.02 m/s (frequency 40 Hz and 5 Hz, respectively, with stroke length 2.4 mm). The normal load was set to 50 N (90 -150 MPa maximum Hertzian pressure). Before each test, the polymer samples were completely immersed in the selected lubricant. The thermoset was placed in the lubricant bath, with a heating element under the bath to control the temperature and keep it constant during the test (at elevated temperature of 80 °C). Special care was taken to maintain the constant lubricant level during the test so that the tribo-pair was always fully immersed in the selected lubricant. After each test, the contact sliding area was marked on the ball with the help of an electric pen.

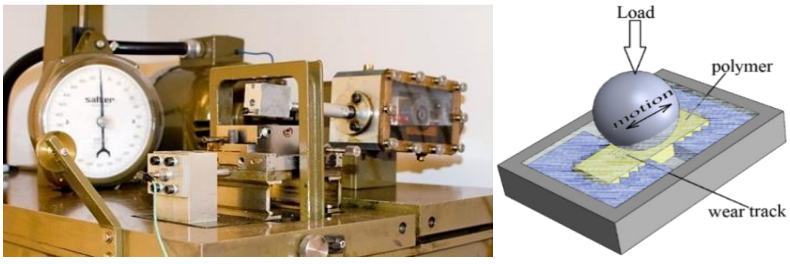


Figure 1: Scheme of tribological tests

During the tests, the current value of the friction coefficient was automatically recorded. The tests were performed for 90 minutes based on preliminary tests, which showed a stable value of the coefficient of friction was reached. Each test was repeated three times, and the average value of the steady-state friction coefficient is used for the comparison analysis.

2.3. Wear evaluation

The wear volume of the polymer discs was calculated from the dimensions of the wear tracks, which were measured using a 3D digital microscope with an additional nano-point confocal profilometer Hirox HRX-01 & NPS after the tribological tests. We determined a 3D profile for each calotte, from which the dimensions of the wear calotte and characteristic cross-sections were read on several characteristic locations along the wear track profile with Mountain Map software [14].The procedure is presented on Figure 2. The digital microscope also made it possible to determine the shape and dimensions of the surface films on the steel ball.

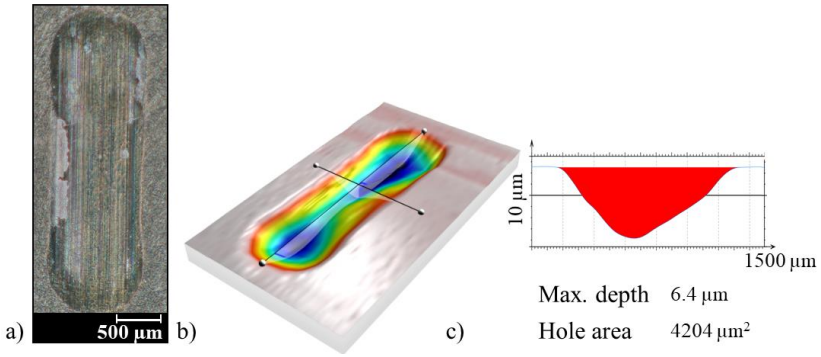


Figure 2. Procedure for wear analyses with a 3D digital microscope and Mountains Map software a) 2D digital image of worn surface's topography; b) 3D profile of worn volume; c) Measurement of cross-sectional area at a specific location.

3. RESULTS

3.1. Running in time

Figures 3 and 4 show the evolution of measured coefficient of friction during individual tests. When lubricated with ISO VG46 and G+W mixture, COF reaches a stable value very quickly, within the first 100 s of the test regardless the loading speed or temperature. The curves measured for POM follow the values measured for POM CF30. However, the COF measured in the water increases within the first 1000 s for POM, and for about 3000 s for POM CF30. Measured values for reinforced POM are considerably higher, and unstable compared to both pure POM and measurements in the other two lubricants.

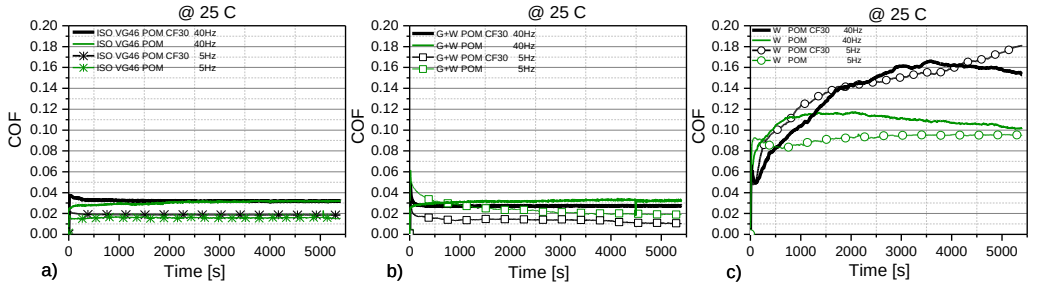


Figure 3: Coefficient of friction evolution during the tribological test of POM CF30 and POM in three different lubricants: a) ISO VG46; b) Glycerol + Water; c) Water all at room temperature

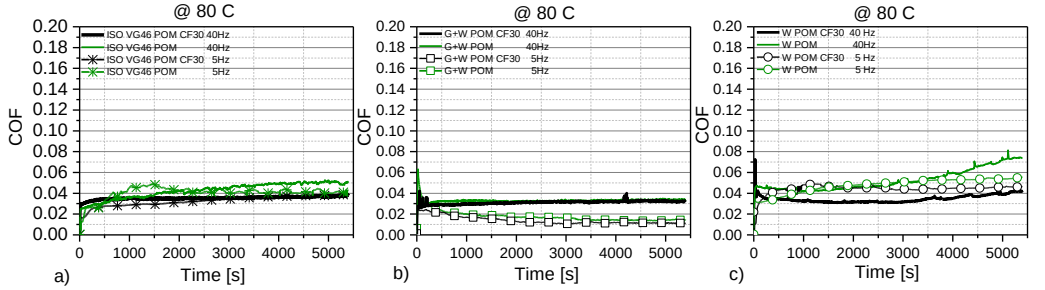


Figure 4: Coefficient of friction evolution during the tribological test of POM CF30 and POM in three different lubricants: a) ISO VG46; b) Glycerol + Water; c) Water all at 80 °C

Based on the results of measurements, the G+W mixture enables a good lubricating film, comparable to hydraulic oil. The increase in the measured average coefficient of friction observed in water indicates that the lubricating fluid film was not able to effectively separte friction surfaces of polymer plates and steel counterpart. Fluctuation observed in case of POM CF30 indicate, that reinforcements under such severe wear conditions incorporated additional instabilities.

3.2. Coefficient of friction

The average values of the steady-state coefficient of friction in ISO VG46 oil, G+W mixture and water are shown in Figure 5. The results showed comparable and low values of the coefficient of friction of 0.033 and 0.027 in the case of POM 30CF in oil and G+W mixture at room temperature. The measured values of the friction coefficient of POM CF30 in water are up to 6 times higher at room temperature.

It can also be seen from Figure 4 that the friction coefficients of the considered contacts in hydraulic oil and G+W mixture are comparable even in the case of lower sliding speed and at higher temperature. At a lower loading speed, there is a 40 % decrease in the coefficient of friction in the considered POM CF30 contacts in oil and G+W mixture. The decrease is smaller (10 %) at higher temperature in oil, and higher (60 %) in G+W mixture. Overall, COF increases at higher temperature in oil and G+W mixture, but not significantly.

On the contrary, in water COF does not change significantly at lower loading speed. However, measured COF decreases at higher temperature for both POM CF30 (4 times) and pure POM (2 times) in water, and measured values become comparable to values measured in hydraulic oil and G+W mixture.

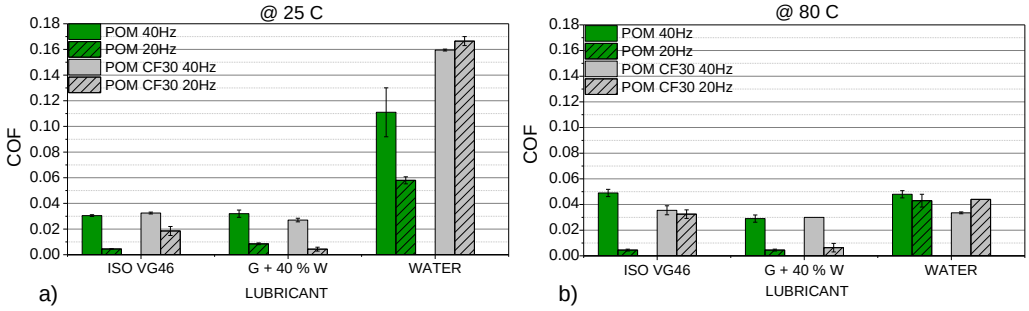


Figure 5: Average values of coefficient of friction ± sd of POM CF30 tested in three different lubricants:

a) at room temperature; b) at 80 °C.

3.3. Specific wear rate

Based on the cross-section measured along the 3D profile and overall dimensions of the wear track, we calculated the specific wear rate of the polymer discs in contact with the steel ball, in the case of using hydraulic oil ISO VG46, G+W mixture and demineralized water as lubricants, which is shown in Figure 6.

Specific wear, like the coefficient of friction, is comparable when using ISO VG46 hydraulic oil and G+W mixture, and most values for POM CF30 are in the range of 10^{-7} mm^3/Nm, and in the range of 3 x 10^{-8} mm^3/Nm for pure POM. The specific wear rate in both lubricants is lower than measured in water (~ 3 - 6 x 10^{-6} mm^3/Nm for POM CF30, ~ 3 x 10^{-7} mm^3/Nm for pure POM), at room temperature.

Specific wear rate significantly increased in hydraulic oil and glycerol-water mixture at elevated temperature, and became comparable to the specific wear rate measured in water (~ x 10^{-6} mm^3/Nm).

Specific wear is in general higher at lower sliding speed in for all lubricants tested.

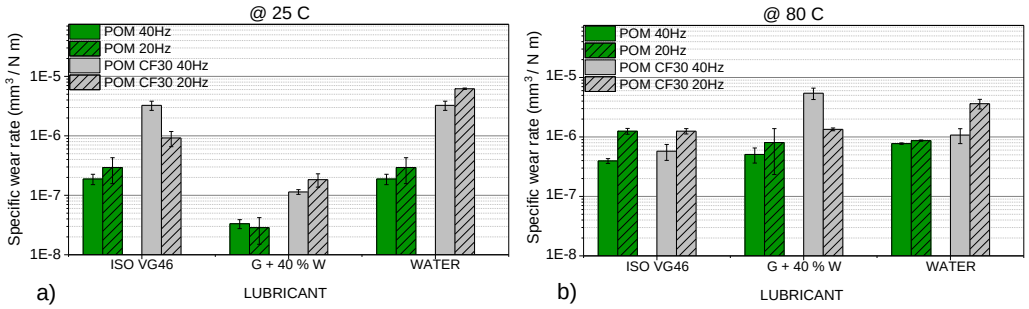


Figure 6: Average value of specific wear rate ± sd of POM CF30 tested in three different lubricants

a) at room temperature; b) at 80 °C

3.4. Worn surfaces analyses

Selected worn surfaces of the polymer samples observed by optical microscopy are presented in Figures 7 and 8. The surface appearance of POM 30CF is similar in shape and size (width ~900 μm, length ~3500 μm) in ISO VG46 oil (Figure 7.I.a) and G+W mixture (Figure 7.I.b). The wear track is only approximately 2 times wider at elevated temperature (width ~1400 μm). The wear mechanisms of POM composites in oil and G+W mixture are similar; lubricants form film thick enough to separate

surfaces in contacts. However, the narrowest wear track was observed in case of oil at room temperature and at higher frequency (~ 936 µm, Figure 6.I.a). The wear track is not significantly wider in glycerol + water mixture (~ 8 % increase) at room temperature, although the scratches' along the sliding direction are observed, especially in the middle of the wear scar (Figure 7.I.b). In case of pure water used at room temperature, both wear track width (~ 2.5 times increase), the number and intensity of scratches is significantly higher (Figure 7.I.c).

At elevated temperature, the worn surfaces are on average 50 % wider at approximately the same length (Figure 7.II.a/b) in oil and G+W mixture. On contrary, in the water scratches become more intense, although overall wear track width does not change significantly.

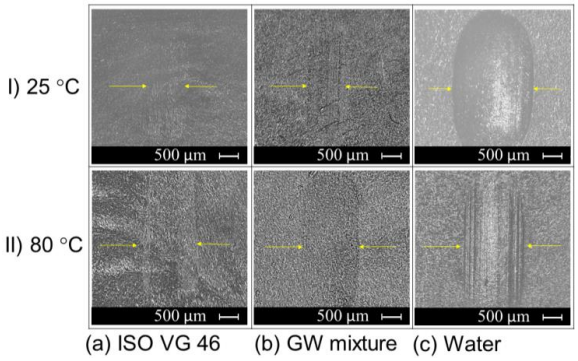


Figure 7: Digital images of worn POM CF30 polymer surfaces at (I) room and (II) elevated temperature, lubricated with: (a) Oil ISO VG46; (b) Glycerol-water mixture; (c) Water all at 40 Hz

At lower frequency (Figure 8.a-c) there was no significant difference in the wear track dimensions among oil and G+W mixture at both temperatures, in water the wear track was slightly (~ 10 %) wider. The main difference was observed in wear mechanism since the wear scratches along the sliding direction are more intense in case of G+W mixture, especially at higher temperature (Figure 8.II.b). This agrees with considerably higher specific wear measured in POM CF30 at elevated temperature, compared to oil or water (Figure 6.b). At elevated temperature, wear track width was comparable among lubricants (~ 1522 µm, Figure 8.II.a-c), and not significantly changed compared to room temperature conditions. The difference was in the scratch's intensity and depth, being the least intense in case of oil.

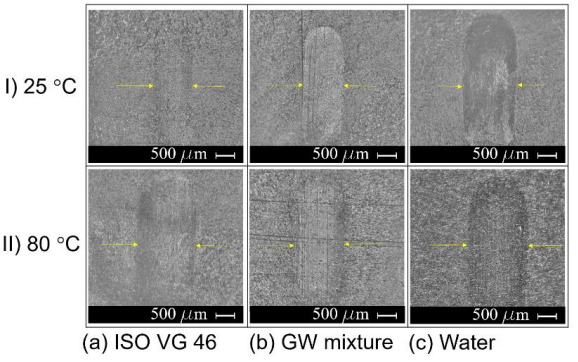


Figure 8: Digital images of worn POM CF30 polymer surfaces at (I) room and (II) elevated temperature, lubricated with: (a) Oil ISO VG46; (b) Glycerol-water mixture; (c) Water all at 5 Hz

4. DISCUSSION AND CONCLUSONS

In this study we have analysed the possibility of using an affordable engineering polymer POM CF30 in combination with a G+W mixture as a green lubricant as a potential tribo-pair for hydraulic applications. For reference and comparison, the same contacts were tested in ISO VG46 hydraulic oil and demineralised water, as the most widely available green base lubricant. The experiments were performed with parameters that correspond to seat on-off valves, and according to the experiment equipment limits. Using such polymer composites could potentially enable lighter hydraulic components, with excellent tribological properties. Additionally, new additive technologies could enable relatively quick prototyping accompanied with new component design. The tests were conducted at room temperature and elevated temperature, as expected in hydraulic applications, with the samples fully immersed in the selected lubricant. The tests were performed at two different sliding speeds.

Glycerol is an alternative lubricant whose annual production exceeds the demand for the same. Our recent study proved good tribological properties of pure glycerol for five different polymer composites [14]. Among observed composites, POM reinforced with 30 % carbon fibres successfully followed high performance PEEK reinforced with 30 % carbon fibres by tribological performance.

The results of this study show low values of the coefficient of friction of POM CF30 when G+W mixture is used as lubricant, at room temperature and elevated temperature and at both frequencies tested. The values of the coefficient of friction were similar when comparing G+W and ISO VG46 hydraulic oil. At room temperature, the coefficient of friction was up to 6 times higher in water than in oil and G+W mixture. At elevated temperature, however, the difference was not significant, although the lowest value was measured in the G+W mixture (0.03). At lower frequencies, a similar trend of COF decrease in the G+W mixture and oil was observed.

The specific wear rate was also lower in the G+W mixture and hydraulic oil, especially at room temperature (~ 10^{-7} mm^3/Nm). When comparing G+W mixture and hydraulic oil with water as lubricant, we measured one order of magnitude higher specific wear rate. At higher temperatures, the difference between the lubricants decreased as an increase in the specific wear rate was observed. At lower frequencies, a higher specific wear rate was observed, which is related to the more intense scratches observed. However, at higher temperatures, all specific wear rate values increased and there was no significant difference between the lubricants (the order of x 10^{-6} mm^3/Nm). Water proved to be a less effective lubricant at both frequencies and at room temperature compared to a G+W mixture and compared to water. At higher temperature, however, the difference decreased significantly.

Compared to the measurements of pure POM, in our recent study, a similar trend was observed in the measured coefficients of friction (influence of water, higher temperature, or lower frequency), but higher values of COF were measured overall for pure POM than for POM reinforced with 30 % CF. In contrast, lower values of specific wear rate were observed for pure POM compared to reinforced POM. This effect is probably due to the fact that in the case of reinforced POM the carbon fibres carry most of the applied load, but at the same time the thinning of the fibres indicates a fracture of the POM matrix and a higher wear rate [27]. However, further elemental or spectroscopic analysis of the worn surfaces is required to discuss the difference in detail. Based on the current measurements, POM CF30 gives excellent tribological results in both hydraulic oil and a G+W mixture and can be considered as a potential material or even a combination of material and lubricant for hydraulic applications where low load and high frequency are expected.

NOMENCLATURE

EAL	Environmentally adapted lubricants
POM	Polyoxymethylene
PEEK	Polyetheretherketone
PPS	polyphenylene sulphid
UHMWPE	ultra-high molecular weight polyethylene
PTFE	polytetrafluoroethylene
CF	Carbon Fibre
G+W	Glycerol + Water mixture
W	Water
COF	Coefficient of friction /

REFERENCES

1. Deaconescu T, Deaconescu A (2022) Experimental Research on Polymer-Based Coaxial Sealing Systems of Hydraulic Cylinders for Small Displacement Velocities. Polymers 2022, Vol 14, Page 290 14:290. https://doi.org/10.3390/POLYM14020290
2. Deuster S, Schmitz K, Widomski K, et al (2021) Bio-Based Hydraulic Fluids and the Influence of Hydraulic Oil Viscosity on the Efficiency of Mobile Machinery. Sustainability 2021, Vol 13, Page 7570 13:7570. https://doi.org/10.3390/SU13147570
3. Friedrich K (2018) Polymer composites for tribological applications. Advanced Industrial and Engineering Polymer Research 1:3–39. https://doi.org/10.1016/J.AIEPR.2018.05.001
4. Matkovič S, Pogačnik A, Kalin M (2021) Wear-coefficient analyses for polymer-gear life-time predictions: A critical appraisal of methodologies. Wear 480–481:203944. https://doi.org/10.1016/J.WEAR.2021.203944
5. Stryczek J, Banaś M, Krawczyk J, et al (2017) The Fluid Power Elements and Systems Made of Plastics. Procedia Eng 176:600–609. https://doi.org/10.1016/J.PROENG.2017.02.303
6. Marciniak L, Banaś M, Stryczek J (2016) The Design and Theoretical and Experimental Study of the Plastic Hydraulic Valves. 9th FPNI PhD Symposium on Fluid Power, FPNI 2016. https://doi.org/10.1115/FPNI2016-1511
7. Li Y, Shang Y, Wan X, et al (2022) Design and experiment on light weight hydraulic cylinder made of carbon fiber reinforced polymer. Compos Struct 291:115564. https://doi.org/10.1016/J.COMPSTRUCT.2022.115564
8. Majdič F, Kalin M (2012) Vpliv tesnil na obnašanje vodno-hidravličnega valja. In: Vižintin J., Sedalček M. (eds) International Conference on Tribology: Conference on Tribology, Cooling Fluids and Technical Diagnostics. Slovenian Society for Tribology, Ljubljana
9. Kurdi A, Chang L (2018) Recent Advances in High Performance Polymers—Tribological Aspects. Lubricants 2019, Vol 7, Page 2 7:2. https://doi.org/10.3390/LUBRICANTS7010002
10. Ren Y, Zhang L, Xie G, et al (2021) A review on tribology of polymer composite coatings. Friction 9:429–470. https://doi.org/10.1007/S40544-020-0446-4
11. Siddiqui MSN, Pogacnik A, Kalin M (2023) Influence of load, sliding speed and heat-sink volume on the tribological behaviour of polyoxymethylene (POM) sliding against steel. Tribol Int 178:108029. https://doi.org/10.1016/J.TRIBOINT.2022.108029
12. Zhang H, Wu Y, Liang M, et al (2020) Comparative study on the mechanical, tribological, and thermal properties of POM composites filled with different PTFE: https://doi.org/101177/0892705720932620. https://doi.org/10.1177/0892705720932620
13. Gao C, Guo G, Zhang G, et al (2017) Formation mechanisms and functionality of boundary films derived from water lubricated polyoxymethylene/hexagonal boron nitride nanocomposites. Mater Des 115:276–286. https://doi.org/10.1016/J.MATDES.2016.11.016
14. Trajkovski A, Novak N, Pustavrh J, et al (2023) Performance of Polymer Composites Lubricated with Glycerol and Water as Green Lubricants. Applied Sciences 13:7413. https://doi.org/10.3390/APP13137413

15. Hernández-Sierra MT, Bravo-Sánchez MG, Báez JE, et al Improvement Effect of Green Lubricants on the Tribological and Mechanical Performance of 4140 Steel. https://doi.org/10.3390/app9224896
16. Zhang T, Liu C, Gu Y, Jérôme F (2021) Green Chemistry CRITICAL REVIEW Glycerol in energy transportation: a state-of-the-art review. https://doi.org/10.1039/d1gc02597j
17. Li J, Zhang C, Ma L, et al (2012) Superlubricity Achieved with Mixtures of Acids and Glycerol. Langmuir 29:271–275. https://doi.org/10.1021/LA3046115
18. Chen Z, Liu Y, Zhang S, Luo J (2013) Controllable superlubricity of glycerol solution via environment humidity. Langmuir 29:11924–11930. https://doi.org/10.1021/LA402422H
19. Trejo González JA, Longinotti MP, Corti HR (2011) The Viscosity of Glycerol−Water Mixtures Including the Supercooled Region. J Chem Eng Data 56:1397–1406. https://doi.org/10.1021/JE101164Q
20. Liu C, Qiao Y, Lv B, et al (2020) Glycerol based binary solvent: Thermal properties study and its application in nanofluids. International Communications in Heat and Mass Transfer 112:104491. https://doi.org/10.1016/J.ICHEATMASSTRANSFER.2020.104491
21. Ma Q, He T, Khan AM, et al (2021) Achieving macroscale liquid superlubricity using glycerol aqueous solutions. Tribol Int 160:107006. https://doi.org/10.1016/J.TRIBOINT.2021.107006
22. Shi Y, Minami I, Grahn M, et al (2014) Boundary and elastohydrodynamic lubrication studies of glycerol aqueous solutions as green lubricants. Tribol Int 69:39–45. https://doi.org/10.1016/J.TRIBOINT.2013.08.013
23. Björling M, Shi Y (2019) DLC and Glycerol: Superlubricity in Rolling/Sliding Elastohydrodynamic Lubrication. Tribol Lett 67:1–8. https://doi.org/10.1007/S11249-019-1135-1/FIGURES/2
24. Habchi W, Matta C, Joly-Pottuz L, et al (2011) Full film, boundary lubrication and tribochemistry in steel circular contacts lubricated with glycerol. Tribol Lett 42:351–358. https://doi.org/10.1007/S11249-011-9778-6/FIGURES/6
25. Somberg J, Saravanan P, Vadivel HS, et al (2021) Tribological characterisation of polymer composites for hydropower bearings: Experimentally developed versus commercial materials. Tribol Int 162:107101. https://doi.org/10.1016/J.TRIBOINT.2021.107101
26. Habchi W, Matta C, Joly-Pottuz L, et al (2011) Full film, boundary lubrication and tribochemistry in steel circular contacts lubricated with glycerol. Tribol Lett 42:351–358. https://doi.org/10.1007/S11249-011-9778-6/FIGURES/6
27. Zhang L, Qi H, Li G, et al (2017) Impact of reinforcing fillers' properties on transfer film structure and tribological performance of POM-based materials. Tribol Int 109:58–68. https://doi.org/10.1016/j.triboint.2016.12.005

Chapter 7

Mobile Applications

METHODOLOGY OF SYSTEM PARAMETER OPTIMIZATION FOR PARALLEL ELECTRIC HYDRAULIC HYBRID MOBILE MACHINE VIA CONVEX PROGRAMMING

peer reviewed

Zichang Lin[1], Jiaming Wu[1], Zhenchuan Lin[1], Feng Wang[1]*, Bing Xu[1]

[1] *State Key Laboratory of Fluid Power Components and Mechatronic Systems, School of Mechanical Engineering, Zhejiang University, 866 Yuhangtang Road, 310030, Hangzhou, China.*

* Corresponding author: E-mail address: dieter@zju.edu.cn

ABSTRACT

Hydraulic hybrid powertrain is widely investigated for its high power density of hydraulic power system. In designing a hybrid vehicle, finding the combined optimality of component sizing and energy management is essential for minimizing vehicle costs and maximizing energy efficiency. Simultaneous optimization framework is an effective and important method due to its computational efficiency and resolution. In this paper a convex programming-based system parameter optimization framework is proposed for hydraulic hybrid vehicle. This technique allows simultaneous optimization of component sizing and energy management by converting it into a convex problem. To illustrate this, the system optimization problem in a parallel electric hydraulic hybrid wheel loader is posed over a fixed loading cycle. The Pareto front of PEHH system parameter optimization problem is obtained. The HM size is the main factor of system performance trade-off between battery aging and energy consumption. PEHH can reduce the battery capacity loss in a loading cycle by 26.4% compared to the pure electric drive with a 13.0% increase of energy consumption. With the same grid number of 7, the CP-based simultaneous method consumes 99% less computing time than DP-based bi-level method and provides the optimal solution with 1.5% less battery capacity loss.

Keywords: System parameter optimization, wheel loader, electric hydraulic hybrid system, convex programming

1. INTRODUCTION

The electrification of mobile machines is becoming a well-established trend. The electric construction equipment market is expected to grow to $19.9 billion in 2027 at a compound annual growth rate of more than 21% according to The Business Research Company. In 2023, China's sales of electric loaders reached 3595 units, accounting for 3.5% of the total sales of wheel loaders.

Parallel electric hydraulic hybrid powertrain (PEHH) is a promising substitute for pure electric drive and has been widely investigated [1]. The electric hydraulic hybrid powertrain provides hydraulic launching and braking to release battery current stress. For on-road vehicles like city bus, the energy saving of PEHH can reach up to 50% in the best scenario [2]. PEHH has also been applied on mobile machines such as wheel loaders [3], excavators [4] and forklifts [5], improving their energy efficiency and battery lifetime.

System parameter optimization has been an important topic in the field of hybrid vehicle [6]. Unlike the non-hybrid vehicles, system parameter optimization of hybrid powertrain should include the optimization of energy management strategy (EMS). There are four methods for the combined plant and controller optimization problem and global optimum can only be guaranteed with bi-level and

simultaneous strategies.

In the bi-level optimization method, the possible component sizes are enumerated in the outer loop, and the optimal energy management is assessed in the inner loop [7]. By comparison, simultaneous method is an effective and important method due to its computational efficiency and resolution. A simultaneous optimization method is using an adjustable rule-based strategy and optimize the strategy design parameters and system parameters simultaneously [8]. However, for complex problems, the rule-based strategy is not close enough to the global optimum.

Another simultaneous optimization method is the convex programming. By modelling the combined plant and controller optimization problem as a convex programming problem, the optimal control variable sequence and the optimal system parameters are obtained simultaneously with a convex optimization solver [9]. This method is adopted extensively in electric hybrid systems, but has not been performed on hydraulic hybrid systems [10]

In this paper, a system parameter optimization method based on convex programming is proposed for PEHH mobile machines. The design problem of a PEHH wheel loader is used to demonstrate the optimization method. The rest of this paper is organized as follows. Section 2 introduces the PEHH wheel loader. Section 3 illustrates the optimization problem formulation. Section 4 introduces the framework of the proposed optimization method. Section 5 shows the results of the optimization study. The last section is the conclusion and outlook.

2. PARALLEL ELECTRIC HYDRAULIC HYBRID WHEEL LOADER

The PEHH wheel loader schematic is shown in **Figure 1**. The working and other functions are not considered here for simplicity. In the electric wheel loader, these functions are electrically decoupled from drivetrain and are not of the interest of this study. The main energy source is a battery pack, powering the electric motor through an inverter. High-pressure and low-pressure hydraulic accumulators and a variable hydraulic motor are added to provide hydraulic auxiliary power. The electric motor (EM) and hydraulic motor (HM) are mechanically coupled via gears.

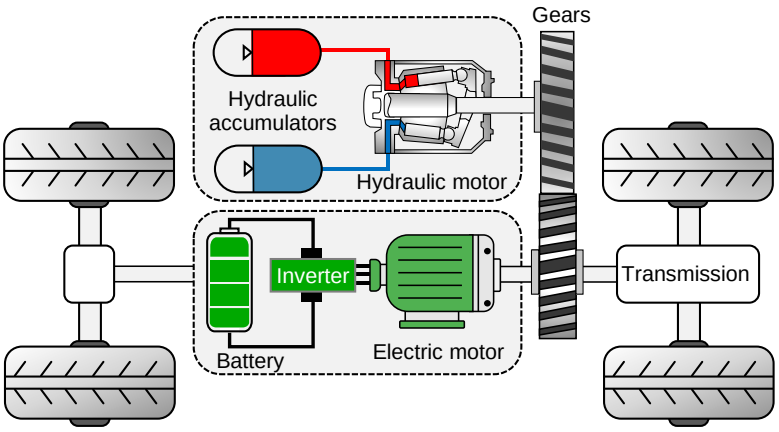


Figure 1: PEHH wheel loader schematic

Generally, the system parameters to be optimized in PEHH wheel loader includes: gear ratio, EM size, hydraulic motor geometric displacement, accumulator size and pre-charge pressure. The transmission ratio is not changed to make use of the original drive train. The main wheel loader parameters are shown in **Table 1**.

Table 1: Parameters of the wheel loader

Wheel loader parameter	Value	Unit
Operating weight	19000	kg
Load weight	5500	kg
Maximum speed	40	km/h

3. PROBLEM FORMULATION

In this section, the problem formulation and modelling details are introduced. The PEHH is expected to provide the same driving performance as the pure electric drive with smaller EM, less energy consumption and longer battery lifetime. The problem constitutes an objective function and constrains. The variables to be optimized are the state vector and the sizing factor for components. The control input variable u includes the control inputs of length N. These variables are explained in more detail at the end of this section.

3.1. Driving cycle

In the optimization problem, a given driving pattern is required to find the optimal component sizes and the energy management variables. The V loading cycle is a typical driving cycle for wheel loaders that can reflect real-life driving. In this study, the EM propel torque $T_{load}(k)$ and speed ω_{EM} of an electric wheel loader during a V loading cycle are used as the input of the optimization problem. The driving cycle is divided into N discrete instants with a time difference of ΔT.

3.2. Modelling

In this section the models of the powertrain and its components are presented. Since the models are used for convex programming, they all guarantee convexity. Quasi-static models are therefore approximated with nonlinear convex functions and some variable change is also used.

Powertrain

The EM and HM are mechanically linked to the drive train in the studied PEHH and can propel the wheels. Having the required EM speed $\omega_{EM}(k)$ and propel torque $T_{load}(k)$, the powertrain model is described by torque balance equations, given as

$$T_{EM}(k)\omega_{EM}(k) + T_{HM}(k)\omega_{HM}(k) = T_{load}(k)\omega_{EM}(k) \quad (1)$$

where $T_p(k)$ and $\omega_{EM}(k)$ is given by loading cycle data, T_{EM}(k) is EM torque, $T_{HM}(k)$ is HM torque and $\omega_{HM}(k)$ is HM speed calculated as $\omega_{HM} = i_G\omega_{EM}$, where i_G is gear ratio.

Battery

The battery capacity is determined by required operation time and is not optimized in this study. It's modelled as an open circuit voltage V_{oc} in series with a constant internal resistance R. The open circuit voltage is approximated to be constant in the allowed state of charge operating region. The terminal power, P_{bat}, and the battery current, I_{bat}, are calculated as

$$P_{bat}(k) = I_{bat}(k)V_{OC}(k) - I_{bat}^2(k)R \quad (2)$$

The battery current I_{bat} is chosen to be positive when charging. The electric energy consumption of battery E_{cons} is calculated by integrating battery power.

The battery cycle aging capacity loss model is adopted from [11]. It is related to the battery full

equivalent cycle FEC, the battery state of charge SOC and the battery current I_{bat}. During a loading cycle, FEC and SOC are approximated to be constant. The battery cycle aging capacity loss Q_{ag} in percentage is given by

$$\begin{aligned} Q_{ag} &= \sum_{k=1}^{N} \Delta FEC(k) \left(a\left(\frac{|I_{bat}(k)|}{Q_b}\right) + b \right) \left(c(0.4 - SOC)^3 + d\right) \frac{FEC^{-0.5}}{2} \\ \Delta FEC(k) &= \frac{1}{3600} \frac{1}{2Q_b} \frac{I_{bat}(k-1) + I_{bat}(k)}{2} \Delta T \end{aligned} \tag{3}$$

where Q_b is the battery capacity in Ah, a, b, c, d is constant model coefficients.

EM

The EM model with the inverter is described by a power loss map, $P_{EM,loss,base}$, where the losses are measured at steady-state for different torque-speed combinations. The power losses for each EM speed are approximated by a second-order polynomial in torque. To vary the size of the EM, the torque limits and losses are scaled linearly by scaling factor s_{EM}. In this way the losses of the scaled EM are calculated at each time instant as

$$P_{EM,loss}(k) = c_1(k) \frac{T_{EM}^2(k)}{s_{EM}} + c_2(k)|T_{EM}(k)| + c_3(k) s_{EM} \tag{4}$$

where the coefficients $c_1 \geq 0$, c_2 and c_3 are functions of $\omega_{EM}(k)$ and are calculated using least squares method for a number of grid points of ω_{EM}. For speed values not belonging to the grid nodes, the coefficients are obtained by linear interpolation. The accuracy of the approximation are high and are discussed in detail in [9]. The maximum EM torque at each time instant $T_{max}(k)$ is obtained by interpolation and considered as constrains to EM torque T_e.

The electric power load of EM, P_{EM}, is given by

$$P_{EM}(k) = \omega_{EM}(k) T_{EM}(k) + P_{EM,loss}(k) \tag{5}$$

The electric power load of hydraulic working functions P_{HW} is not considered here and can be easily included in the problem by considering $P_{bat}(k) = P_{EM}(k) + P_{HW}(k)$.

HM

The HM torque, T_{HM}, and the HM flow rate Q_{HM} are functions of geometric displacement D, displacement fraction x_{HM}, pressure differential p_{HM} and angular speed ω_{HM} of HM. A modified version of Wilson's model [12] for variable-displacement pump-motor is used here.

To preserve the problem convexity[1], the variable change $T_{xpD}(k) = x_{HM}(k) p_{HM}(k) D$ and $T_{pD}(k) = p_{HM}(k) D$ are introduced. The equations (6) and (7) are rewritten as

$$T_{HM}(k) = T_{xpD}(k) + K_\omega \omega_{HM}(k) D + K_p T_{pD}(k) + T_0 D \tag{6}$$

$$P_{HM,Hydr}(k) = p_{HM}(k) Q_{HM}(k) = T_{xpD}(k) \omega_{HM}(k) - k_s(k) \frac{T_{pD}^2(k)}{D} - q_0 T_{pD}(k) \tag{7}$$

where $P_{HM,Hydr}$ is the hydraulic power of HM, k_s is calculated by $k_s = a + b\omega_{HM} + c\omega_{HM}^2$, The coefficients $a, b, c,\ K_w, K_p, T_0, q_0$ are calculated using least squares method with a number of grid points of ω_{HM}, p_{HM} and x_{HM}. The accuracy of the approximation is discussed in the simulation section. The HM displacement x_{HM} is chosen to be positive when motoring.

It is worth noting that even when HM has no power output, there are still losses due to leakage and

[1] $f(x, y) = x * y$ is neither convex or concave.

friction. For example, the volumetric efficiency of HM is negative if the pump leakage is greater than the theoretical pumping flow rate.

Hydraulic accumulator

The low-pressure hydraulic accumulator pressure is constant, equal to atmospheric pressure. For the high-pressure accumulator, the accumulator pressure p_{HA} is approximated as an affine function to simplify the problem and preserve problem convexity, given by:

$$p_{HA}(k) = k_{HA}E_{HA}(k) + p_0 \tag{8}$$

$$T_{pD}(k) = p_{HM}(k)D = K_{HA}E_{HA}(k) + B \tag{9}$$

The coefficient k_{HA}, $K_{HA} = k_{HA}D$ and $B = p_0D$ are related to the accumulator parameter p_0 and V_0 and used as a sizing factor. The accuracy of the approximation is discussed in the simulation section.

The energy stored in the high-pressure accumulator E_{HA} are given by

$$E_{HA}(k) = E_{HA}(k-1) + \frac{P_{HA,Hydr}(k-1) + P_{HA,Hydr}(k)}{2}\Delta T \tag{10}$$

For charge sustaining of hydraulic accumulator, the initial E_{HA} is required to be no higher than the final E_{HA}.

4. OPTIMIZATION FRAMEWORK

The framework schematic of the proposed optimization method is shown in **Figure 2**. With some approximation and variable change, some component sizing factors can be integrated in the convex problem including HM geometric displacement D, EM scaling factor s_{EM} and accumulator pre-charge pressure p_0. Other sizing factors such as gear ratio i_g and accumulator volume V_0, are optimized in the outer loop.

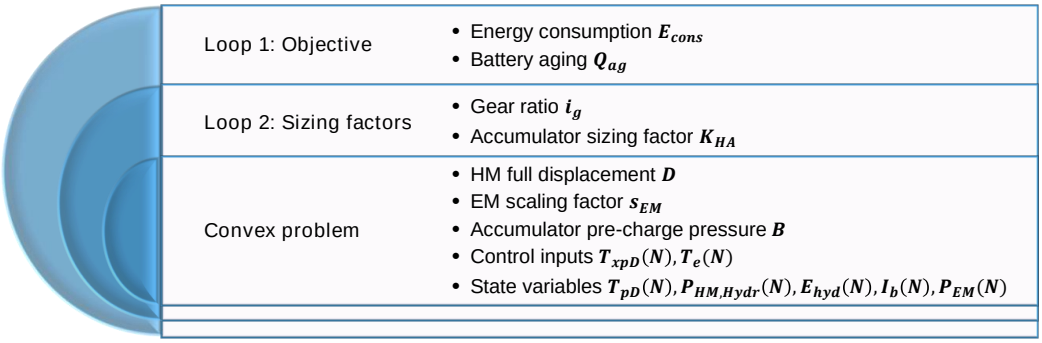


Figure 2: Framework schematic of the proposed optimization method

4.1. Objective function

The objective of the optimization contains two indexes, energy consumption and battery aging. Pareto front stands for a set of solutions where one of the objectives cannot be improved without sacrificing another one. In this paper, the Pareto optimal solutions are obtained by minimizing battery capacity loss under different maximum energy consumption constrains.

4.2. Sizing factors

Some sizing factors are hard to be integrated in the convex programming problem. To address this, an outer loop is added to enumerate the possible solution of these factors. In this study, the sizing factor loop include two index, Gear ratio i_G and Accumulator sizing factor K_{HA}. A number of 2-dimension vectors are generated and enumerated. The convex programming problem is solved for each sizing factor.

4.3. Convex optimization

Combing eq. (1)-(10). Most of the equations preserve the problem convexity since they include only affine functions, except (2), (3), (4) and (7). The original equalities are relaxed to inequations (10a) -(10d) to preserve the problem convexity. Because these inequities include convex non-linear term, including absolute value function and quadratic-over-linear function. The optimal solution will satisfy (10a) -(10d) with equality as it is optimal for energy consumption and capacity loss. When the equality of these inequality does not hold, there are some extra energy loss or battery capacity loss in the system [9]. Therefore, it will be optimal for the non-relaxed problem as well. The convex programming problem is formulated as follow, the variables to be optimized are in bold font:

$$
\begin{aligned}
&min\ \boldsymbol{Q_{ag}}\\
&\boldsymbol{subject\ to}:\\
&T_{load}(k)\omega_{EM}(k) = \boldsymbol{T_{EM}}(k)\omega_{EM}(k) + \boldsymbol{T_{HM}}(k)\omega_{HM}(k)\\
&\boldsymbol{T_{HM}}(k) = \boldsymbol{T_{xpD}}(k) + K_\omega\omega_{HM}(k)\boldsymbol{D} + K_p\boldsymbol{T_{pD}}(k) + T_0\boldsymbol{D}\\
&\boldsymbol{P_{HM,Hydr}}(k) \geq \boldsymbol{T_{xpD}}(k)\omega_{HM}(k) - k_s(k)\frac{T_{pD}^2(k)}{D} - q_0T_{pD}(k) \qquad \text{(a)}\\
&\boldsymbol{E_{HA}}(k) = \boldsymbol{E_{HA}}(k-1) + \frac{\boldsymbol{P_{HA,Hydr}}(k-1) + \boldsymbol{P_{HA,Hydr}}(k)}{2}\Delta T\\
&\boldsymbol{T_{pD}}(k) = K_{HA}\boldsymbol{E_{HA}}(k) + B\\
&\boldsymbol{P_{EM}}(k) \geq c_1(k)\frac{\boldsymbol{T_{EM}^2}(k)}{\boldsymbol{s_{EM}}} + c_2(k)|\boldsymbol{T_{EM}}(k)| + c_3(k)\boldsymbol{s_{EM}} + \omega_{EM}(k)\boldsymbol{T_{EM}}(k) \qquad \text{(b)}\\
&\boldsymbol{P_{EM}}(k) \leq \boldsymbol{I_{bat}}(k)V_{OC}(k) - \boldsymbol{I_{bat}^2}(k)R \qquad \text{(c)}\\
&\boldsymbol{E_{cons}} = \sum_{k=1}^{N}\boldsymbol{I_{bat}}(k)V_{OC}(k)\Delta T\\
&\boldsymbol{Q_{ag}} \geq \sum_{k=1}^{N}\boldsymbol{\Delta FEC}(k)\left(a\left(\frac{|\boldsymbol{I_{bat}}(k)|}{Q_b}\right) + b\right)(c(0.4 - SOC)^3 + d)\frac{FEC^{-0.5}}{2} \qquad \text{(d)}\\
&\boldsymbol{\Delta FEC}(k) = \frac{1}{3600}\frac{1}{2Q_b}\frac{\boldsymbol{I_{bat}}(k-1) + \boldsymbol{I_{bat}}(k)}{2}\Delta T\\
&\boldsymbol{E_{cons}}(k) \leq E_{max}\\
&|\boldsymbol{T_e}(k)| \leq T_{max}(k)\boldsymbol{s_{EM}}\\
&\boldsymbol{s_{EM}} \in [0, s_{EM,max}]\\
&\boldsymbol{T_{pD}}(k) \in [p_0D, p_{max}D]\\
&\boldsymbol{D} \in [D_{min}, D_{max}]
\end{aligned}
\qquad (11)
$$

5. OPTIMIZATION STUDY

In this section, the accuracy of model approximation is first discussed. The Pareto front is then obtained with given optimization framework. The influence of component sizing factors on powertrain performance and the optimal energy management under different optimization objective are also analysed. Comparison between the computing time of the proposed convex optimization method and the bi-level method is carried out to check its computational efficiency.

5.1. Driving cycle and model parameters

The loading cycle used in the study is obtained via field test as shown in **Figure 3**. The loading time is 30.7 seconds long and the maximum vehicle speed is about 12 km/h. The EM propel torque is provided by EM and HM in the PEHH powertrain.

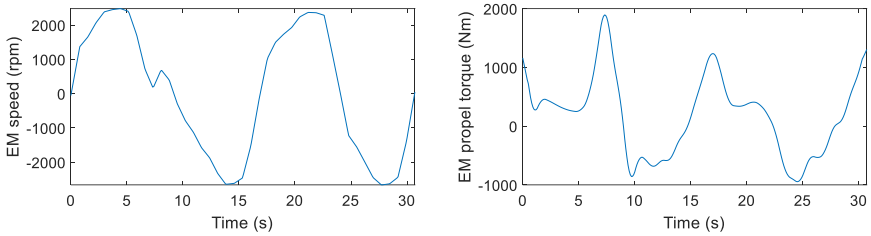


Figure 3: EM speed and propel torque during a loading cycle

The HM efficiency data includes mechanical and volumetric efficiency. The hydraulic pump/motor model coefficients are obtained by fitting the efficiency test data of a variable-displacement pump. The efficiency map obtained after model fitting and the original efficiency map are shown in **Figure 4**. The efficiency test data is measured with displacement fraction of $x = -1, -0.8, -0.6, -0.4, -0.2$, and only the data of $x = -1$ *and* $x = -0.2$ is shown here due to page limit. As shown, the fitted efficiency map is a set of ellipse curve and is close to the original test data.

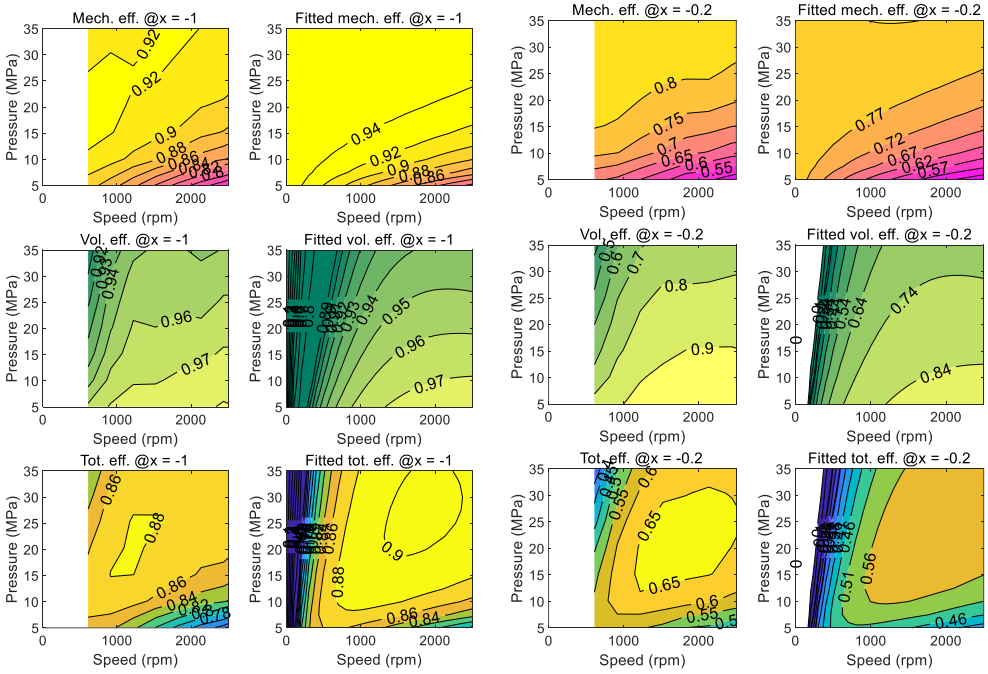


Figure 4: Original and fitted HM mechanical, volumetric and total efficiency map. (a) $x = -1$. (b) $x = -0.2$

The HM efficiency in pumping and motoring mode is shown in **Figure 5**. The maximum efficiency of HM is about 90%.

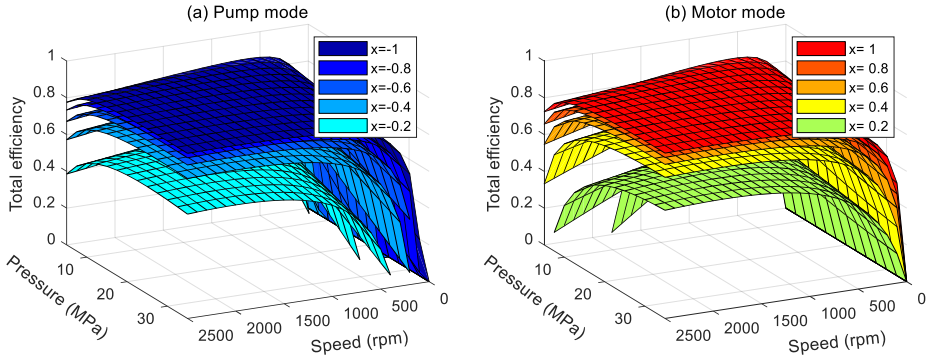


Figure 5: Fitted HM efficiency map with different displacement fraction. (a) pump mode. (b) motor mode

Under the hydraulic accumulator isothermal model, the curve of gas pressure versus volume is an inverse proportional function curve. Within a limited pressure range, the inverse proportional function curve can be approximated with an affine function. The error of pressure of model approximation can affect the maximum torque and power loss of HM, but the error is smaller than 1.0 MPa and its impact on system performance can be ignored.

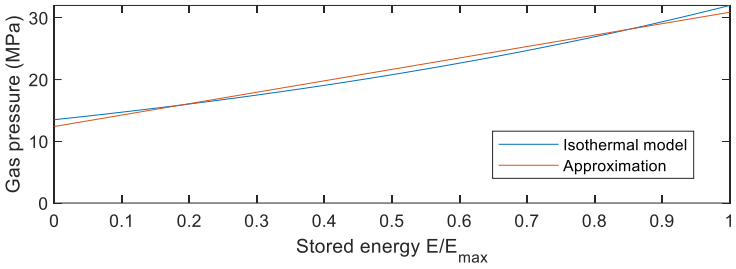


Figure 6: Hydraulic accumulator isothermal model and approximation

The EM and HM model coefficients are obtained by fitting actual machine test data. The EM used in this study is an 80-kW permanent magnet synchronous motor with a maximum speed of 3300 rpm. The efficiency and power loss map obtained after model fitting are shown in **Figure 7**.

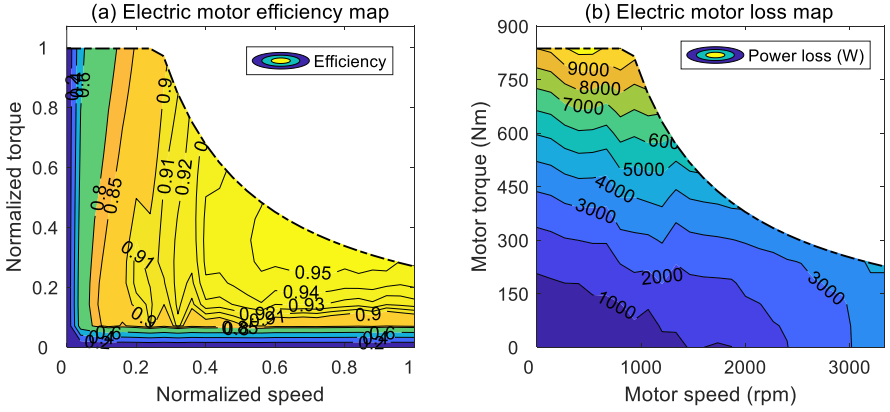


Figure 7: EM efficiency map and loss map.

5.2. System parameter optimization results

The Pareto front and Pareto optimal points for the PEHH parameter design solutions is shown in **Figure 8**. In the $E_{cons} - Q_{ag}$ space, each point corresponds to one system parameter solution. The left boundary in the dashed line of all possible points is the Pareto front consist of different Pareto optimal points with different HM displacement. Among six selected designs, solution with largest displacement $D = 149\ cm^3/rev$ has the lowest battery capacity loss and highest energy consumption, 1.93×10^{-6} % and 1721.3 kJ, respectively. Solution with lowest HM displacement $D = 20\ cm^3/rev$ has the highest battery capacity loss and lowest energy consumption, 2.52×10^{-6} % and 1549.2 kJ, respectively.

Compared to the electric drive solution (2.62×10^{-6} % and 1522.9 kJ), the energy consumption of PEHH is always higher while the battery capacity loss can be reduced by 26.4%. The reason is that HM leads to more leakage and friction loss but the battery charging and discharging is reduced with hydraulic power assist. Solution #3 provides a relatively good trade-off between battery aging and energy consumption, reducing the battery capacity loss by 17% with only 5% more energy consumption. More precise trade-off requires weighted calculation of capacity losses and energy consumption targets, considering battery price and electricity price.

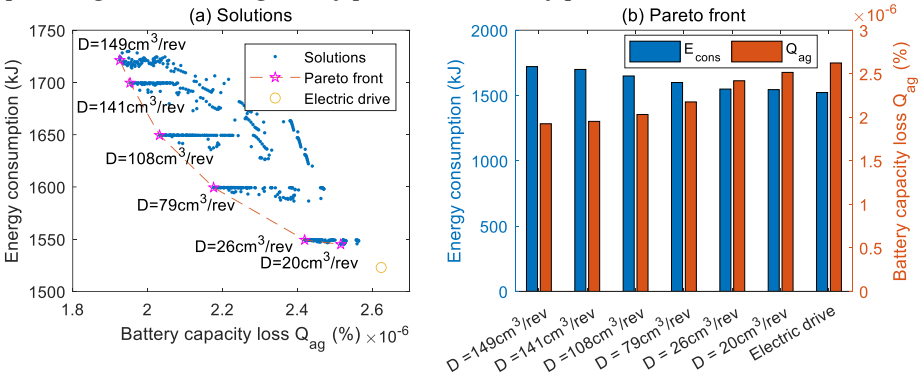


Figure 8: Pareto front and Pareto optimal points for the system design solutions. (a) Solution distribution. (b) System performances of Pareto front

The component energy losses and component sizing factors of Pareto optimal solutions with different energy consumption is shown in **Table 2**. The EM energy loss is reduced while the HM loss increases in PEHH. Firstly, when the weight of battery aging in the objective increases, the optimal HM geometric displacement increases accordingly and the EM size decreases. Therefore, the hydraulic hybrid ratio is the main factor of system performance trade-off.

Table 2: Performances and system parameters of the Pareto optimal solutions

Solution	Energy consumption (kJ)	Battery capacity loss (%)	HM displacement (cm^3/rev)	EM maximum torque (Nm)	Accumulator volume (L)	Gear ratio	Pre-charge pressure (MPa)	Initial pressure (MPa)
#1	1721.3	1.93×10^{6}	149	1554	19.9	0.9	13.5	32
#2	1699.6	1.95×10^{6}	141	1558	20.9	0.9	13.5	32
#3	1649.5	2.03×10^{6}	108	1656	16.0	0.9	13.5	32
#4	1599.4	2.18×10^{6}	79	1735	10.6	0.8	13.5	32
#5	1549.2	2.42×10^{6}	26	1827	4.4	0.9	13.5	32
#6	1545.1	2.52×10^{6}	20	1848	6.4	0.7	13.5	27
Electric Drive	1522.9	2.62×10^{6}	N/A	1900	N/A	N/A	N/A	N/A

The total optimization time of the proposed method varies with the number of sizing factors. As shown in **Figure 9**, for the typical bi-level optimization method based on dynamic programming (DP), the sizing factors are enumerated with six loops. With a grid number of N, the DP problem needs to be solved for N^6 times. While the convex programming problem is solved for N^2 times with two loops. This provides smaller computational burden and high discrete resolution.

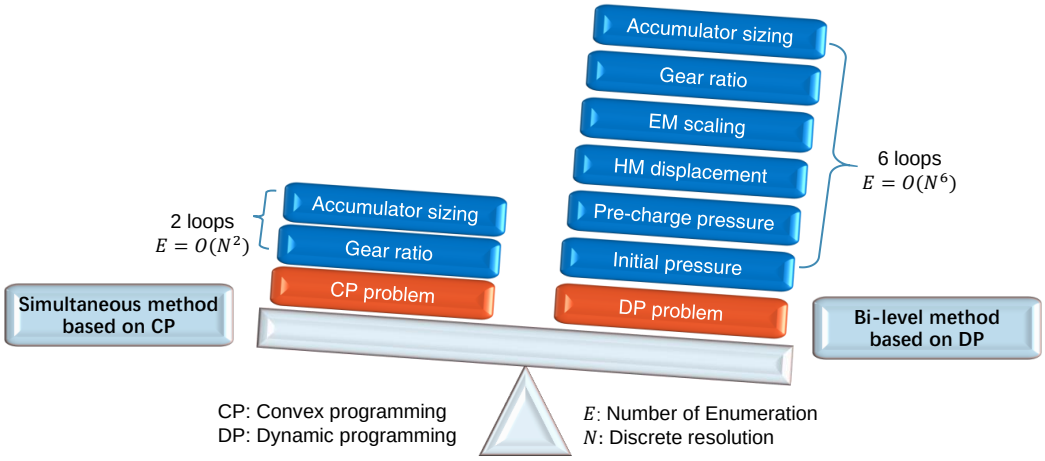


Figure 9: **C**omparison between the proposed simultaneous and DP-based bi-level optimization methods.

The DP-based bi-level method and CP-based simultaneous method were implemented and were run on a computer equipped with an i5-6500 CPU. Both the CP-based and DP-based methods adopts Latin hypercube sampling for the system parameter selection loop. For CP problem (11), we used CVX, a package for specifying and solving convex programs [13, 14]. For DP problem, a standard DP solver, DynaProg, is used to solve the optimal energy management problem [15].

The mean computing time is 3.5 seconds for each convex programming problem and 0.14 second for each DP problem. Considering the enumeration number, when $N \geq 3$, the convex programming method shows computational superiority. As shown in **Figure 10**, The computing time of DP-based bi-level method increases much faster than CP-based simultaneous method. With the same grid number of 7, the CP-based simultaneous method consumes 99% less computing time than DP-based bi-level method and provides the optimal solution with 1.5% less battery capacity loss (1.93×10^{6}% and 1.96×10^{6}%).

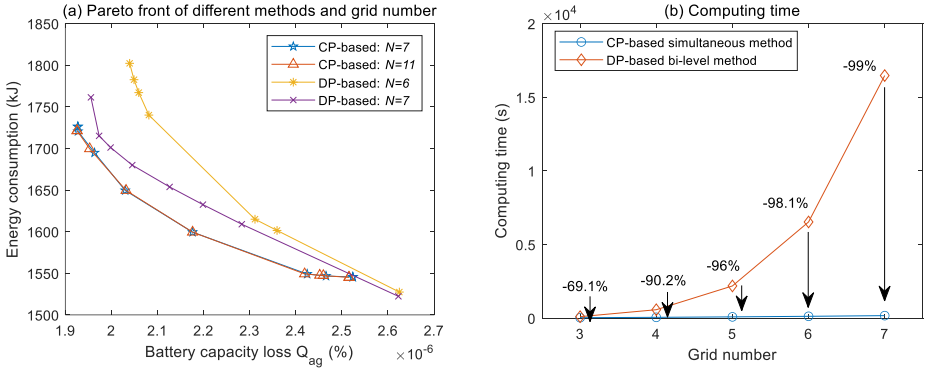


Figure 10: Comparison between CP-based simultaneous method and DP-based bi-level method (a) Pareto optimal solutions (b) Computing time.

6. CONCLUSION AND OUTLOOK

In this paper a system parameter optimization framework based on convex programming is proposed. It is applied on a parallel electric hydraulic hybrid wheel loader for the simultaneous optimization of its component sizing and energy management.

(1) PEHH extends battery life at the cost of energy consumption. The Pareto front of PEHH system parameter optimization problem is obtained. The HM size is the main factor of system performance trade-off between battery aging and energy consumption.

(2) PEHH can reduce the battery capacity loss in a loading cycle by 26.4% compared to the pure electric drive with a 13.0% increase of energy consumption.

(3) The convex programming method have superiority in terms of computational burden and digital resolution over the bi-level method. With the same grid number of 7, the CP-based simultaneous method consumes 99% less computing time than DP-based bi-level method and provides the optimal solution with 1.5% less battery capacity loss.

In the future, several aspects need to be studied. First, more application of this method can be studied such as series hydraulic hybrid powertrain and diesel-hydraulic hybrid powertrain. Secondly, an attractive topic is to incorporate the optimal vehicle speed and life cycle cost into the optimization framework, which will generate more valuable insights.

ACKNOWLEDGEMENT

This work was supported in part by the National Key R&D Program of China under Grant 2022YFE0128900, in part by the Zhejiang Provincial Natural Science Foundation under Grant LR22E050003, and in part by the National Natural Science Foundation of China under Grant 51875509.

REFERENCES

[1] Yang J, Liu B, Zhang T, et al (2022) Application of energy conversion and integration technologies based on electro-hydraulic hybrid power systems: A review. Energy Convers Manag 272:116372. https://doi.org/10.1016/j.enconman.2022.116372

[2] Niu G, Shang F, Krishnamurthy M, Garcia JM (2017) Design and Analysis of an Electric Hydraulic Hybrid Powertrain in Electric Vehicles. IEEE Trans Transp Electrif 3:48–57. https://doi.org/10.1109/TTE.2016.2628792

[3] Zhang H, Wang F, Xu B, Fiebig W (2022) Extending battery lifetime for electric wheel loaders with electric-hydraulic hybrid powertrain. Energy 261:125190. https://doi.org/10.1016/j.energy.2022.125190

[4] Qin T, Quan L, Li Y, Ge L (2021) A Dual-Power Coordinated Control for Swing System of Hydraulic-Electric Hybrid Excavator. In: 2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). IEEE, pp 55–60

[5] Tong Z, Jiang Y, Tong S, et al (2023) Hybrid drivetrain with dual energy regeneration and collaborative control of driving and lifting for construction machinery. Autom Constr 150:104806. https://doi.org/10.1016/j.autcon.2023.104806

[6] Huang Y, Wang H, Khajepour A, et al (2018) A review of power management strategies and component sizing methods for hybrid vehicles. Renew Sustain Energy Rev 96:132–144. https://doi.org/10.1016/j.rser.2018.07.020

[7] Frank B (2016) Using optimal control in concept evaluation and system optimization of diesel-electric hybrid construction machines. 2016 Int Conf Electr Syst Aircraft, Railw Sh Propuls Road Veh Int Transp Electrif Conf ESARS-ITEC 2016. https://doi.org/10.1109/ESARS-ITEC.2016.7841323

[8] Uebel K, Raduenz H, Krus P, De Negri VJ (2018) Design optimisation strategies for a hydraulic hybrid wheel loader. BATH/ASME 2018 Symp Fluid Power Motion Control FPMC 2018 1–11. https://doi.org/10.1115/FPMC2018-8802

[9] Murgovski N, Johannesson L, Sjöberg J, Egardt B (2012) Component sizing of a plug-in hybrid electric powertrain via convex optimization. Mechatronics 22:106–120. https://doi.org/10.1016/j.mechatronics.2011.12.001

[10] Li Y, Tang X, Lin X, et al (2022) The role and application of convex modeling and optimization in electrified vehicles. Renew Sustain Energy Rev 153:111796. https://doi.org/10.1016/j.rser.2021.111796

[11] Zhang H, Wang F, Xu B, Fiebig W (2022) Extending battery lifetime for electric wheel loaders with electric-hydraulic hybrid powertrain. Energy 261:125190. https://doi.org/10.1016/j.energy.2022.125190

[12] Costa GK, Sepehri N (2018) Understanding overall efficiency of hydrostatic pumps and motors. Int J Fluid Power 19:106–116. https://doi.org/10.1080/14399776.2018.1476306

[13] Grant M, Boyd S (2008) Graph implementations for nonsmooth convex programs. In: Blondel V, Boyd S, Kimura H (eds) Recent Advances in Learning and Control. Springer-Verlag Limited, pp 95–110

[14] CVX Research I (2012) {CVX}: Matlab Software for Disciplined Convex Programming, version 2.0

[15] Miretti F, Misul D, Spessa E (2021) DynaProg: Deterministic Dynamic Programming solver for finite horizon multi-stage decision problems. SoftwareX 14:100690. https://doi.org/10.1016/j.softx.2021.100690

CONTROL OF REAR-WHEEL STEERING FOR A FOUR-WHEEL STEERED AGRICULTURAL STANDARD TRACTOR

peer reviewed

Ruben Hefele[1]*, Johannes Burth[1], Michael Maier[1], Timo Oksanen[1,2]

[1]*Chair of Agrimechatronics, Technische Universität München, Dürnast 8, 85354 Freising*
[2]*Munich Institute of Robotics and Machine Intelligence (MIRMI), Technische Universität München, Georg-Brauchle-Ring 60-62, 80992 München*

* Corresponding author: Tel.: +49 (8161) 71-4903; E-mail address: ruben.hefele@tum.de

ABSTRACT

Steering is crucial for automation in agriculture. Four-wheel steered tractors are better suited for highly automated agricultural machinery compared to front-wheel steered tractors. Common tractors with four-wheel steering have equal large tire diameters on the front and rear axles. In comparison, this study uses a standard tractor, which means that the tires of the rear axle have a larger diameter than the tires of the front axle, thus significantly affecting the rear-axle steer control. An off-the-shelf front-wheel steering system is used, while the rear-wheel steering control is built from scratch. Steady-state accuracy and fast convergence are the control objectives. A double-acting differential cylinder is used as a steering actuator, supplied by a current controlled proportional directional valve with feedback from an angle sensor. The tractor's load sensing variable displacement pump provides hydraulic power. A SIL2 functional safety variant electric control unit with redundancy is used for the rear-axle controller. In-vehicle network communication is via CAN bus based on SAE J1939. System identification and modelling of the steering system are carried out, followed by controller setup and tuning using first order lag plus integrator plus delay rules while considering dead-zone and valve saturation. Further on, this paper introduces two novel controller designs, utilizing gain scheduling and model-based state space control techniques, which lead to enhanced performance.

Keywords: four-wheel steering, servo hydraulic control, system identification, highly automated agricultural machinery, FOLIPD process, gain scheduling , 2DOF control, PID control

1. INTRODUCTION

Highly automated tractors have emerged as a promising solution in modern agricultural practices, providing efficient and precise operations without relying on human intervention. These advanced machines combine the power of automation, navigation systems, and sophisticated control algorithms to carry out various tasks. Among the critical aspects of achieving effective autonomous operation, precise control over the tractor's heading plays a pivotal role in ensuring accurate trajectory tracking and manoeuvrability.

In many agricultural scenarios, especially on sloping terrains, maintaining the desired heading of the tractor and implement becomes particularly challenging due to inherent disturbances and variations in ground conditions. Conventional tractors typically rely solely on front axle steering mechanisms, resulting in limited control capabilities and reduced stability, especially under side slip conditions. However, integrating rear-axle steering into the control system provides an innovative solution to these challenges, allowing dynamic control of the tractor's heading and compensating for side slip at slopes. Likewise, when considering mounted implements, it is important to consider the precise tracking of paths in curves and the kinematic constraints to avoid placing excessive mechanical strain on the mounting links.

In our research-tractor project “AMX G-trac”, we have modified and customized a four-wheel steerable standard tractor to perform in future highly automated field operations. While the base tractor is built by Traktorenwerk Lindner, Austria, the rear-wheel system was redeveloped to enable more accurate steering servo control than the original system was able to perform. The new system is based on the original valves and a Parker IQAN mobile control system. Newly developed control algorithms compensate for system nonlinearities and provide smooth actuation. In contrast, the front-wheel steering is based on the Danfoss electro-hydraulic steering valve.

The paper presents two system identification methods in section **3.1** and **3.2**. While the first is a state of the art method, the second is a novel proposal, representing the system to a more precise extend. In the subsequent stage a systematic design approach for a PD controller with first order lag plus integrator plus delay (FOLIPD) tuning rules (section **4.1**), a two-degree-of-freedom (2DOF) controller combining model-based state space control and PD control (section **4.2**) as well as a gain scheduled PD controller (section **4.3**) is described. The gain scheduling controller shows better performance.

2. ELECTROHYDRAULIC STEERING SYSTEM

The problem analysed in the present work is the steering servo control of a tractor for future autonomous operation. **Figure 1** depicts the composition of the steering system, comprising electric control units (ECUs) and sensors. The hydraulic system of the tractor is based on a load-sensing circuit. Double-acting differential cylinders are actuated through directional proportional valves in order to realize different steering angles.

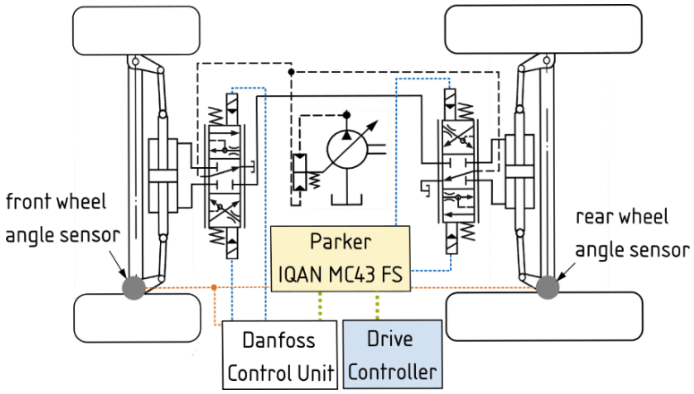


Figure 1: Simplified steering system

For control of the front-axle steering the built-in Danfoss system is utilized via CAN bus-based J1939 communication. To simplify the illustration, the manual steering system is omitted from **Figure 1**.

The Parker IQAN controller serves as a gateway between the main drive controller and the tractor system to control both, rear and front steering angles, which enables functional safe operation. The IQAN MC43 FS is a rugged master controller providing multiple CAN bus channels, analog, digital and PWM in- and outputs as well as SIL2 safety features.

Programming of the IQAN is facilitated by IQANdesign, a high-level graphical design tool that also supports MATLAB Simulink model import. The IQAN controller is directly used for rear-axle control, accomplished through two PWM outputs that command the directional proportional valve. Additionally, lock valves (which are not displayed) are controlled by the IQAN. The analog voltage input from the rear wheel angle sensor serves as control feedback. The IQAN has a built-in PWM current controller, delivering a precision of 1 mA, while the operating range is from roughly -2500 to 2500 mA. To calibrate the wheel angle sensors, accurate values for the true steering angle were

measured using a total station and subsequently mapped to the sensor voltage outputs within IQANdesign.

3. PROCESS MODEL AND SYSTEM IDENTIFICATION

In the present work a FOLIPD process model (1) is used for modeling the hydraulic rear steer plant. Literature shows that this model is acceptably accurate to approximate the servo-valve dynamics [1]. Utilizing the FOLIPD process model is also suggested in [2–5].

$$G(s) = \frac{1}{s} \cdot \frac{K_v}{1 + T_F\, s} e^{-sL} \tag{1}$$

The transfer function consists of an integrator part, an open-loop transform function between valve current input and cylinder speed output as well as an input-output time delay. K_v represents the process gain, T_F the time constant and L the process time delay.

To ensure consistent results in system identification, the tractor was lifted so that the tires didn't have ground contact while measurements were taken. This was done to maintain a constant load profile and under the assumption that the tractor's steering behavior during this setup closely resembles its steering behavior during actual field operations. Measurements were taken with a sample rate of 50 ms.

3.1. Model identified from open-loop step response

While the identification can be performed utilizing an impulse response [5], in the first approach of the present work for model identification a step response procedure is used as suggested in [6].

First, the dead zone of the valve must be determined using a ramp input signal as presented in **Figure 2**. It can be expressed as follows:

$$u_{dz}(t) = \begin{cases} u(t), & if\ u(t) \geq u_{dz,+} \\ 0, & if\ u_{dz,-} < u(t) < u_{dz,+} \\ u(t), & if\ u(t) \leq u_{dz,-} \end{cases} \tag{2}$$

Where negative dead zone limit $u_{dz,-} = -850\ mA$ and positive limit $u_{dz,+} = 965\ mA$.

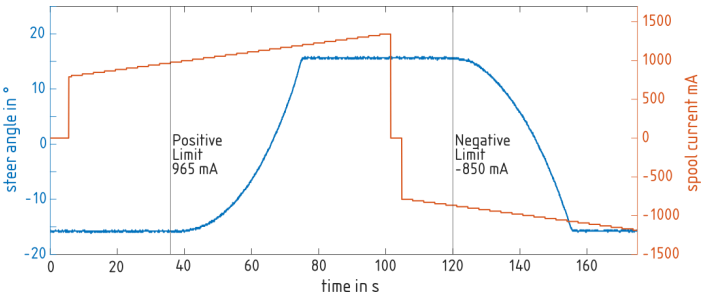


Figure 2: Dead zone investigation

In the subsequent stage, the nonlinear process gains are determined. In order to achieve plug-and-play tuning between the tractor and implement, Oksanen [6] suggests increasing the step in increments of 10%. However, since there is no need for plug-and-play functionality in the rear-axle control and a more precise resolution of the nonlinear process gains is preferred, we have chosen to use increments of 2%.

Figure 3 illustrates the step input with 2% increments, starting from the dead zone and rising to a maximum of 2500 mA, which corresponds to 100% spool current. By determining the steady-state speed for each operating point, **Figure 4** is generated. This figure considers both the dead zone and the saturation, with a negative saturation level of 2386 mA and positive saturation level of 2234 mA.

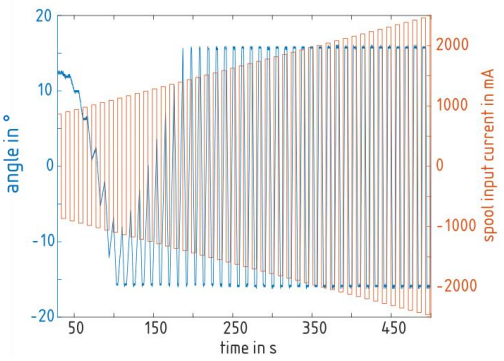


Figure 3: Spool current and steer angle

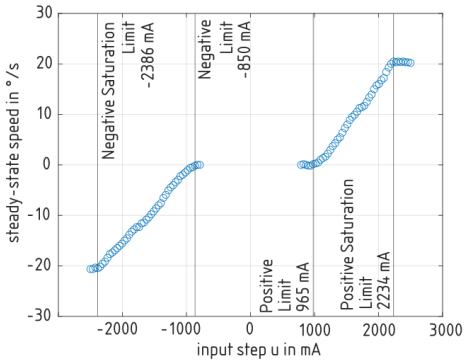


Figure 4: Steady-state speed versus spool current

As observed in **Figure 3**, it is evident that the allowable steering angle is confined within a range of 16 degrees in both the positive and negative directions.

The time delay L and time constant T_F are identified from six step responses (three positive, three negative) and averaged. An exemplary step response for integrating processes and the time delay L and time constant T_F is presented in **Figure 5**.

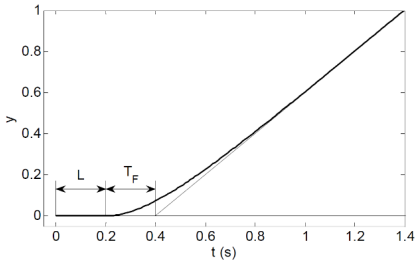


Figure 5: Step response for integrating process with delay L and time constant T_F [6]

Chosen operating points were 55 %, 75 % and 95 % of respective saturation limit to calculate the averaged parameters.

(3) presents the identified model:

$$G(s) = \frac{1}{s} \cdot \frac{K_v}{1 + 0.0385\,s} e^{-0.2658\,s} \tag{3}$$

The process gain K_v is non-constant due to the nonlinear process, which is why averaging K_v would not adequately reflect the process. In the MATLAB Simulink model a look-up table is used between plant and transfer function input to represent the nonlinear process gain incorporating both the dead zone and the saturation, presented by **Figure 6** which is a polished version of **Figure 4.** This means

that any currents exceeding the saturation limit have been eliminated, and currents within the dead zone have been set to zero.

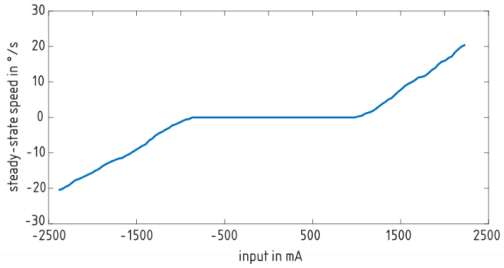


Figure 6: Look-up table for nonlinear process gain K_v including dead zone and saturation

3.2. Greybox modelling with time-varying delay and time constant

In an alternative modelling approach, greybox identification was investigated as proposed by Galuppini et al. [7]. They propose to inject a series of sinusoidal inputs at different frequencies to the plant. However, it was observed that the resulting transfer function deviated considerably from the transfer function stated in (3), indicating nonlinearities in addition to the dead zone and saturation. Consequently, a different input-output method was employed.

In this study, the greybox system was stimulated with various step inputs, and the measured angle was recorded as the output, similar to the procedure outlined in **Figure 3**. (1) served as the underlying model for this investigation. With 50 operating points in both the negative and positive directions, a total of 100 different transfer functions result, varying in process gain K_v, time constant T_F and delay L. **Figure 7** depicts the resulting varying process gain. A double exponential curve was fitted to the positive and negative directions, to obtain continuous monotone behaviour.

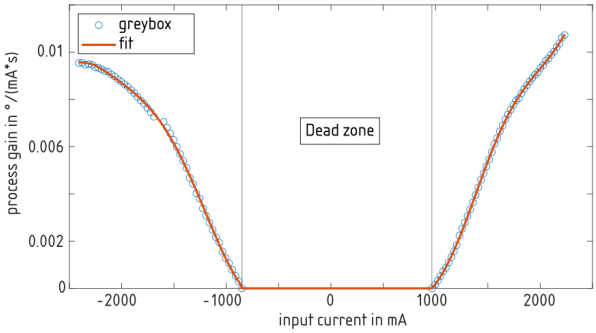


Figure 7: Time-varying process gain identified from greybox

Figure 8 and **Figure 9** present the time-varying time constant and delay identified with greybox modelling. The measurement values for the time constant exhibit considerable scattering, while the delays demonstrate exponential growth towards the dead zone with less scattering. To account for this, an averaged value is adopted for the time constant (**Figure 8**), distinguishing between positive and negative spool displacements. Within the dead zone, a linear interpolation is employed, as no significant impact on the simulated valve behaviour is observed.

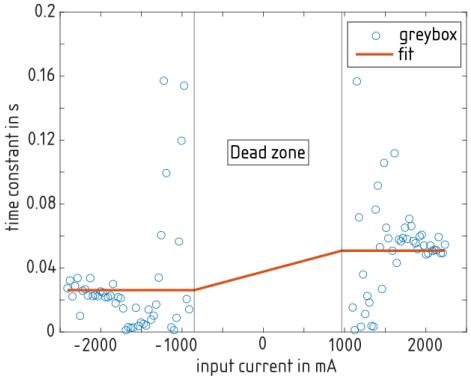


Figure 8: Time-varying time constant identified from greybox

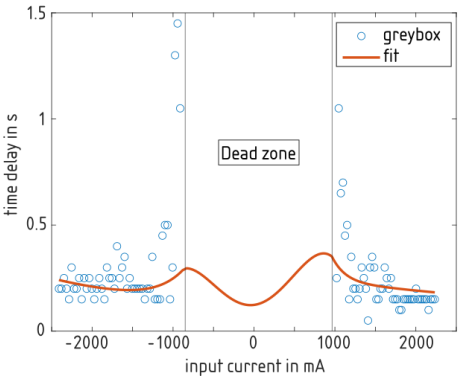


Figure 9: Time-varying delay identified from greybox with fitted polynomial to dead zone to consider the delay at 0 mA while maintaining continuous behaviour

For the delay (**Figure 9**), double exponential curves are fitted to both positive and negative spool displacements outside of the dead zone, ensuring a continuous and monotonic behaviour. In contrast, the delay for the transition to 0 mA currents differs significantly from delays for currents outside of the dead zone. The measured delay for transitions to 0 mA is 0.17 s. To incorporate this into the model and maintain continuous behaviour, a fourth-degree polynomial is fitted to the dead zone, with a value of 0.17 s at 0 mA and a smooth transition at the boundaries of the dead zone. The smooth transition is obtained by taking into account the function value and its derivation at the boundaries. A polynomial of degree four yields the most accurate fitting results.

3.3. Model evaluation

With adding random Gaussian noise to both models, which has similar characteristics as the position sensor of the rear-axle steering (max. 0.1 °), simulation results presented in **Figure 10** and **Figure 11** are obtained for a step and sinusoidal input signal. The simulated results are compared to the measurement of the real steer plant again with lifted tractor rear axle.

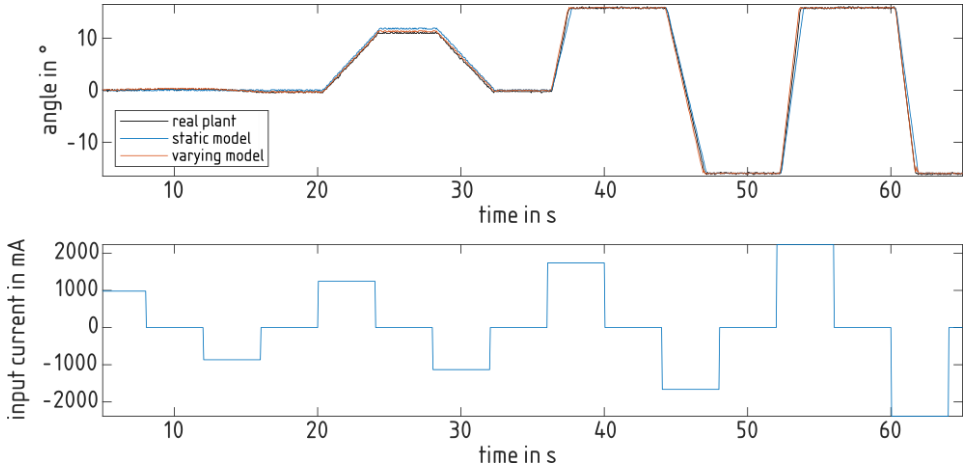


Figure 10: Simulation results for static and varying model for different step inputs

It can be observed that the model incorporating varying delay and varying time constant (section **3.2**) approximates the real steer process with better accuracy. This is particularly noticeable during the

step input signal for low currents, as well as during transitions to a 0 mA current when the steering angle is not restricted.

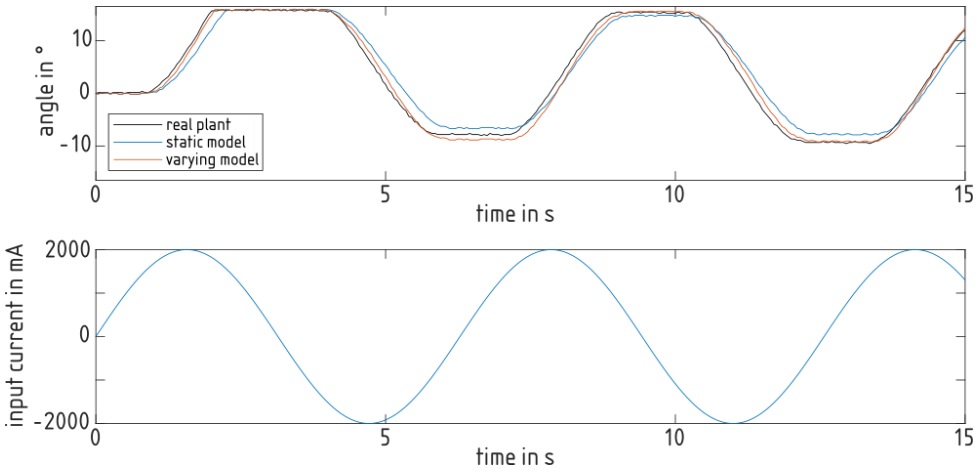


Figure 11: Simulation results for static and varying model for sinusoidal input

4. CONTROLLER DESIGN

In the following subsections, the architecture and parameter determination of different controllers is presented. There are: a PD controller tuned with FOLIPD tuning rules, a 2DOF state space controller combined with PD control and a PD gain scheduling controller.

To obtain a smoother feedback signal, a linear Kalman filter is utilized due to the presence of significant noise in the input signal received from the sensor, as presented in **Figure 12**.

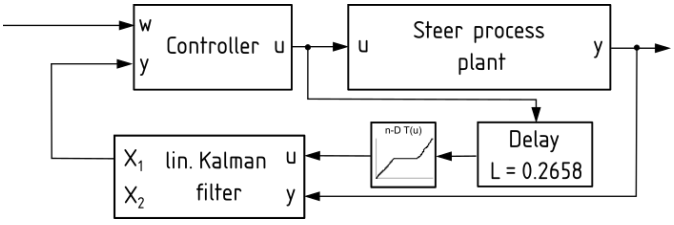


Figure 12: Implementation of Kalman filter to the steer process

Whereby set point w is steer angle, the control variable u is current and feedback y is the measured steer angle. The model of the Kalman filter is in a continuous state space representation (4) with the states position (X_1) and velocity (X_2):

$$A = \begin{pmatrix} 0 & 1 \\ 0 & -1/0.0385 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1/0.0385 \end{pmatrix}, \quad c = (1 \quad 0). \tag{4}$$

The nonlinear process gain (**Figure 6**), which accounts for the nonlinear behavior and dead zone, has been incorporated as a cascade block preceding the linear Kalman filter in order to accurately capture these characteristics. Additionally, to account for the delay introduced by the steer plant, a delay approximation has been incorporated before the Kalman filter (**Figure 12**).

4.1. PD controller based on FOLIPD tuning rules

Eriksson and Oksanen [4] propose a robust linear PID controller for FOLIPD processes with

unknown varying delay. Their proposed tuning rules are:

$$k = \frac{10^{f(L,T_F)}}{K_v L}, k_i = 0, k_d = \frac{T_F^{g(L,T_F)}}{K_v} 10^{h(L)}, \tag{5}$$

where

$$\begin{aligned} f(L,T_F) &= 0.0027\left(\frac{T_F}{L}\right)^2 - 0.0794\frac{T_F}{L} - 0.34, \\ g(L,T_F) &= 0.02 + (0.51 - 0.076\log_{10}(T_F))L^{0.15}, \\ h(L) &= 0.97 - 1.48\,L^{0.15}. \end{aligned} \tag{6}$$

The PID controller becomes effectively a PD controller and is to be used as a linear controller. For our PD controller, (5) and (6) result in $k = 1.6749$ and $k_d = 0.1029$, while $K_v = 1$. In order to obtain a linear process with $K_v = 1$, the nonlinear process gain (**Figure 6**) is compensated with its inverse function shown in **Figure 13** as proposed in [6, 8]. In order to obtain a bijective function, which is necessary for stability reasons, the function is modified around the zero-crossing.
Figure 14 depicts the integrated inverse element and the full architecture of the PD controller.

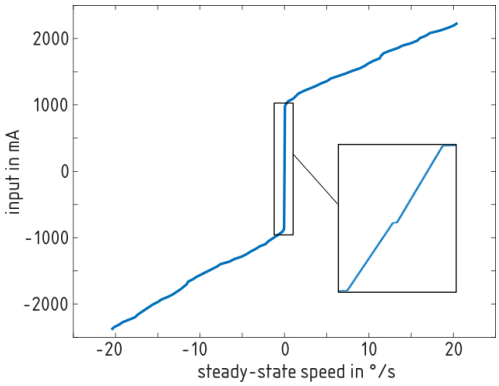


Figure 13: Inverse of nonlinear process function

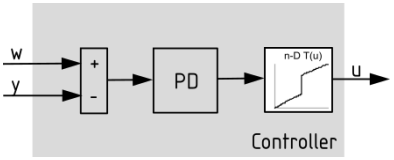


Figure 14: Architecture of PD Controller

4.2. Two degree of freedom state space controller combined with PD control

A second controller structure which has been studied is a 2DOF state space controller combined with a PD controller, presented in **Figure 15**:

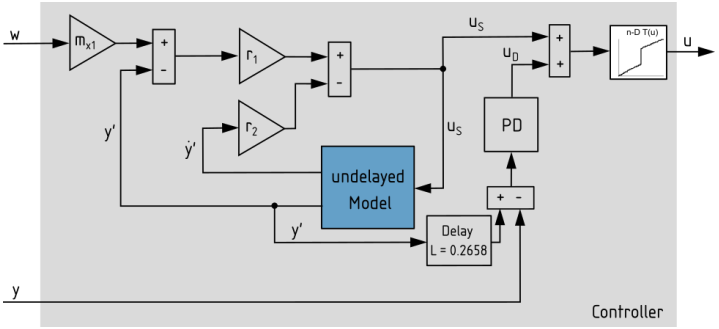


Figure 15: Architecture of 2-DOF controller comprising state space and PD controller

The state space controller exclusively operates on the undelayed model. For this, the model described in section **3.1** is utilized, omitting the delay. However, to address the discrepancy arising from the delay between the actual plant and the undelayed model, a PD controller, identical to the one

described in section **4.1**, is employed. The combination of both control signals is then adjusted with the inverse element explained in section **4.1**, ultimately yielding the control signal required for the steer plant.

According to Föllinger [9] the controller parameters are determined with (7-11). m_{x1} is calculated by applying (7):

$$\begin{pmatrix} m_{x1} \\ m_{x2} \\ m_u \end{pmatrix} = \begin{pmatrix} A & b \\ c^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ I \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{7}$$

with

$$A = \begin{pmatrix} 0 & 1 \\ 0 & -1/0.0385 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1/0.0385 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 0 \end{pmatrix}. \tag{8}$$

The controller values r^T are calculated using following equations and pole placement [9]:

$$(s - \lambda_{R1}) \dots (s - \lambda_{Rn}) = s^n + p_{n-1}s^{n-1} + \dots + p_0 \tag{9}$$

$$t_1^T = e_n Q_s^{-1} \tag{10}$$

$$r^T = p_0 t_1^T + p_1 t_1^T A + \dots + p_{n-1} t_1^T A^{n-1} + t_1 A^n \tag{11}$$

Where λ are the Eigenvalues of the closed loop control system of the model, and Q_s the controllability matrix. In our case $n = 2$ is valid. From tuning, it resulted that with $\lambda_{R1} = -3.3$ and $\lambda_{R2} = -50.0$ the best control results are obtained.

4.3. Gain scheduling PD controller

A third controller for FOLIPD processes with time-varying gain, delay and time constant being tested, as described in the introduction, is a gain scheduled controller. Gain scheduling is commonly used for controlling nonlinear processes. The classical gain scheduling approach uses a family of controllers for multiple operating points [10]. An observable scheduling variable is used to determine the right linear controller depending on the operating point.

However, in our process, the only feedback available is the steering angle signal, which is not suitable as a scheduling variable, because of the integrator. Since the varying transfer function parameters, e. g. delay, are dependent on the actual current applied to the valve, the valve current control signal is used as scheduling variable, which results in an inherent problem. To overcome this, we propose below a procedure to determine the most appropriate controller for each operating point, despite the inherent problem.

For the controller design, a family of proportional gains k and the derivative gains k_d depending on the valve input current u_{sv} (scheduling variable) are calculated from (5). In our case, we used 100 different tuning sets for the PD controller. This represents a compromise between accuracy and the required computing power. The parameters for gain, time constant and delay result from the model described in section **3.2**.

The controller itself is then structured as follows:

1. Calculation of a vector of theoretically resulting valve input currents u_{th} based on the error e between the setpoint and feedback:

$$u_{th} = ke + k_d \dot{e} \tag{12}$$

2. u_{isp} results from finding the intersection point between u_{sv} and u_{th}.
3. To compensate dead zone and saturation, the following filter is applied:

$$u(t) = \begin{cases} u_{max}, & if\ u_{isp} > u_{max} \\ u_{isp}, & if\ u_{dz,+} \leq u_{isp} \leq u_{max} \\ u_{dz,+}, & if\ 0 < u_{isp} < u_{dz,+} \\ 0, & if\ u_{isp} = 0 \\ u_{dz,-}, & if\ u_{dz,-} < u_{isp} < 0 \\ u_{isp}, & if\ u_{min} \leq u_{isp} \leq u_{dz,-} \\ u_{min}, & if\ u_{isp} < u_{min} \end{cases} \quad (13)$$

With u_{max} being positive saturation limit, u_{min} negative saturation limit, $u_{dz,+}$ positive dead zone limit and $u_{dz,-}$ negative dead zone limit.

The controller gain tuning sets vary for k between roughly 160 and 8700 and for k_d between 8 and 1700. Results showed that the controller performance increases with higher values for the controller gains. Hence, k is multiplied by a factor of 1.3 and k_d by 1.4.

5. SIMULATION AND MEASUREMENT RESULTS

The controllers were implemented to the IQAN using the Simulink export for IQANdesign. In the following the control results from simulation and measurement are presented. **Figure 16** depicts the results for step responses with the varying simulation model described in section **3.2**.

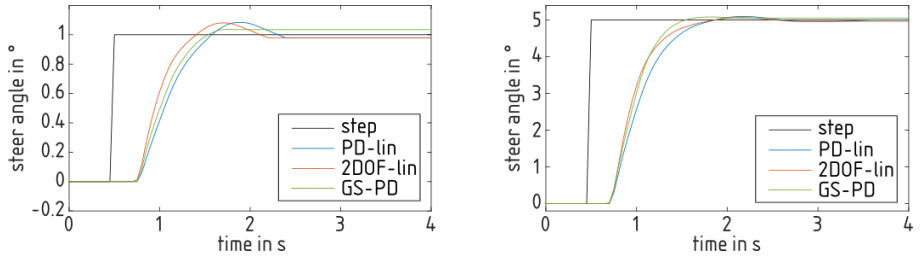


Figure 16: Step responses of controllers with varying simulation model from section 3.2

The PD controller and 2DOF controller with a linearized plant demonstrate superior steady-state accuracy, with $e_\infty = -0.02\,°$, when compared to the gain scheduling PD controller with $e_\infty = 0.03\,°$ for a step of 1 °. However, when it comes to overshoot, the gain scheduling controller outperforms the others. In terms of rise time, the 2DOF controller proves to be the fastest for a step of 1°, while the gain scheduling controller exhibits less rise time for a step of 5 °. The gain scheduling controller excels in terms of settling time for both steps.

When adding noise to the simulation model, especially the overshoot criteria changes for all of the controllers as shown in **Figure 17**.

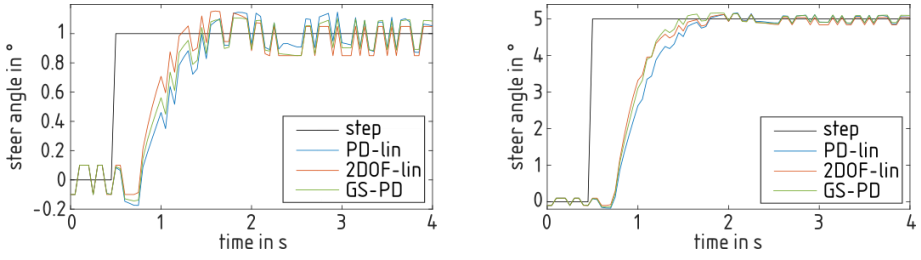


Figure 17: Step responses of controllers with varying noisy simulation model from section 3.2

Figure 18 depicts the unfiltered results obtained when the controllers are implemented to the actual plant. The PD controller outperforms the others in both step scenarios, although no noticeable distinction can be observed between the controllers. For a step of 1 ° the 2DOF controller shows higher rise time especially compared to the pure PD controller.

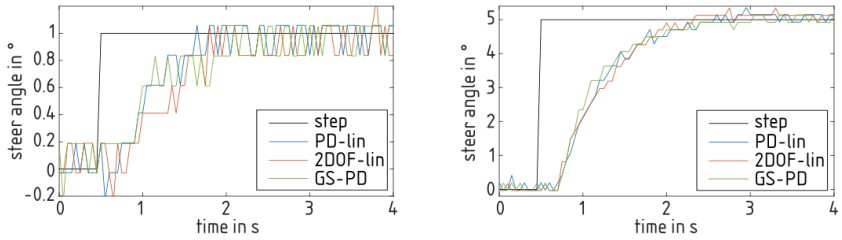


Figure 18: Measured unfiltered step responses of controllers at true steer plant

Figure 19 presents the comparison between simulation and measurement for the step responses of the gain scheduling PD controller. The difference in response time between the true plant and the noisy model is evident. Although the model exhibits some slight overshooting, such behavior is not observed in the measured data. This suggests that the controllers for the true plant could potentially be designed to operate even faster.

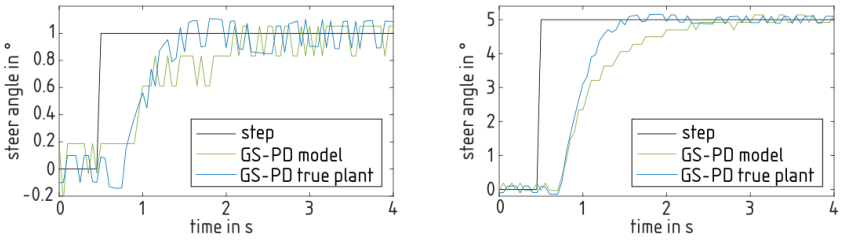


Figure 19: Simulated versus measured step responses (unfiltered) for gain scheduling PD control

Controller performance parameters are summarized in **Table 1** for the simulated response and **Table 2** for the measured response.

Table 1: Controller performance parameters from simulation with time-varying model without noise

Controller at step	Overshoot [%]	Settling time [s]	e_∞ [°]
PD-lin at 1 °	10.8	2.32	-0.022
2DOF-lin at 1 °	10.4	2.12	-0.021
GS-PD at 1 °	0.2	1.58	0.035
PD-lin at 5 °	1.9	1.72	0.002
2DOF-lin at 5 °	2.0	2.19	0.034
GS-PD at 5 °	0.6	1.49	0.050

Table 2: Controller performance parameters measured at true plant

Controller at step	Overshoot [%]	Settling time [s]	e_∞ [°]
PD-lin at 1 °	0.0	1.63	-0.086
2DOF-lin at 1 °	0.0	1.71	-0.146
GS-PD at 1 °	0.0	2.05	-0.080
PD-lin at 5 °	0.0	2.55	-0.006
2DOF-lin at 5 °	0.0	2.32	-0.218

GS-PD at 5 °	0.0	2.45	0.000

6. CONCLUSION AND OUTLOOK

To conclude, this paper presents two modelling approaches, with the new proposal containing varying delay and time constant proving to be better than the one with constant delay and time constant. In simulation, the gain scheduling exhibits the best results. However, in measurement scenarios, all controllers are operating similarly, with the PD controller and linearized plant performing slightly better than the others. All the controllers are stable.

It is worth mentioning that the model itself can still be improved. Furthermore, by utilizing a better model, the controllers can be made faster since the controllers highly rely on the accuracy of model parameters. Especially the gain scheduling controller could be improved with dedicated tuning rules.

Looking ahead, future research can focus on enhancing the proposed model and exploring alternative methods for improving controller performance, allowing for even more efficient and reliable control systems.

For our research tractor all the three presented control algorithms compensate system nonlinearities and provide smooth actuation of the rear-axle steering. Hence, all of them could be used for the “AMX G-trac” in future autonomous operation.

REFERENCES

[1] W. J. Thayer (1958 (revised in 1965)) Transfer function for MOOG servovalves. Controls Division, East Aurora, NY. technical bulletin

[2] Moosavi H, Mirza Hessabi R, Mercan O (2015) Numerical simulation and analysis of nonlinear state-space control design for hydraulic actuator control. Struct Control Health Monit 22:1068–1085. https://doi.org/10.1002/stc.1731

[3] Krata D, Ochman M, Panek M et al. (2021) Adaptive Smith Predictor Control Scheme for a Nonlinear Hydraulic System. In: 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, pp 1–6

[4] Eriksson L, Oksanen T, Mikkola K (2009) PID controller tuning rules for integrating processes with varying time-delays. Journal of the Franklin Institute 346:470–487. https://doi.org/10.1016/j.jfranklin.2009.01.003

[5] Możaryn J, Winnicki A, Suski D (2019) Impulse Identification and Discrete P/PD Control of Electro-Hydraulic Servodrive. JAMRIS 12:51–57. https://doi.org/10.14313/JAMRIS_4-2018/25

[6] Oksanen T (2010) Closed loop control over ISO 11783 network – challenges of plug-and-play. IFAC Proceedings Volumes 43:203–208. https://doi.org/10.3182/20101206-3-JP-3009.00035

[7] Galuppini G, Magni L, Raimondo DM (2018) Model predictive control of systems with deadzone and saturation. Control Engineering Practice 78:56–64. https://doi.org/10.1016/j.conengprac.2018.06.010

[8] Deng W, Yao J, Ma D (2017) Robust adaptive precision motion control of hydraulic actuators with valve dead-zone compensation. ISA Trans 70:269–278. https://doi.org/10.1016/j.isatra.2017.07.022

[9] Föllinger O (2022) Regelungstechnik. VDE Verlag

[10] Leith DJ, Leithead WE (2000) Survey of gain-scheduling analysis and design. International Journal of Control 73:1001–1025. https://doi.org/10.1080/002071700411304

OPTIMAL SPEED TRAJECTORY OF ELECTRIC WHEEL LOADERS AIMING AT EXTENDING BATTERY LIFETIME

peer reviewed

Haoxiang Zhang[1], Qi Zhang[1], Jiajia Wang[1], Yihan Qiao[1], Feng Wang[1]*, Bing Xu[1]

[1]*School of Mechanical Engineering, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang 310027, P. R. China*
** Corresponding author: E-mail address: dieter@zju.edu.cn*

ABSTRACT

The electrification of wheel loaders is considered a leading trend due to its advantage of zero-carbon emissions. However, the inevitable phenomenon of battery degradation has led to increased battery usage and maintenance costs. This study first extends the battery lifetime by optimizing the speed trajectory based on the typical loading cycle of the wheel loader. The optimal control problem is formulated by systematically modelling the wheel loader's powertrain and using a precise semi-empirical battery aging model. To reduce computational costs, the modified optimal control problem includes a weighted penalty on travel time. A combined algorithm of dynamic programming and Brent's method (DP-BM) is introduced to provide a numerical solution to the optimization problem with a reduced computational burden. Simulation results demonstrate that the optimized trajectory can decrease the average power consumption of the battery and reduce the number of full equivalent cycles, resulting in a 4.48% improvement in the average battery lifetime compared to the typical trajectory. Furthermore, the proposed approach significantly reduces computation time compared to the conventional dynamic programming method, with an average reduction of 95%.

Keywords: Speed trajectory optimization, Dynamic programming, Brent's method, Battery lifetime, Electric wheel loader

1. INSTRUCTION

The electric wheel loader has been gradually rolling out in recent years due to its advantage of zero-carbon emission. However, the inevitable battery degradation phenomenon influences the battery lifetime and hinders the widespread application of the electric wheel loader. The high load and high inertia characteristics of the electric wheel loader result in a high average discharge current, exacerbating battery degradation. Additionally, the wheel loader is usually designed for a single application such as loading material, meaning that the driving path is relatively short and fixed. The high repetition of the duty cycle causes the battery to charge and discharge frequently, which further accelerates battery degradation.

Aiming at extending battery lifetime, this paper attempts to find the optimal speed profile of electric wheel loaders in the typical loading cycle. The optimization problem can be defined as follows: given the target route and duration time of the vehicle, calculate the speed at each moment to form a velocity trajectory, aiming at maximizing the battery lifetime.

Currently, there is no work for either wheel loaders or on-road vehicles on extending battery lifetime through speed trajectory optimization. However, some studies around speed optimization focus on minimizing energy consumption, which can provide references to this paper. Mello and Bauer optimize speed trajectories between stops for electric vehicles by considering real-world constraints such as following distance to the next vehicle and jerk [1]. Wu et al. [2] address similar issues and focus on avoiding stops in arterial corridors considering the impact of the presence of

intersection queues in both temporal and spatial dimensions for an electric vehicle. In addition, some works focus on the speed profile of trams [3, 4]. The specific operating scenario gives the optimization problem extra constraints such as precise green signal time windows.

Numerical methods in dealing with trajectory optimization problem greatly affect the calculation efficiency and accuracy. Dynamic programming has been widely and extensively used because it can find a global optimum even for nonlinear systems with nonlinear constraints. However, it is difficult to obtain an accurate solution within an acceptable time. Since the operating distance and duration vary for each loading cycle, the optimization problem must be addressed frequently, which puts forward higher requirements on the algorithm's efficiency. The conventional dynamic programming algorithm takes substantial computation time and must be improved.

The contribution of this study is to extend battery lifetime of an electric wheel loader through speed trajectory optimization. To achieve this target, dynamic programming method is employed to give a numerical and optimal speed-versus-time trajectory that satisfies the given constraints. In order to reduce the computational cost, the optimal control problem is modified such that a weighted penalty on travel time is included. Brent's method is first used to calculate the weighted penalty (achieve time constraint in the cost function). The dynamic programming combining Brent's method (DP-BM) proposed in this paper provides a very useful tool that can calculate the global optimal speed trajectory with a reduced computational burden. The significant saving of operation time provides the possibility for the application of trajectory optimization in real working cycles.

2. MATHEMATICAL MODELING OF ELECTRIC WHEEL LOADER

2.1. System Overview

The conventional wheel loader generally uses one internal combustion engine that simultaneously drives the hydraulic pump and the drive shaft. Taking advantage of the electric motor's small size and flexibility in arrangement, the distributed solution with two independent electric motors that separately drive the working hydraulic system and propulsion system is widely used in the electric wheel loader. **Figure 1** shows the powertrain architecture analyzed in this study. The two systems are powered simultaneously by one battery pack, then achieve their respective functions.

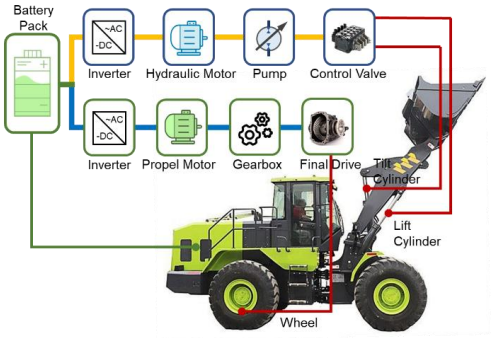


Figure 1: Powertrain of the electric wheel loader

2.2. System Modelling

Drivetrain

The time-based longitudinal dynamics of a wheel loader can be described as:

$$\begin{aligned} \dot{s} &= v \\ \dot{v} &= \frac{1}{m}\left(F_t + F_b - F_r\right) \end{aligned} \tag{1}$$

where m is the operating mass of the wheel loader, v is the wheel loader speed, s is the position of the wheel loader. F_t, F_b and F_r are the motor traction force, electrical braking force and resistance force respectively.

The basic resistance force of an operating wheel loader includes the aerodynamic friction, the rolling friction, and the thrust force, as given by:

$$F_r = \mu mg + \frac{1}{2}\rho v^2 C_d A_v + F_{th} \tag{2}$$

where μ is the rolling friction coefficient, g is the gravitational acceleration, ρ is the air density, C_d is the aerodynamic drag coefficient, A_v is the frontal area of the vehicle body.

The relationship between electric motor torque T_m and the motor traction force F_t is:

$$T_m = \frac{r_w}{\gamma_g \gamma_f} F_t \tag{3}$$

where r_w is the wheel radius, γ_g is the gear ratio of gearbox, and γ_f is the final drive ratio.

Electric motor

The electric motor power in the steady-state condition, P_m, is given by:

$$P_m = T_m \omega_m \eta_m^{-\mathrm{sgn}(T_m)} \tag{4}$$

where ω_m is the motor speed and η_m is the motor efficiency. The motor outputs mechanical power if $T_m > 0$ and generates electric power if $T_m < 0$.

Battery

The battery pack is developed by an equivalent circuit model composed of a voltage source U_{oc} and a resistance R_b accounting for Joule losses. The battery current can be derived as:

$$i_b = \frac{U_{oc} - \sqrt{{U_{oc}}^2 - 4R_b P_b}}{2R_b} \tag{5}$$

The battery C-rate is defined as:

$$I_c = \frac{|i_b|}{Q_b} \tag{6}$$

where Q_b is the capacity of battery and $Q_b = Q_{init} - Q_{loss}$. Q_{init} is the battery initial capacity at the manufactory (nominal capacity) and Q_{loss} is the total capacity loss.

The battery SOC accounts for the current battery capacity and its dynamic model is given by:

$$\dot{SOC}_b = -\frac{i_b}{Q_b} \tag{7}$$

The equivalent voltage source has a nonlinear connection with the battery SOC.

3. SPEED TRAJECTORY OPTIMIZATION

3.1. Typical Loading Trajectory

The commonly used movement trajectory for wheel loaders to carry material from material pile to truck is the V-type loading cycle (V-cycle) [5]. The V-cycle consists of six operating phases: bucket filling, backward 1, forward 1, dumping, backward 2, and forward 2. The speed trajectory optimization in this study focuses on the forward/backward phases, where the short moving in bucket filling and dumping phases are not part of the research. In addition, the prerequisite of optimizing the speed trajectory for a wheel loader is to determine a moving route. In this study, the forward/backward routes are taken as known conditions for optimizing the speed trajectory.

3.2. Battery Aging Model

For battery lifetime optimization, the first step is to set up an accurate and useful battery aging model. In this study, an accurate semi-empirical model from Naumann's previous works [6, 7] is adopted. The battery degradation contains cycle aging capacity loss Q_{cyc} and calendar aging capacity loss Q_{cal}:

$$Q_{loss} = Q_{cyc} + Q_{cal} \tag{8}$$

The expressions of Q_{cyc} and Q_{cal} are:

$$\begin{cases} Q_{cyc} = k_c k_{SOC_cyc} FEC^z \\ Q_{cal} = k_T k_{SOC_cal} \sqrt{t} \end{cases} \tag{9}$$

where k_c and k_{SOC_cyc} are the C-rate influence factor and the SOC influence factor for cycle aging capacity loss, $k_c = aI_c + b$ and $k_{SOC_cyc} = c(0.4 - SOC_b)^3 + d$. k_T and k_{SOC_cal} are the temperature influence factor and the battery SOC influence factor for calendar aging capacity loss, $k_T = k_{ref} \cdot \exp\left[k\left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right]$ and $k_{SOC_cal} = e\left(SOC_b - 0.5\right)^3 + f$. FEC is the full equivalent cycle and defined as the ratio of accumulated charge throughput to the current battery capacity, k_{ref} and T_{ref} are the reference aging rate and storage temperature, a, b, c, d, e, f, k and z are model based parameters and their specific values are referred to reference [6, 7].

3.3. Problem Formulation

The aim of the optimal control problem formulated and solved in this study is to minimize the loss of battery capacity. The primary step is developing a model that accurately measures battery deterioration and incorporates it into the cost function. The concept of severity factor, defined by Onori [8], is utilized to quantify the relative aging effect with respect to a nominal operating condition. The battery lifetime Γ with respect to a nominal cycle can be characterized by the total full equivalent cycle (total FEC) when the battery reaches its end of life, which is expressed as:

$$\Gamma = \int_0^{EOL} \frac{|I_{nom}(t)|}{2Q_b} dt \tag{10}$$

where $I_{nom}(t)$ is the current profile under nominal conditions. EOL represents the time when the battery life reaches its end, and at this time the battery capacity has dropped 20% from its initial

value. Since one FEC stands for a complete discharge process and a complete charge process, the total FEC is calculated as Ah-throughput divided by two times of battery capacity.

Then, the relative aging effects of any other loading cycle the battery is subject to can be reflected by the severity factor:

$$\sigma(I,T,SOC)=\frac{\Gamma}{\gamma(I,T,SOC)}=\frac{\int_0^{EOL}\left|I_{nom}(t)\right|dt}{\int_0^{EOL}\left|I(t)\right|dt} \tag{11}$$

where $\gamma(I,\theta,SOC)$ is the battery lifetime with specific operating conditions given in terms of current I, temperature T, and SOC. When the battery is undergoing a more severe load cycle, the severity factor is greater than one and a shorter life is expected.

In order to give the effective lifetime depletion due to charge exchange within the battery, an effective battery full equivalent cycle (effective FEC) FEC_{eff} is defined as:

$$FEC_{eff}(t)=\int_0^{t_f}\sigma(I,T,SOC)\cdot\frac{\left|i_b(\tau)\right|}{2Q_b}d\tau \tag{12}$$

where t_f is the operational time. Effective FEC gives the effective lifetime depletion with respect to the nominal lifetime defined by Γ. The battery will reach the end of life when $FEC_{eff}(t)=\Gamma$, and the objective of minimizing battery aging is equivalent to minimizing $FEC_{eff}(t)$. Therefore, the cost function to be minimized is defined by:

$$J_E=\frac{1}{2Q_b}\int_0^{t_f}\sigma(t)\left|i_b\right|dt \tag{13}$$

The optimal control problem is subjected to the longitudinal dynamics of the wheel loader. The state equations in the time-domain give a direct description of the dynamic system, however, it is difficult to represent some distance-dependent parameters, e.g., the speed restriction. In this study, the problem of optimizing a speed profile is formulated in the distance domain to make it robust. The dynamics of the wheel loader in (1) is rewritten as:

$$\begin{aligned}\dot{x}_1&=\frac{1}{x_2}\\ \dot{x}_2&=\left[\frac{F_t}{m}-\left(\mu g+\frac{\rho v^2 C_d A_f}{2m}+\frac{F_{th}}{m}\right)\right]\frac{1}{x_2}\end{aligned} \tag{14}$$

where x_1 is the travel time of the wheel loader, x_2 is the wheel loader speed. Thus, the battery current i_b with respect to the state variables can be expressed by:

$$i_b=\frac{P_b}{U_b}=\frac{x_2 u}{U_b}\eta^{-\mathrm{sgn}(u)} \tag{15}$$

where P_{bat} is the battery power. η is the overall efficiency and $\eta=\eta_{dri}\cdot\eta_m$, where η_{dri} is the overall efficiency of drivetrain.

Finally, the optimal control problem takes the following mathematical form, considering the boundary conditions and the physical constraints for states and control inputs:

$$\begin{aligned}
&\min J_E = \frac{1}{2Q_b U_b}\int_0^{s_f} \sigma(s)\left|\frac{x_2 u}{U_{oc}}\eta^{-\mathrm{sgn}(u)}\right|\frac{1}{x_2}ds \\
&s.t.\quad \dot{x}_1 = \frac{1}{x_2},\ \dot{x}_2 = \left[\frac{F_t}{m} - \left(\mu g + \frac{\rho v^2 C_d A_f}{2m} + \frac{F_{th}}{m}\right)\right]\frac{1}{x_2} \\
&x_1(0) = 0,\ x_1(s_f) = t_f,\ x_2(0) = 0,\ x_2(s_f) = 0 \\
&x_2 \in [0, v_{\max}],\ F_t \in [F_{t,\min}, F_{t,\max}]
\end{aligned} \tag{16}$$

where s_f is the distance traveled. t_f and s_f are bounded, meaning that the vehicle stops at time t_f after traveling a given distance, s_f. $v_{\max}$ is the maximum speed of the wheel loader. $F_{t,\min}$ and $F_{t,\max}$ are the braking and driving force limits of the wheel loader.

In order to further reduce the computational cost, the optimal control problem (16) is modified such that a weighted penalty on travel time is included as follows:

$$\begin{aligned}
&\min J_E' = \int_0^{s_f}\left(\frac{1}{2Q_b U_b}\sigma(s)\left|\frac{x_2 u}{U_{oc}}\eta^{-\mathrm{sgn}(u)}\right|\frac{1}{x_2} + \lambda\frac{1}{x_2}\right)ds \\
&s.t.\quad \dot{x}_2 = \left[\frac{F_t}{m} - \left(\mu g + \frac{\rho v^2 C_d A_f}{2m} + \frac{F_{th}}{m}\right)\right]\frac{1}{x_2} \\
&x_2(0) = 0,\ x_2(s_f) = 0,\ x_2 \in [0, v_{\max}],\ F_t \in [F_{t,\min}, F_{t,\max}]
\end{aligned} \tag{17}$$

where λ is a weighting factor to penalize a travel time. Apparently, this formulation can effectively eliminate the state variable x_1 from the optimal control problem. The index of each state is changed from time to distance. The desired travel time can be achieved by adjusting the weighting penalty factor λ.

3.4. Combined DP-BM Method

Dynamic Programming is a technique that is applied to the very wide field of optimal control in multi-stage decision problems. It solves complex problems by breaking them down into simpler subproblems [9]. The motivation of using dynamic programming mainly relies on its ability in finding the global optimal solution. In this study, dynamic programming is implemented with the open-source software DynaProg which was developed by Miretti [10].

As for the optimization problem shown in (17), the desired operating time should be achieved by adjusting the weighting factor λ, then finding the best solution by DP. The operating time is negatively correlated with the weighting factor, where a large λ will force the wheel loader to run at a relatively high average speed. Take $t_f = g(\lambda)$ as the actual final time, and define $f(\lambda) = g(\lambda) - t_d$, where t_d is the target operating time. Therefore, the time constraint is transferred to find the single root of $f(\lambda)$.

To find the root of a function, a hybrid approach called Brent's method is applied in the study, where it does just that by applying a speedy open method wherever possible, but reverting to a reliable bracketing method if necessary, thereby is applied in this study, to achieve time constraint in the cost function. Brent's method combines the bisection method, the secant method and IQI, where the bisection method is used for increasing the possibilities of convergence, the secant

method is used for faster convergence, and the IQI is applied for solving the parabola-typed equation. Combining dynamic programming and Brent's method (DP-BM), the numerical solution of the optimal control problem can be obtained. Details of the algorithm is shown with pseudo-code.

Pseudo-code of the DP-BM algorithm.

Given a stopping tolerance $\delta > 0$

Given points a and b such that $f(a)f(b) < 0$, make sure $|f(a)| \geq |f(b)|$ so that b is regarded as the better approximate solution. $f(a)$ and $f(b)$ are calculated by using dynamic programming.

A third point c is initialized by setting $c = a$.

A flag is initialized by setting flag = Ture.

Repeat until $f(b) = 0$ or $f(\hat{b}) = 0$ or $|b - \hat{b}| < \delta$

  If $a = c$ **then** $\hat{b}$ is determined by linear (secant) interpolation: $\hat{b} = \dfrac{af(b) - bf(a)}{f(b) - f(a)}$.

  Otherwise a, b, and c are distinct, and $\hat{b}$ is determined using inverse quadratic interpolation:

$$\hat{b} = \frac{af(b)f(c)}{(f(a)-f(b))(f(a)-f(c))} + \frac{bf(a)f(c)}{(f(b)-f(a))(f(b)-f(c))} + \frac{cf(a)f(b)}{(f(c)-f(a))(f(c)-f(b))}$$

  If (condition 1: $\hat{b}$ is not between $\dfrac{3a+b}{4}$ and b) or

  (condition 2: flag = True and $|\hat{b} - b| > \dfrac{1}{2}|b - c|$) or

  (condition 3: flag = False $|\hat{b} - b| > \dfrac{1}{2}|c - d|$) or

  (condition 4: flag = True and $|b - c| < \delta$) or

  (condition 5: flag = False and $|c - d| < \delta$) **then**

  $\hat{b} = \dfrac{a+b}{2}$ and set flag = True. (Bisection method).

  Otherwise set flag = False.

  Calculate $f(\hat{b})$ using dynamic programming.

  $d = c$, $c = d$. (d is assigned for the first time here; it won't be used above on the first iteration because the flag is set as True)

  If $f(a)f(\hat{b}) < 0$ **then** $b = \hat{b}$, otherwise $a = \hat{b}$.

  If $|f(a)| < |f(b)|$ **then** swap (a, b).

end repeat

Output the optimal $\lambda = \hat{b}$.

4. SIMULATION STUDIES

4.1. Simulation Parameters

In the simulation, a typical medium wheel loader with an operating weight of 19000 kg is adopted and basic parameters including the drivetrain, the electric motor, the battery, and the working hydraulic system are shown in **Table 1**. The efficiency map of the electric motors is obtained from Advisor 2002. The working hydraulic pump efficiency map comes from Eaton and can be found in previous work [11]. Efficiencies of the other parts of the drivetrain, such as the inverter, gearbox, and final drive, are assumed to be constant.

Table 1: Basic parameters of the wheel loader

Property	Symbol	Value	Unit
Operating weight	m	19000	kg
Radius of tire	r_w	0.789	m
Tire rolling friction coefficient	μ	0.06	-
Drag coefficient	C_d	0.24	-
Front area	A_v	10	m2
Gearbox ratio/Final drive ratio	γ_g / γ_f	11.6/5.1	-
Drivetrain efficiency	η_{dri}	0.95	-
Electric motor maximum power	$P_{m_\max}$	157	kW
Nominal battery capacity	U_{oc}	618	V
Battery pack initial resistance	R_0	1.9	mΩ
Battery pack initial capacity	Q_{init}	456	Ah

The calculation environment is MATLAB R2021b with an Intel Core i7-8750H at 2.2GHz and 16GB of RAM. It should be noted that the accuracy of the global optimal solution of dynamic programming depends on the grid number. In the simulation, the distance interval of two adjacent stages is set as 0.01m. The grid number of the state variable is 4000 and the grid number of the control variable is 1000.

In addition, typical speed trajectories are extracted and summarized from the real trajectories of skilled drivers and used for comparison as baselines. The typical trajectory is constructed by three phases: acceleration, constant velocity, and deceleration. In the acceleration phase, the wheel loader starts at constant torque, then accelerates at rated power. After reaching its maximum speed at low gear, the wheel loader will operate with constant velocity and finally decelerates with a constant negative acceleration.

4.2. Simulation Results under One Typical Distance and Average Speed

The typical and optimized speed trajectories as well as corresponding characteristics with respect to time are shown in **Figure 2**. At the first half of the trajectory, the optimal trajectory uses tapered torque to accelerate the wheel loader. Since the change of torque at low speed has little effect on the motor efficiency (**Figure 2** (d)), higher torque can improve vehicle speed with the same energy consumption. After the wheel loader reaches a high speed, the typical trajectory will keep at this speed and then slow down with energy recovery braking, while the optimal trajectory still operates at high motor efficiency region first, then coasts in the deceleration process and finally adopts the braking action. The braking process of the typical trajectory is much longer than that in the optimized trajectory, which makes the motor operate at a low-efficiency region thus aggravating the battery burden. In addition, coasting is adopted for the optimized trajectory in the deceleration process and this will temporarily reduce battery usage and significantly decrease the accumulated FEC.

As is depicted in **Figure 2** (c) and (f), after driving one complete distance, the optimized trajectory can reduce average power and accumulated FEC with 19.9% and 23.8%, respectively. In addition, the energy consumption of the optimized trajectory also shows a slight decrease, which is shown in **Figure 2** (e). This may benefit from the transfer of the electric motor operating points, where the inefficient energy recovery during long braking of the typical trajectory is replaced by coasting and efficient short-time braking.

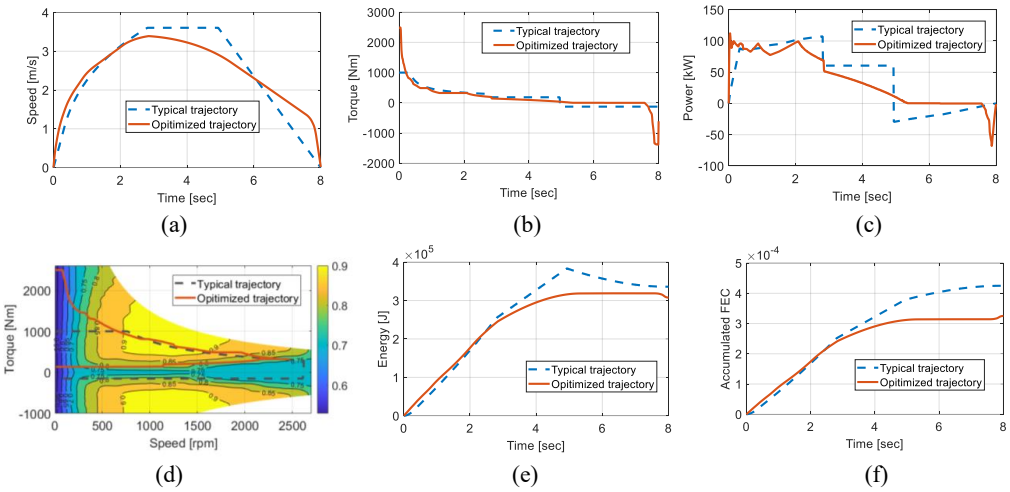


Figure 2: The typical and optimized speed trajectory with respect to time (a) Speed (b) Torque (c) Power (d) Electric motor operating points (e) Energy (f) Battery usage equivalent cycle

Based on the typical and optimized trajectories, a one-day working cycle is designed, which includes double 3-h repeated working cycles, double 3-h charging time (around 0.33C), and 12-h rest time. The accumulated FEC result is depicted in **Figure 3**. It can be found that over the 12-h charging/discharging process, the FEC is greatly down from 1.36 to 1.25 times, which reduces the usage time of the battery with a drop of 8.09%. The decline is smaller than that in one single movement shown in **Figure 2** (f). This is because the FEC is not only influenced by the discharge process but also charge process. And the optimization object in the discharge process does not include the working system. In a one-day cycle, as can be seen in **Figure 4**, the propulsion system takes 51.75% of discharge energy and only 25.85% of total energy. In other words, the accumulated FEC of the battery can be reduced by 8.1% in a full day's cycle just from speed trajectory optimization, without changing the wheel loader hardware configuration.

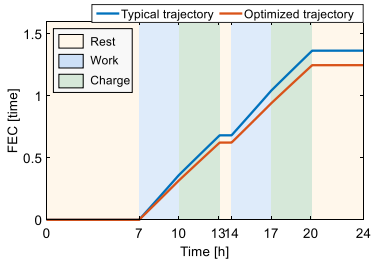


Figure 3: Accumulated FEC in one-day work under the typical and optimized trajectories

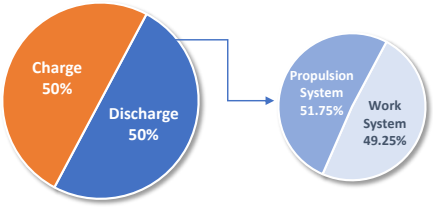


Figure 4: The FEC proportion of propulsion system in one-day cycle

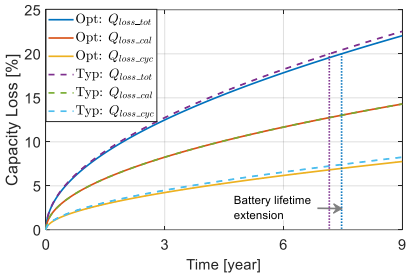


Figure 5: Cycle and calendar aging capacity losses under the typical and optimized trajectories

Based on the typical and optimized trajectories, a 9-year working period with 365 continuous working days in one year is applied in the virtual experiments and simulation results are depicted in **Figure 5**. The total battery capacity loss is decreased by 2.40% after 9 years of continuous operations with the optimized trajectory. The two calendar aging curves are very close since the given temperatures as well as the time periods are the same, and the battery SOCs are also very close for the two simulations. Taking the battery capacity decreases to 80% of the initial capacity as its end of life, the battery pack can be used for only 2529 days under the typical trajectory while 2635 days under the optimized trajectory, meaning that the battery lifetime can be extended by 4.19%.

4.3. Simulation Results under Different Distances and Average Speeds

The speed trajectories and the corresponding variations in battery power for various average speeds are shown in **Figure 6**. As the average speed decreases, the peak power gradually decreases and the power curves are getting smooth. The recovery power also decreases with the average speed decreases. In addition, all the speed trajectories contain 4 parts: speed up, cruising, coasting and deceleration, meaning that such a speed trajectory composition can minimize the battery usage frequency and average discharge current. Furthermore, the power curve shows fluctuation for the trajectories with low average speed. This is due to the coarse grid of state variables in dynamic programming and can be smoother by increasing the grid density.

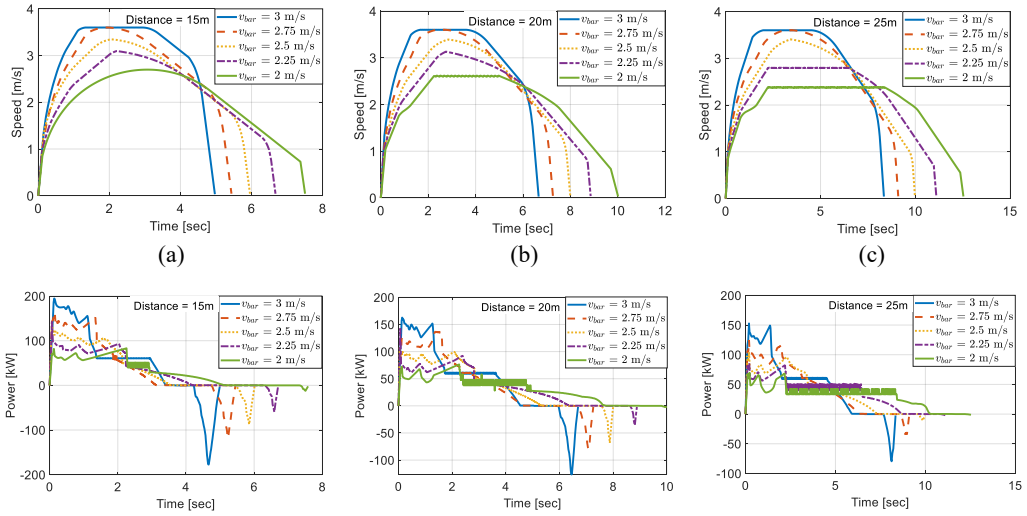
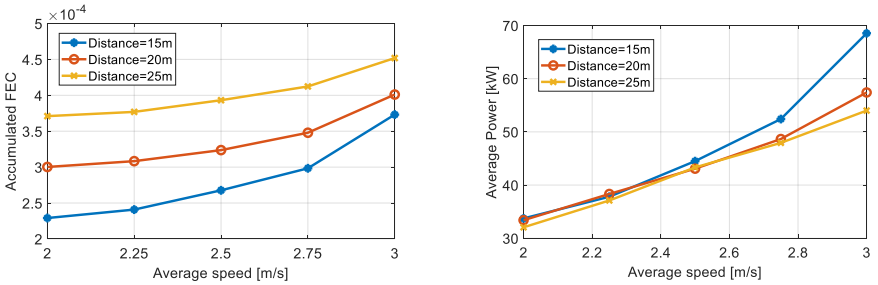


Figure 6: Cycle and calendar aging capacity losses under the typical and optimized trajectories

Numerous simulations utilizing different distances and average speeds are performed and results show that the optimized trajectories reduce accumulated FEC and average power in all tested cases, with average savings of 20.90% and 19.24%. Following prolonged continuous operation, the battery can be utilized for an average of 2632 days when using the optimized trajectory, resulting in a 4.48% extension of the battery's lifetime.

The accumulated FEC and average power of the optimal trajectory at different average speeds are shown in (a) (b)

Figure 7 (a) and (b), respectively. As average speed increases, the wheel loader must speed up and down in a short time, thus both the accumulated FEC and average power show an increasing trend, therefore burdening the battery degradation. On the other hand, under the same speed, the accumulated FEC shows positively correlated with the distance. This is because a long distance consumes more energy, thus increasing the accumulated FEC. The average speed shows small differences for wheel loaders with the same driving distance and average speed, due to the similar work intensity and tasks.



(a) (b)

Figure 7: Accumulated FEC and average power of the optimal trajectory at different average speeds

4.4. Comparison of Computational Time

When dealing with trajectory optimization tasks, the classical optimal control problem usually takes travel time and speed as two state variables and distance sequence as the index. The problem can be solved by conventional dynamic programming while needing an unacceptable amount of time. The DP-BM algorithm is designed to give a rapid numerical solution to the upgraded problem with achieving the time constraints.

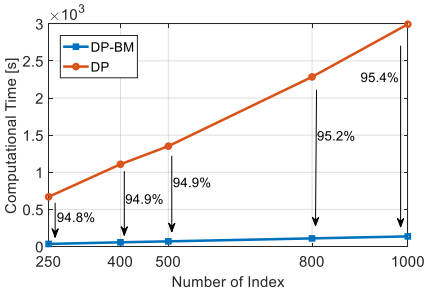


Figure 8: Comparison of computational time with direct dynamic programming

Figure 8 illustrates the computational time comparison between the conventional dynamic programming method and the proposed DP-BM approach for the identical trajectory optimization

problem. The proposed approach demonstrates a significantly reduced computation time compared to conventional dynamic programming, with an average reduction of 95%. It should be noted that the calculation accuracy and calculation time of dynamic programming are contradictory. To further reduce computation time and simultaneously guarantee the accuracy of optimal trajectory, the speed trajectories under different driving distances and average speeds can be calculated offline and save the results in memory. The wheel loader planning system can directly look up the trajectory in the database based on the perceived distance and elapsed time.

5. CONCLUSIONS

This paper presents an optimization of the speed trajectory for an electric wheel loader, with the aim of extending battery lifetime. The optimization problem is formulated based on the modelling of wheel loader's powertrain and battery degradation. To reduce computational costs, a weighted penalty on travel time is included in the optimal control problem. An optimization algorithm, DP-BM, is developed to calculate the global optimal speed trajectory with a reduced computational burden.

A set of typical distances and average speeds is selected to provide a detailed characteristic description of the optimized trajectory. Numerous simulation experiments are conducted to provide an intuitive description of the optimized trajectories. Comparative results demonstrate that with the optimized trajectories, the average battery lifetime can be extended by 4.48% after 9 years of continuous operations. Moreover, the proposed approach significantly reduces computation time by an average of 95% compared to conventional dynamic programming.

In the future, the effectiveness of the DP-BM algorithm in extending battery lifetime will be verified through experiments, and to enhance the algorithm’s robustness. For the existing wheel loader equipped with visual perception, the algorithm will be integrated into the vehicle controller, and display the optimal speed trajectory of the vehicle through the dashboard or on-board screen to guide the driver in manoeuvring the vehicle properly. For future autonomous wheel loaders, it will be possible to directly track the speed trajectory output by this algorithm, achieving the maximum extension of battery lifetime.

REFERENCES

1. Mello EF and Bauer PH. Energy-Optimal Speed Trajectories Between Stops. *IEEE Transactions on Intelligent Transportation Systems* 2020; 21: 4328-4337.
2. Wu XK, He XZ, Yu GZ, et al. Energy-Optimal Speed Control for Electric Vehicles on Signalized Arterials. *IEEE Transactions on Intelligent Transportation Systems* 2015; 16: 2786-2796.
3. Xiao Z, Feng XY, Wang QY, et al. Eco-driving control for hybrid electric trams on a signalised route. *IET Intelligent Transport Systems* 2020; 14: 36-44.
4. Yuan WC and Frey HC. Potential for metro rail energy savings and emissions reduction via eco-driving. *Applied Energy* 2020; 268.
5. Oh K, Yun S, Ko K, et al. Gear ratio and shift schedule optimization of wheel loader transmission for performance and energy efficiency. *Automation in Construction* 2016; 69: 89-101.
6. Naumann M, Schimpe M, Keil P, et al. Analysis and modeling of calendar aging of a commercial LiFePO4/graphite cell. *Journal of Energy Storage* 2018; 17: 153-169.
7. Naumann M, Spingler FB and Jossen A. Analysis and modeling of cycle aging of a commercial LiFePO4/graphite cell. *Journal of Power Sources* 2020; 451.
8. Onori S, Spagnol P, Marano V, et al. A new life estimation method for lithium-ion batteries in plug-in hybrid electric vehicles applications. *International Journal of Power Electronics* 2012; 4: 302-319.
9. Bellman R and Lee E. History and development of dynamic programming. *IEEE Control Systems Magazine* 1984; 4: 24-28.
10. Miretti F, Misul D and Spessa E. DynaProg: Deterministic Dynamic Programming solver for finite horizon multi-stage decision problems. *Softwarex* 2021; 14.

11. Zhang H, Wang F, Xu B, et al. Extending battery lifetime for electric wheel loaders with electric-hydraulic hybrid powertrain. *Energy* 2022; 261.

AUTOMATED SYSTEM SYNTHESIS FOR ELECTRIFIED MOBILE MACHINERY

peer reviewed

Bernhard Sender[1]*, Andreas Opgenoorth[1], Univ.-Prof. Dr.-Ing. Katharina Schmitz[1]

[1]*Institute for Fluid Power Drives and Systems (ifas), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen*

* Corresponding author: Tel.: +49 241 80-47724; E-mail address: bernhard.sender@ifas.rwth-aachen.de

ABSTRACT

The electrification of mobile machines leads to new drive solutions. Not only the electric drive components are newly deployed, but also the mechanical and hydraulic drive components can be utilized in a new or different way. This allows more efficient systems to be built and new advantages to be exploited. Conventional development processes require the developer to select a specific drive concept to be investigated at an early stage. The drive concept is then examined by means of simulations and/or experimental tests. The suitability of the drive system for the application is only evaluated after its detailed investigation and an iterative process. The overall process leads to a high effort and requires a large amount of time. Thus, it is economically not possible to evaluate many different system combinations with conventional methods. This paper proposes a tool to improve the development process. A methodology for an automated drive system evaluation is presented both structurally and in terms of how individual components are modelled. The methodology aims to deduce functioning system topologies based on load cycles. An analysis on the influence of different load cycles is therefore conducted.

Keywords: mobile machinery, system synthesis, electrification, excavator, algorithm

1. INTRODUCTION

In this paper the proposed methodology by Opgenoorth [1, 2] for algorithm-based automated evaluation of mobile machine drivetrain topologies, called TopoSelect is applied to the use case of an excavator. The presented methodology is suitable for the design of mobile electrical machines in general. This paper focuses specifically on excavators as an application example. Since excavators account for a large portion of mobile machinery in use [3–5] and their common drivetrain topologies are generally energy inefficient [3, 6–8], they are responsible for a large portion of CO_2 emissions in the mobile machinery sector [3, 9, 4]. Excavators are therefore an excellent example to show the challenges and opportunities in the electrification of mobile machinery.

2. STATE OF THE ART

2.1. Excavator topologies

A common drivetrain topology for excavators is shown in simplified form in **Figure 1**. It consists of three main parts, the energy source, the energy conversion and the energy output [10, 11]. Since electric powertrains have yet not been adapted on a large scale, the diesel internal combustion engine is still the dominant technology for the energy source for Non-Road Mobile Machinery (NRMM) and excavators in particular [12–14].

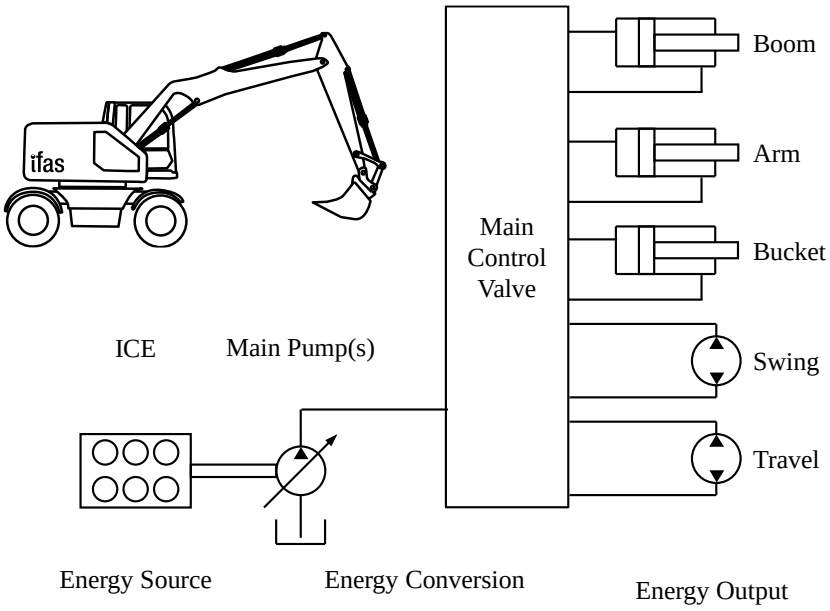


Figure 1: Simplified drivetrain topology of a traditional excavator [10]

The energy conversion consists of the hydraulic pump(s) and the Main Control Valve (MCV). The pump is powered by the engine and supplies hydraulic energy to the system. Traditionally, hydraulic systems of excavators rely on valve-controlled closed center or open center systems. Valve based systems allow the supply of multiple hydraulic drives using just a single pump. The main control valve in such a system, usually featuring multiple valve plates, divides and directs hydraulic flow based on operator joystick commands. It supplies energy to hydraulic drives (energy output), including the hydraulic cylinders of the boom, arm, and bucket, the swing drive for upper carriage rotation, and travel drives. These drives transform the hydraulic energy back into mechanical energy. This conversion powers the linear and rotational movement of the excavator components, enabling its various tasks like digging and grading. [3, 10, 11]

Common systems based on internal combustion engines have proven to be cost effective and reliable in the past [5, 15]. Despite this, electrifying mobile machines offers the advantage of using sustainable energy sources by integrating electrical storage or conversion systems like batteries or fuel cells, eliminating or reducing local emissions. The cost of the components required for this, specifically batteries and fuel cells, are one of the main challenges in the implementation of electric or hybrid drivetrains in mobile machines [2, 12].

As shown by Opgenoorth et al [10] for a 9 t-excavator, electrification can lead to efficiency gains by enabling decentralized direct electric drives and dynamic motor speeds. Machine design plays a crucial role in determining investment costs, particularly in terms of battery storage. In this context, efficiency gains through system design or better components become essential, as they can now lead to cost advantages. [1, 2, 10]

Figure 2 shows the energy flow through the drive train of an 18t wheeled excavator. It becomes apparent that significant losses occur in the whole drivetrain, mainly induced by the low efficiency of the internal combustion engine and systematic hydraulic losses. This means that only 6.4% of the input chemical energy of the diesel fuel are transformed into the mechanical energy used to perform the working task of the excavator. [3] Similar results have been found by Sturm for a comparable

machine [8].

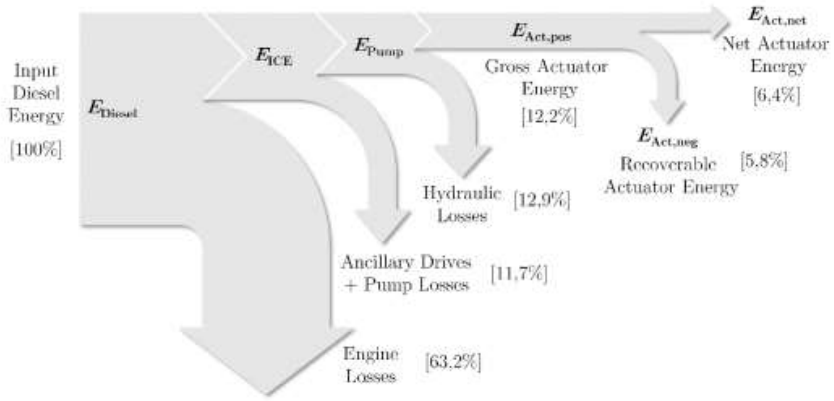


Figure 2: Sankey plot of a 18 t excavator performing a 90° dig and dump cycle [3]

Contemplating on this state of the art many possible topologies for integrating electric drives into mobile machines can be considered. To enable a holistic approach for designing electrified mobile machinery methods for systematic exploration of the possible design space are needed. The presented methodology, called TopoSelect, aims to assist engineers in this task by algorithm based automated evaluation of drivetrain topologies. [1, 2]

3. METHODOLOGY

3.1. System Topology Generation

The TopoSelect methodology separates the drivetrain of a mobile machine into eight subsystem layers. These consist either of mechanical subsystems (mechanical energy conversion), hydraulic subsystems (hydraulic or hydromechanical energy conversion) or electric subsystems (electric or electromechanical energy conversion). **Figure 3** shows the eight layers that are considered in the generation of a drive train topology. The system is calculated backwards from drives to energy source, i.e., the first layer starts with the force and speed, or torque and speed load cycle of the considered mechanical drives. [1, 2]

Between subsystems, the important internal variables are transferred. For the mechanical connections this is speed and torque, for the hydraulic pressure and volume flow and for the electrical current and voltage. Power and efficiencies can then be calculated from these internal variables. [1, 2]

TopoSelect is meant for the development of electric mobile machines, so in all variations the main power supply is always electric. In this application only batteries are considered as the electric power source. Fuel cells or other electrical power sources are not included. This means, the last three subsystems for every topology variation in this application are always electric motor, inverter, and battery, as shown in Figure 3. [1, 2]

A certain path, i.e., a certain drive topology, does not have to include all eight subsystems. This means, that the hydraulic subsystems are generally optional, despite being often used in mobile applications. This enables the option of direct electric drives (optionally paired with a mechanical transmission). For example, for the excavator application, drive trains can be considered in which the linear actuators are hydraulic, but the upper carriage is rotated by an electric swing drive. This is the case with some tracked excavators such as the Kobelco HB365LC/NLC-3 [16] or the Komatsu

HB215LC-3 [17]. The use of electric track drives can also be evaluated with the presented methodology. If a certain path includes a hydraulic system, the linear or rotational mechanical movement is converted into hydraulic variables in the subsystem of the hydraulic converter. These are transferred to the hydraulic system. Constant pressure, open center and load sensing systems as well as displacement control are implemented. [1, 2]

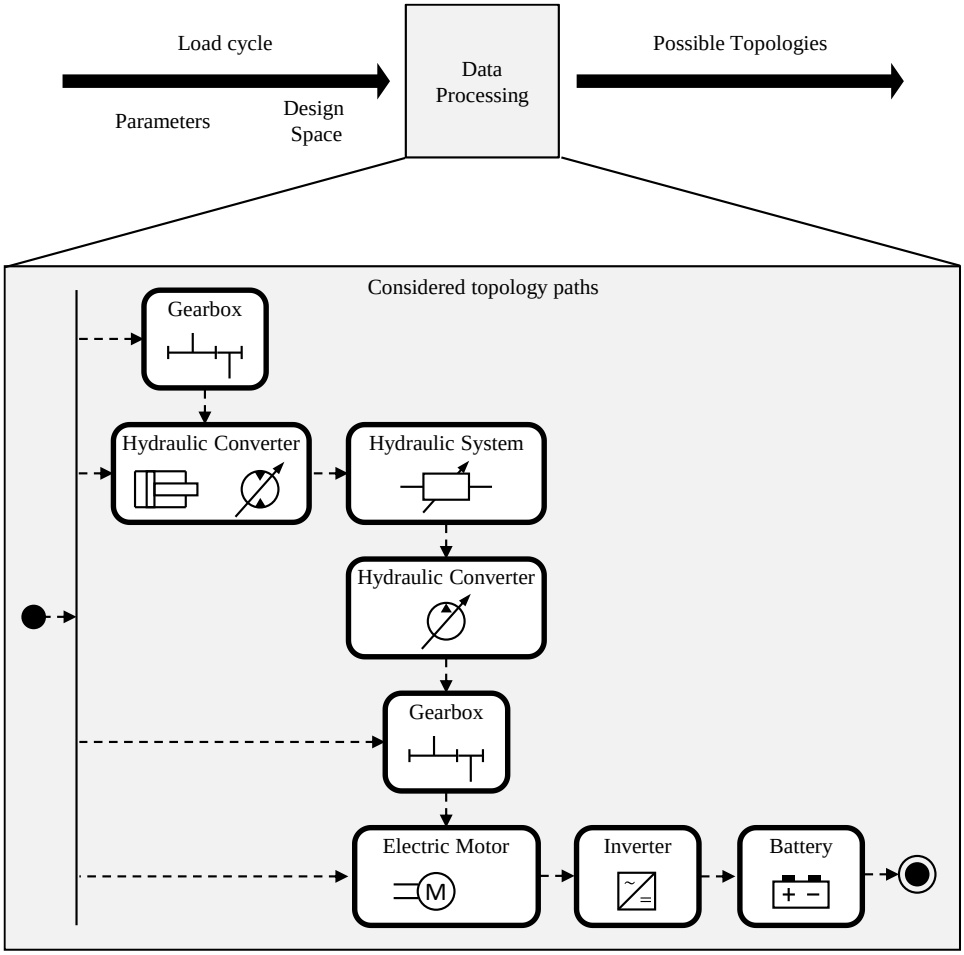


Figure 3: Topology generation based on subsystems

For each subsystem shown in Figure 3 various components of different type and nominal power with different parameters are implemented in an underlying data base. Based on the number of selected components per layer, the number of possible variations can be deduced with equation (1)

$$n_{comb} = \prod_{i}^{8} n_{sub,i} \tag{1}$$

For example, five components selected for every layer result in the calculation of $5^8 = 390625$ variations. Because of this large number of possible paths, the performed calculations within the subsystems must be efficiently implemented regarding computation time, to achieve acceptable performance of the optimization. This is why the calculations within TopoSelect are based on static

analysis of operating points in which design parameters like cylinder diameters or pump displacement volumes are determined, resulting in fast to calculate operations without algebraic loops. This means the methodology focusses onto the energy balance rather than the dynamic properties of the evaluated systems.

For the size of the hydraulic cylinders, for example, a minimum required rod diameter is calculated from the maximum acting force in relation to the buckling safety. The system pressure and the acting forces from the load cycle determine the minimum required piston area. These two variables are used to determine the minimal cylinder size from a range of standardized cylinder sizes.

3.2. Software implementation

The presented methodology is realized through a MATLAB based Software tool. The overall architecture is shown in simplified form in **Figure 4**.

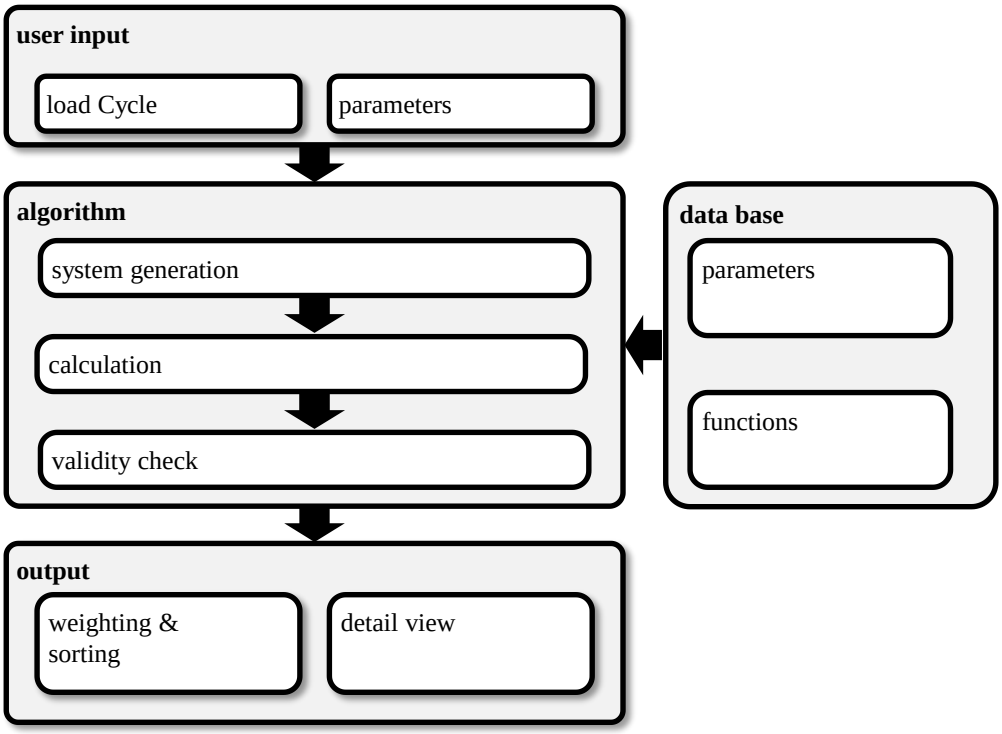


Figure 4: TopoSelect software architecture [2]

The software tool can be broken down into four mayor sections. The user input, the algorithm, the data base, and the output.

User Input

In this interface the user specifies the system requirements with the input of a specific load cycle. Which set of subsystems should be considered in the design space can also be chosen to limit the number of possible variations. The user can also define a set of constraints such as the emission equivalents and cost of electricity, the inverter DC-link voltage, and details regarding the planned workload of the machine.

Algorithm

The software tool generates all possible combinations according to the method explained in chapter 3.1 and calculates the internal variables. Invalid combinations are eliminated. The calculation process places special demands on the software. One of the biggest challenges is the interaction with the database of functions and properties. In particular, the calculation of individual components is associated with significant difficulties. This is because the calculations must be performed quickly, considering the numerous possible combinations, and complicated calculation tasks. To solve this problem, the use of software that can perform parallel calculations with multidimensional data structures is essential. Since most system computations follow a unidirectional flow, a direct implementation is feasible. The unidirectional structure is established by predefined interfaces. However, certain constraints are indispensable in the detailed analysis. For example, the calculation of the inverter losses depends on the battery voltage, which is not known at the time of calculation. In such cases, initial assumptions are required.

Data Base

The performed computations require a data base of components, filled with parameters regarding size, cost, efficiency, and emitted emissions during production. Information about the compatibility with other subsystems is also included. Various drive systems are created in the methodology. These include, for example, gearboxes, hydraulic components, electric machines, power converters, energy storage systems and various types of hydraulic systems. The selection of stored components is illustrated here using the example of hydraulic systems. The hydraulic systems implemented are the constant pressure system, the load sensing system, the open center system with one and two circuits and the displacement control system. The system calculation includes the determination of the pressure losses depending on the required volume flows and also, in the case of the multi-circuit open-center systems, the determination of the complex valve logic. In other categories, however, specific components are stored. In the case of the e-machine, for example, machines with rated outputs of 20 to 400 kW are available. Three variants with different nominal and maximum speeds are included for each of these power classes. A data set with lookup tables for the electrical operating points and the losses as a function of the mechanical operating points is stored for each variant consisting of nominal power and nominal speed.

Output

The output section includes an interface, where the user can make choices regarding the importance of certain properties of the system through associating weights to parameters which are used to rank all the calculated system topology variations. For example, the user can specify if the efficiency or the cost of a specific application is more important. As a result, the software supplies a ranked list of systems to the user based on the made choices. Details, which components form a specific topology, its parameters, and plots of the internal variables, are displayed through a graphical user interface.

4. WEIGHTING METHOD

The following chapter describes in detail how the ranked list of systems as the output is determined. In TopoSelect the Method of Weighted Sums is used. This method is simple and precise, but requires the calculation of a performance value [18]. The following list of criteria can be used to evaluate the performance of the systems:

- Efficiency [%]
- Hourly energy cost [€/h]
- Total cost of ownership [€]

- Investment cost [€]
- Operating cost [€/h]
- Total $C0_2$-emmisions [g]
- Initial $C0_2$-emmisions [g]
- Operating $C0_2$-emmisions [g/h]

Within the algorithm calculation (compare Figure 4) the investment cost and initial C02-emmisions of a specific combination are obtained from the component database using the calculated component sizes. The system efficiency can be calculated from the internal variables (compare section 3.1) of each subsystem layer, the hourly energy cost and operating C02-emmisions are determined with the calculated required power and the user specified cost of energy and C02-emmisions of the used power mix. The operating cost and operating C02-emmisions describe the produced cost and emissions from machine maintenance and can also be calculated for a specific combination of subsystems/components using the underlying component database and the intended daily and annual usage and designed lifetime defined by the user. The total cost of ownership and total C02-emmisions can be calculated from the other criteria. This means that for every valid system topology a value describing its performance regarding the mentioned criteria is obtained.

Through another GUI the user can define the importance of each individual criterion j within the scope of the desired application by inputting a value from 0 to 10 for the criterion weights w_j. By normalising the adjusted weights w_j^{adj} are then calculated (2).

$$w_j^{adj} = w_j \cdot \left(\sum_{j=1}^{n_{crit}=8} w_j \right)^{-1} \tag{2}$$

The Score WS_i of a Combination i is determined according to equation (3) from the adjusted criterion w_j^{adj} and the performance value a_{ij} in terms of the criterion j [18].

$$WS_i = \sum_{j=1}^{n_{crit}=8} w_j^{adj} \cdot a_{ij} \quad \text{with } a_{ij} \in [0,1] \tag{3}$$

With this calculation a value ranging from 0 to 1 is obtained for WS_i,which is used to rank all valid combinations. The required performance value a_{ij} of a combination i with respect to criterion j is calculated by one of two linear cost functions, depending on whether the value of a particular criterion is to be maximized (efficiency) or minimized (costs, emissions). These cost functions can be expressed with equation (4) and represent an interpolation between (0,1) or (1,0) and the maximal and minimal value of the criterion j for all valid topologies.

$$a_{ij} = \begin{cases} \dfrac{c_{ij} - \min(\boldsymbol{c_j})}{\max(\boldsymbol{c_j}) - \min(\boldsymbol{c_j})}, for\ efficiency \\ 1 - \dfrac{c_{ij} - \min(\boldsymbol{c_j})}{\max(\boldsymbol{c_j}) - \min(\boldsymbol{c_j})}, else \end{cases} \tag{4}$$

In this notation $\boldsymbol{c_j} = \{c_{1j}, c_{2j}, c_{1j}, \dots c_{nj}\}$ is the set of values of all n valid combinations regarding the criterion j and c_{ij} the specific value for the topology i. For example, considering the investment costs criterion, this calculation results in a performance value of 1 for the topology with the lowest investment costs and a performance value of 0 for the topology with the highest investment costs. All other performance values are relative to these and therefore lie between 0 and 1.

By sorting the list of topologies according to their calculated score a sorted and ranked list of drivetrain topologies suited to the used load cycle and operating conditions specified is obtained.

5. APPLICATION: EXCAVATOR

To showcase the capability of the proposed method, the data of a dig and dump cycle of an 18 t wheeled excavator is used as input data. It contains the load forces and velocities of the three hydraulic cylinders powering the motion of the boom, arm and bucket and the torque and the rotation speed of the upper carriage. These load cycles are shown in state space representation in Figure 5

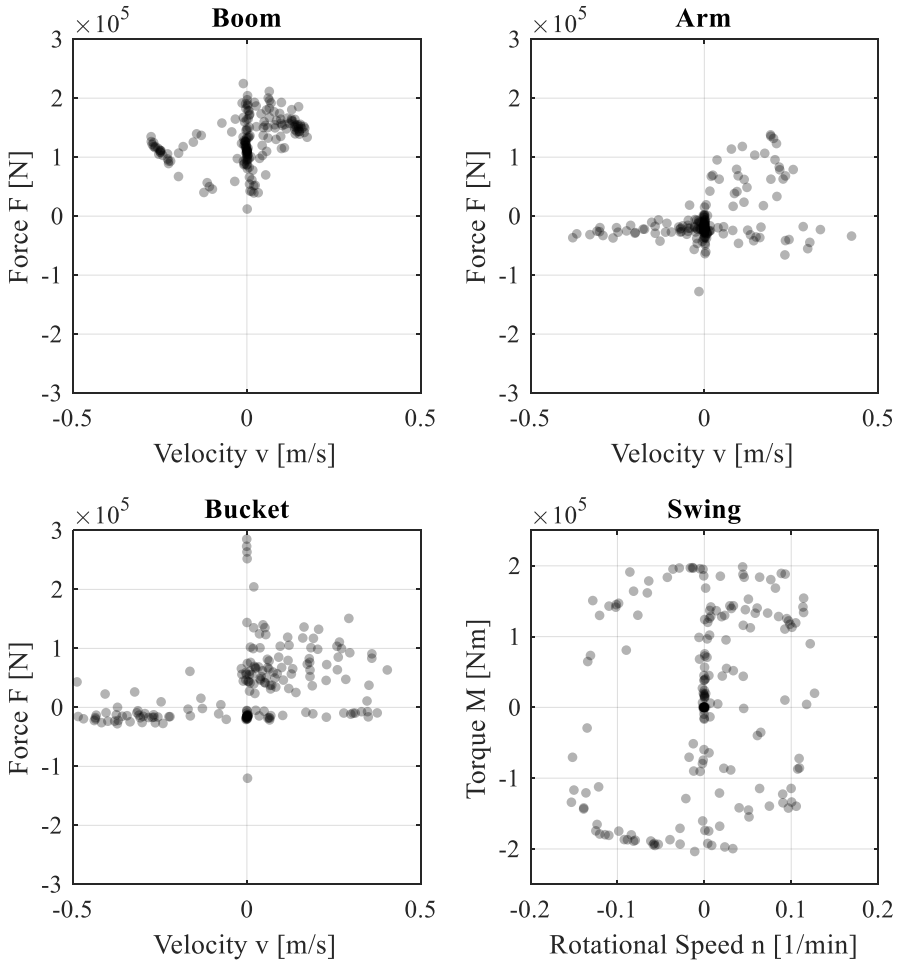


Figure 5: Load Cycle Dig and Dump

Figure 6 shows the Graphical User Interface (GUI) of the data input interface of the TopoSelect tool. The shown load cycle has been selected, displayed as force, velocity, torque, and rotation speed over time in the right half of the window. The left half of the GUI enables the configuration of parameters and definition of the design space. As values for the C02-emission parameter describing the used

power mix 375 g/kWh was chosen. In addition to that, an electricity cost of 0.3 €/kWh, a daily usage of 5 hours, a designed lifetime of 10,000 hours, an annual usage of 150 days was set. The DC-voltage between the inverter and battery is specified with 750 V. The “toggle components”-button in the lower left section of the window opens another GUI in which the limitations of the design space can be specified by including or excluding specific subsystems/components from the calculation. For this application five different output mechanical transmissions and the direct electrical and hydraulic drives were considered, also two options for the hydraulic converters, three different hydraulic systems (open center, load-sensing and constant-pressure), six different hydraulic pumps, six different input mechanical transmissions, eight different electric machines, seven different inverters and three different batteries were selected.

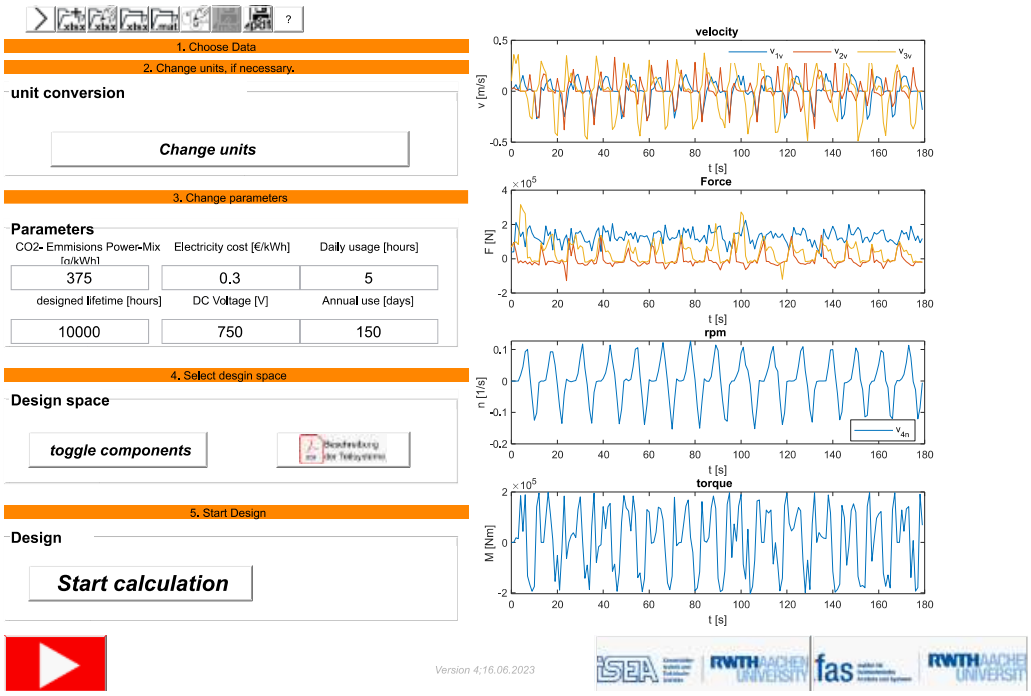


Figure 6: Graphical User Interface

According to (1) this means 254016 possible combinations. From this TopoSelect determined and calculated 405 valid combinations with the algorithm presented in chapter 3. In this example application two of eight possible parameters were used to evaluate the combinations, efficiency (2/10) and total cost of ownership (6/10).

The list is sorted according to the specified parameters. The chosen parameters in this application, efficiency and total cost of ownership are displayed for all valid combinations in a bar plot shown in **Figure 7**. In this representation combination 1 has the best and combination 405 the lowest score. The best combination has the highest efficiency paired with the lowest total cost of ownership. The suggested topology of this system consists of a mechanical transmission paired with a secondary controlled hydraulic motor as a swing drive and hydraulic cylinders to drive the boom, arm, and bucket. These are connected to the hydraulic pump through an Open-Center hydraulic system. A hydraulic pump with a constant displacement volume is selected. The variation of the supplied hydraulic flow is achieved by a variable drive speed. The pump is connected via an electric motor to

the inverter and battery.

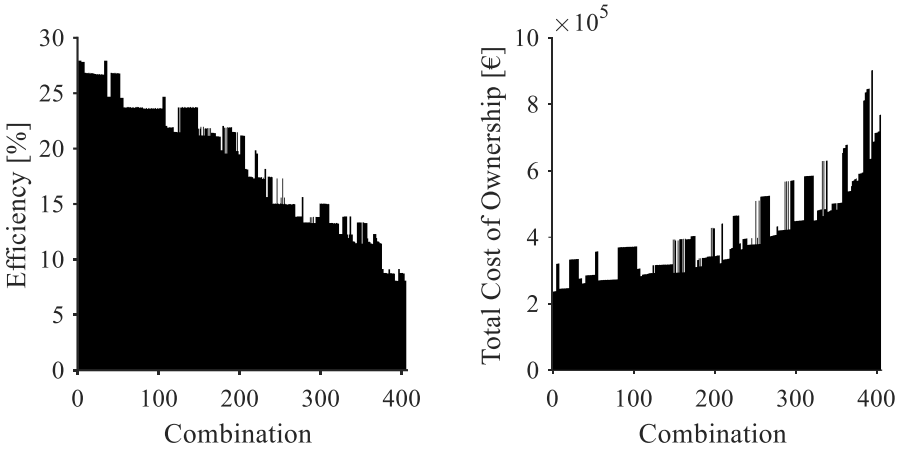


Figure 7:Sorted combinations

5.1. Interpretation

The results obtained must be interpreted considering the underlying database. The more accurate the parameters describing the components, the more accurate the calculated results will be. Since some of the relevant parameters, e.g., the exact costs, detailed efficiency maps, or emissions generated in the production of a particular component, are not easily specified or available in this context, most of the underlying database in this application example does not have the desired accuracy. A more accurate data base would therefore facilitate the applicability of the calculated results. This means that the presented results should only be interpreted as a proof of concept of the capability of the used method rather than a universally applicable recommendation for the drivetrain design of electrified excavators. In addition to that, it must be noted that excavators are often used for a wide range of tasks [4]. That is especially true for smaller machines. This means a proper design approach should include the use of multiple load cycles.

6. SCOPE AND INTENDED USE

The added value contributed by the presented methodology is the systematic and extensive implementation of standard dimensioning calculations, combined with a weighting and sorting algorithm. This enables the systematic investigation of all possible solution combinations based on a component database, certain boundary conditions and a load cycle. In contrast to classical methods, this makes it possible to consider a massive number of variants at an early stage in the design process of an electrically driven mobile machine, where the requirements but no functional structure have yet been defined. The methodology is not intended to optimize a fixed topology of a drive train, but to identify a group of promising topologies. This is particularly important in the development of electrically powered mobile machinery, as completely new drive topologies are now feasible. The method presented is therefore intended to contribute to a holistic view of the entire solution space. Once a selection of topologies has been identified, the design process must be continued with detailed design and other methods, e.g. dynamic simulations.

7. CONCLUSION AND OUTLOOK

As stated above the successful application depends heavily on the availability of detailed and precise component data. However, such databases are not publicly accessible or are unsuitable for such a connection with the current state of development. Therefore, the major task in applying the methodology to a real machine design process is to gather and implement an accurate component data base. In principle, when connected to a comprehensive database system of component manufacturers, the presented methodology is a good approach for the holistic discussion of possible powertrain topologies of electrified mobile machines.

To further complement the design process future work could focus on the implementation of the automated evaluation of multiple load cycles or load collectives. Efforts to improve computation time might enable the inclusion of more sophisticated calculations within the subsystems itself. Both approaches would facilitate the practicability of the presented design methodology.

8. ACKNOWLEDGEMENTS

The project was accompanied and supported by MOBIMA e.V. The authors would like to thank all participants for the financial support and the professional assistance.

NOMENCLATURE

a_{ij}	Performance value	[-]
$\boldsymbol{c_J}$	Set of criterion values	[-]
c_{ij}	Criterion value	[varying]
n_{comb}	Number of possible combinations	[-]
n_{crit}	Number of criteria	[-]
$n_{sub,i}$	Number of selected components in layer i	[-]
w_j	Criterion weight	[-]
w_j^{adj}	Adjusted criterion weight	[-]
WS_i	Combination Score	[-]
DC	Direct current	
GUI	Graphical User Interface	
NRMM	Non Road Mobile Machinery	
MCV	Main Control Valve	

REFERENCES

[1] Opgenoorth A, Loewenherz R, Doncker RW de et al. (2023) Untersuchung einer automatisierten Systemsynthese elektrifizierter mobiler Arbeitsmaschinen. In: Geimer M (ed) Hybride und energieeffiziente Antriebe für mobile Arbeitsmaschinen Karlsruhe: 9. Fachtagung, 28. Februar 2023. KIT Scientific Publishing, Karlsruhe, pp 49–71

[2] Opgenoorth A (2023) ALGORITHM BASED SYSTEM SYNTHESIS OF ELECTRIFIED MOBILE MACHINES. In: Tatiana Minav (ed) The 18th Scandinavian International Conference on Fluid Power, SICFP'23, Tampere, Finland

[3] Vukovic M, Leifeld R, Murrenhoff H (2017) Reducing Fuel Consumption in Hydraulic Excavators—A Comprehensive Analysis. Energies 10:687. https://doi.org/10.3390/EN10050687

[4] Fecke M (2018) Bewertung der Energieeffizienz von Baumaschinen mithilfe einer praxisnahen Lastzyklusentwicklung für einen In-Situ-Test. Bergische Universität Wuppertal

[5] Leifeld R (2022) Hybridmodul für elektrohydraulische Antriebssysteme von Baggern. Shaker Verlag, Düren

[6] Casoli P, Scolari F, Minav T et al. (2020) Comparative Energy Analysis of a Load Sensing System and a Zonal Hydraulics for a 9-Tonne Excavator. Actuators 9:39. https://doi.org/10.3390/act9020039

[7] Salomaa V, Minav T, Mattila J et al. (2018) Efficiency Study of an Electro-Hydraulic Excavator. In: Murrenhoff H (ed) 11th International Fluid Power Conference: Fluid Power Networks. RWTH Aachen University, Aachen, pp 372–385

[8] Sturm C (2015) Bewertung der Energieeffizienz von Antriebssystemen mobiler Arbeitsmaschinen am Beispiel Bagger. KIT Scientific Publishing, Karlsruhe

[9] Fecke M, Helmus M (2015) Standardisierung definierter Lastzyklen und Messmethoden zur Energieverbrauchsermittlung von Baumaschinen. In: Forschungsvereinigung Bau- und Baustoffmaschinen e.V. (ed) Proceedings of the 6th Fachtgaung Baumaschinentechnik: Maschinen, Prozesse, Vernetzung, pp 113–122

[10] Opgenoorth A, Quabeck S, De Donker RW et al. (2020) Challenges and possibilities of the integration of electric drives in mobile machinery. In: Technische Universität Dresden (ed) 12th International Fluid Power Conference, vol 1. Dresdner Verein zur Förderung der Fluidtechnik e. V. Dresden, Dreseden, pp 471–480

[11] Murrenhoff H, Eckstein L (2014) Fluidtechnik für mobile Anwendungen: Umdruck zur Vorlesung, 6., überarb. Fassung. Reihe Fluidtechnik, vol 5. Shaker, Aachen

[12] Lajunen A, Sainio P, Laurila L et al. (2018) Overview of Powertrain Electrification and Future Scenarios for Non-Road Mobile Machinery. Energies 11:1184. https://doi.org/10.3390/EN11051184

[13] Lindgren M, Hansson P-A (2004) Effects of Transient Conditions on Exhaust Emissions from two Non-road Diesel Engines. Biosystems Engineering 87:57–66. https://doi.org/10.1016/j.biosystemseng.2003.10.001

[14] Radermacher T, Merx M, Sitte A et al. (2021) Potentialstudie Energie-/ Kosteneinsparung in der Fluidtechnik: Abschlussbericht. Climate Change

[15] Vukovic M (2017) Hydraulic hybrid systems for excavators. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen; Shaker Verlag

[16] Kobelco (2021) Fully equipped Hybrid SK210HLC-10 with 2 piece boom in the Netherlands. https://www.kobelco-europe.com/news/fully-equipped-hybrid-sk210hlc-10-with-2-piece-boom-in-the-netherlands/

[17] Komatsu (2016) Komatsu Europe International N.V. Presents New HB365LC-3 Hybrid Excavator. https://www.komatsu.eu/Assets/GetPressReleaseByProductName.aspx?id=HB365LC/NLC-3%20Hybrid&langID=de

[18] Bender B, Gericke K (2021) Pahl/Beitz Konstruktionslehre. Springer Berlin Heidelberg, Berlin, Heidelberg

Chapter 8

Pumps

PREDICTIVE MAINTENANCE FOR AXIAL PISTON PUMPS: A NOVEL METHOD FOR REAL-TIME HEALTH MONITORING AND REMAINING USEFUL LIFE ESTIMATION

peer reviewed

Dr. Anik Kumar Samanta*, Dr. Shrinivas Kulkarni, Nitin Hande

[1]*Danfoss Technologies Private Limited, Pune, India.*

* Corresponding author: Tel.: +91 8617872775; E-mail address: anik.samanta@ieee.org

ABSTRACT

In this paper, a novel method to estimate and predict the condition of an open circuit piston pump is presented. We introduce the concept of the 'Pump health index' which can assess the health of the pump in real time and use it to estimate the remaining useful life of the pump. The solution is agnostic to pump size, make, and application duty cycle. The solution has been tested with different levels of degradation that were simulated on the physical pump. The algorithms were implemented on multiple embedded platforms to illustrate the agnostic nature of the developed technology.

Keywords: Open circuit axial piston pump, pump health index, edge computing, predictive analytics, machine learning, pump health monitoring, remaining useful life.

1. INTRODUCTION

Piston pumps – the heart of any hydraulic application is critical to various industrial and mobile applications. Condition monitoring of piston pumps has been a key area of industrial research for many years. A detailed customer profile and value chain analysis is shown in Figure 1, which indicates that health monitoring solution for piston pumps offers incentive for each stakeholder in the value chain.

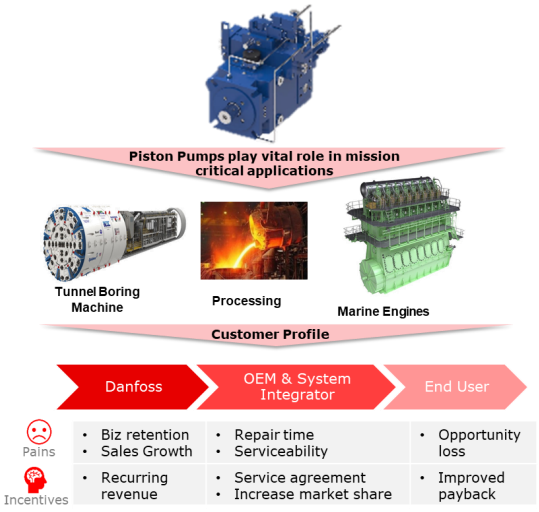


Figure 1: Market potential and customer profile for solution (*All the images are copyrighted to their respective owners)

1.1. Objectives and scope

Piston pumps find applications in mission-critical areas like marine engines, tunnel boring machines, steel manufacturing, and discrete manufacturing to name a few. Piston pumps operate under adverse environments and can operate without any discernible change in performance under incipient fault conditions. However, with continual operation, the fault condition deteriorates and can cause major disruption and downtime [1]. Any downtime on these machines incurs huge losses to the customers. The customer challenges related to these machines could be:

1. Revenue loss due to pump failures: As the pump is the heart of any hydraulic system, failure related to the pump causes non-operation of the machine and hence loss of revenue.
2. Operational cost increase: As customers do not have a mechanism to access pump health status, they may continue using the same pump irrespective of its degradation status. Due to the usage of a degraded pump, the amount of energy consumed is higher than a new pump, for the same amount of useful work. Hence operational cost increases due to the usage of degraded pump.
3. Unplanned system stoppage: Continual operation of degraded pumps can result in pump failures, causing severe downtime for the entire machine. The unplanned system stoppage is a huge challenge for customers.

To overcome these challenges, some of the existing strategies are 1. backup & redundancies, 2. periodic maintenance, and 3. reactive maintenance. All these strategies would be costly and would not solve the problem completely. Hence there is an unmet need to provide real-time health indications for a critical component like a pump.

The objective of the paper is to develop a technology that analyses fundamental signatures from hydraulic pumps and provide insights for continuous monitoring of pump performance as shown in **Figure 2**. For this solution, key signatures from a pump are captured through sensors and their real-time analysis is carried out on a controller. The processed output of the algorithms is in the form of Pump Health Index (PHI), shown as the hypothetical curve with respect to operating hours. The PHI is estimated till current time and using the previously computed values, a forecast is performed.

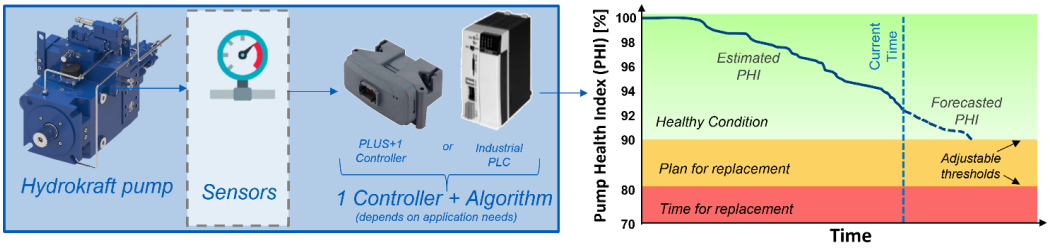


Figure 2: Pump health monitoring solution with the indicated health index

1.2. Literature review and state of art

Most methods estimate volumetric efficiency using outlet flow. However, measuring the main flow affects the output of the pump. Multiple papers have proposed methods that use pressure [2] and vibrations [3] signals to derive fault features. Some of the previous works from Danfoss also show promising results [4]–[6], where grey box modeling is used to predict system performance in the form of motor speed. This work evaluates the pump performance as a standalone product when rotary or linear velocity is not available. Recent advances in machine learning (ML) have been widely used for detecting the defects of piston pumps [7]. As the evolution of ML techniques is rooted in domains other than industrial applications (like computer science), application-specific tuning and capturing

application-specific domain knowledge is not warranted. An attempt to integrate the knowledge into neural networks has been made by Kulkarni and Guha [8]. Most of the ML algorithms require specialized embedded hardware for their implementation. Machine learning-based methods assume the generalizability of the trained models as fault class data are not always present for the pump under test. Apart from the development of these advanced techniques, their deployment strategies concerning computational architectures are also critical. A recent review of condition monitoring of axial piston pumps for mobile applications can be found in [9].

Many researchers have evaluated computational architectures from various perspectives [10], [11]. Particularly for PHM (Pump Health Monitoring) solutions, [12] indicated critical points to be considered like, data processing and storage capability [13], asset overview, domain knowledge, and robustness for estimating the 'Remaining Useful Life (RUL)' [14]. This work also proposes computational architectures for the effective deployment of condition monitoring solutions. The work focuses on the aspects of effective implementation, including the time of response from data acquisition to decision-making and computational resources for data processing in real-time.

Alternatively, residue-based methods can be used to fix a baseline model under the healthy state of the pump, but it requires online training of the models on resource-constrained embedded systems. The condition monitoring of axial piston pumps has matured in the last two decades from a signal processing problem to a data science problem. In [15] spectral analysis and wavelet transformation of outlet pressure were used for pump signature analysis. RUL prediction using different leakage models was proposed in [16] using case flow measurement and Weiner filter. Low and high Reynolds flow losses are estimated as states of the pump model using an extended Kalman filter (EKF) with pressure as a measurement for RUL prediction [17]. The use of particle filters for estimating the RUL was proposed in [18]. A non-linear unknown input observer using swash-plate angle and outlet pressure input was proposed in [19]. However, the lack of correction due to unavailable future data leads to linear regression-type estimation with state estimators. The volumetric efficiency can be estimated using such a state estimator. However, the prediction method will remain unchanged.

The relation between oil contamination and the RUL of the pump was established in [20]. Fault isolation for different parts of the pump like cylinder, valve plate, slipper, sliding boot, and spring wear in the lab environment using vibration signature was demonstrated in [21]. The classification of the faults was carried out with a convolutional neural network. However, the absence of adequate and publicly available data on the fault classes makes it challenging for the implementation of such machine learning-based methods in industrial cases for generalized fault classification. Degradation characteristics of port plate pairs extracted from flow by monitoring the volumetric efficiency were exhibited in [22]. The use of compressed sensing for fault detection was proposed in [23]. However, the spectral estimation of reconstructed, compressively sensed signal is lossy and may lead to loss of fault information. The use of Eigenvectors as indicators of pump degradations using a pre-filtered vibration signal was demonstrated in [24].

Case-flow of an axial open circuit piston pump indicates tribological interface component of wear. However, methods based solely on the case flow can't determine the condition of the pump. Case flow is dependent on the pressure, swash-angle, speed, and fluid viscosity of the pump. In this paper, we propose a method that considers these factors to determine the condition of the pump. The solution is agnostic to pump size, make, and application duty cycle, which has been tested with different levels of degradation that were simulated on the physical pump. The prediction has been tested using different degradation patterns generated mathematically and from actual pump data. The overall estimation error is less than ±2%.

The estimated PHI values are used for predicting the future health of the pump. Estimating the pump health as one unique number has provided the possibility to use a machine learning model, which can

be deployed through resource-constrained edge hardware. This unique number (PHI) is fed to the prediction model, which gets updated for every new value. This provides the health predictions which are updated as per the previous degradation pattern.

This continuous learning methodology using machine learning models provides not only real-field pump degradation but also its variation and prediction to reach certain thresholds set by customers as per application needs. To deploy the methodology for customer application, non-linearity should be handled carefully. The non-linear behaviour of PHI is taken care through continuous learning methodology. An advanced outlier removal method provides the required robustness for actual field deployment.

The estimation and forecast algorithms were implemented on various controllers including Danfoss's Plus+1® MC024™ [25]. The validation has been carried out in two phases. In the first phase, a piston pump has been artificially degraded in the lab by artificially wearing the valve plate. In the second phase, different tests have been carried out using a wide range of pumps (low to high displacement) to validate the algorithms. The overall system is ready to be deployed as a prototype solution at customer sites.

2. METHODOLOGY

One of the challenges faced by the operator/plant manager is the evaluation of the present health status of the pumps and how the performance of the pump might degrade in the future. The proposed method works by calculating the PHI in real-time, storing historical data of the PHI, and predicting the values of PHI up to a certain interval in the future. The PHI is related to the volumetric efficiency of the pump and can be estimated using the leakage flow. However, it is well known that the leakage flow is also correlated with the outlet pressure, fluid viscosity, and speed of the pump. To develop an index that is only affected by the degradation of the pump we measure the correlated sensors for normalization. The overall method is shown in Figure 3.

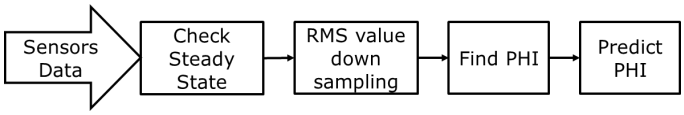


Figure 3: The algorithmic block diagram

Assuming the theoretical flow to be Q_{Th}. The main flow is a fraction of Q_{Th} given by $Q_m = \eta Q_{Th}$. Therefore, the rated case flow or the leakage flow is given by $Q_c^{rated} = (1-\eta)Q_{Th}$. The case flow is directly proportional to the outlet pressure [26]. Therefore, $Q_c \propto P$. Therefore,

$$\frac{Q_c}{Q_c^{rated}} = \frac{P}{P^{rated}} \tag{1}$$

$Q_c = Q_c^{rated} P / P^{rated}$. Using the relation of Q_c^{rated}, we get the leakage flow as

$$Q_c = \frac{P}{P^{rated}}(1-\eta)Q_{Th} \tag{2}$$

Now $Q_{Th} = \omega d$, in which, ω is the rotational speed per minute of the pump and d is the displacement in cubic centimeter of the pump. Therefore, Q_c is given as follows:

$$Q_c = \frac{P\omega d}{P^{rated}}(1-\eta). \tag{3}$$

The total leakage flow is solely not due to degradation. There is control flow which is represented as

a fraction of the theoretical flow given by βQ_{Th} [26]. Hence,

$$Q_c^{act} = \frac{P}{P^{rated}}[\omega d(1-\eta) - \beta Q_{Th}] \tag{4}$$

The quantity β can be obtained empirically from initial data or end-of-line testing of the pump. Once we obtain the actual case flow due to leakage given by Q_c^{act}, we find the relation between the outlet pressure and Q_c^{act} for a window of data. The PHI is a function of the slope of the fitted curve for that window of the data.

3. EXPERIMENTAL SETUP

With the introduction of the novel concept of 'pump health index', it is imperative to verify and validate the method with situations replicating real-field duty cycles and data. For this purpose, an experimental test bed in our facility has been utilized. The experimental test bench schematic with Danfoss Plus+1 controller and display is shown in Figure 4. In this setup, the hydraulic pump is driven by an electric motor. Different loading conditions on the pump were experimented using relief valves.

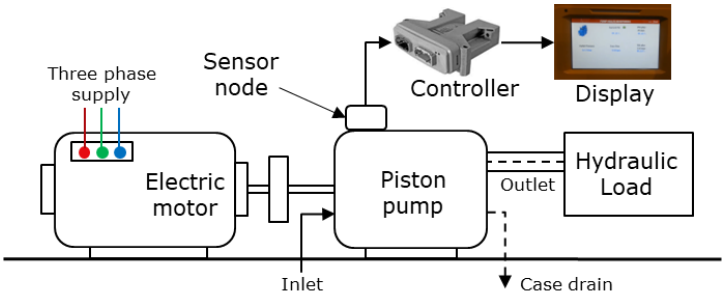


Figure 4: Schematic representation of the experimental test bench

The sensor data is communicated through CAN protocol to the edge hardware from the test-bed programmable logic controller (PLC), where the algorithm is deployed. The output of the algorithm was displayed on the display. A photograph of the pump, driven by the electric motor is shown in Figure 5. Real-field duty cycle is replicated in this setup. Five sensor measurements were captured, i.e., pump discharge pressure, case flow, swash angle, fluid temperature, and speed as shown in Figure 6.



Figure 5: One of the pumps under test

Pump discharge pressure shows the loading of the pump with respect to time. Real-field dynamic duty cycle was achieved, which shows the variation from 20 bar to 210 bar. Case flow values captured

in real-time show the variation as indicated in Figure 6. The swash angle sensor shows a variation of 0 to 18.5 degrees, which is the maximum swash angle for this pump. These variations are as per pump controls. Fluid temperature variation is not dynamic in nature, as the same is expected in field operations. Hence experiments have been conducted at 50 °C, 75 °C, and 93 °C fluid temperature.

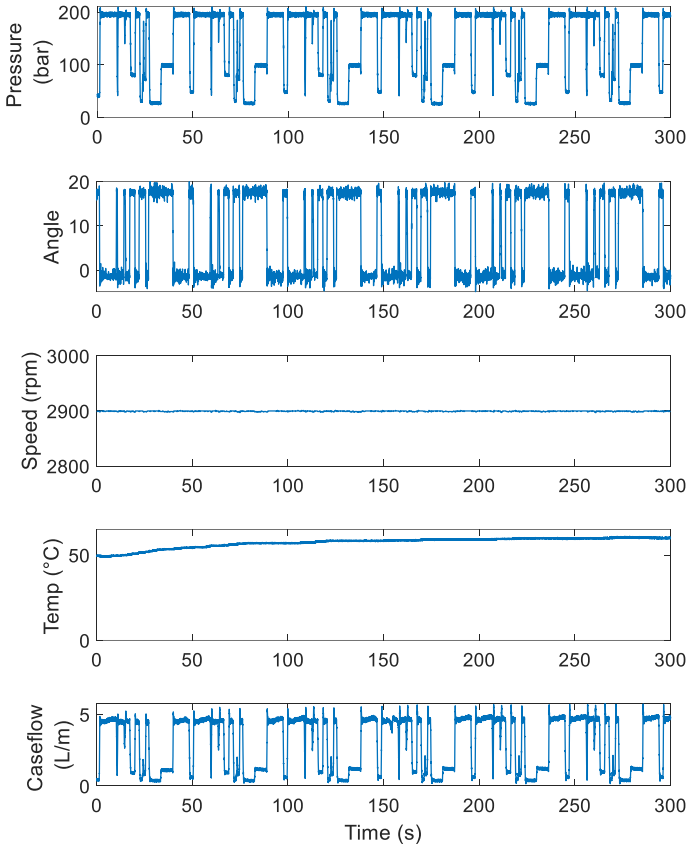


Figure 6: Visualization of different recorded signals of the pump. From top: Pressure, Swash-angle, pump-speed, temperature, and case-flow.

Many industrial applications run the pump at constant speed, with the possibility of changing speed as needed. Hence, PHI algorithms have been tested at constant speed and the same experiment has been repeated with variation in speed at 1000 rpm, 1500 rpm, 2200 rpm, and 2900 rpm. Generally degraded pump has lower efficiency. To collect the data of degraded pumps, tribological parts were degraded and then data was collected. The data is used to estimate and predict the performance of the pump.

4. RESULTS AND DISCUSSIONS

With the experimental setup described in Section 3, detailed test scenarios have been formulated to capture and test the robustness of the solution. The purpose of this testing is to capture the performance of the algorithm with real-field scenarios. The different test scenarios with a variety of pumps are described below. The actual and estimated PHI has been compared.

a. Piston pump sizes (cc): 250 cc, 90 cc, 66 cc, and 28 cc.
b. Speed (rpm): Speed variations have been tested from 500 rpm to 2900 rpm.
c. Temperature (°C): Temperature variations have been captured from 50°C to 93°C.

PHI (Actual): This is the volumetric efficiency of the pump at the rated condition at 50°C. To capture this index, volumetric efficiency is calculated, using the main flow measurement of the pump.

PHI (Estimated): This is the estimated PHI as an output from the developed algorithm, without using the main flow. Calculation of estimated PHI is taking place using the actual sensor data with dynamic variations with time, like the real field scenario.

4.1. End-of-line Testing with Production Pumps

To validate the proposed method statistically, the algorithm was tested on multiple Hydrocraft pumps with end-of-line test data. A total of 13 PVX-66 pumps were tested. The maximum error observed was 1.808 %. Similar tests were carried out with 85 number of 90 cc PVX-90 pumps and 15 number of 250 cc PVX-250 Pumps. The actual versus the estimated PHI are shown in Figure 7. The mean percentage error with PVX-250 was found to be 1.54 %. The higher error was expected for end-of-line test data as this data serves as the baseline for the control flow as discussed in Section 2.

Table 1: Estimation performance of 66 cc Hydrocraft pump

Pump #	Speed (RPM)	Temperature (°C)	Estimation Error (%)
1	1500	44	0.75
2	1500	44	0.87
3	1500	40	0.37
4	1500	40	0.72
5	1500	44	0.14
6	1500	40	0.13
7	1500	40	1.57
8	1500	40	1.16
9	1500	44	0.16
10	1500	40	1.79
11	1500	44	0.17
12	1500	44	**1.80**
13	1500	44	0.87

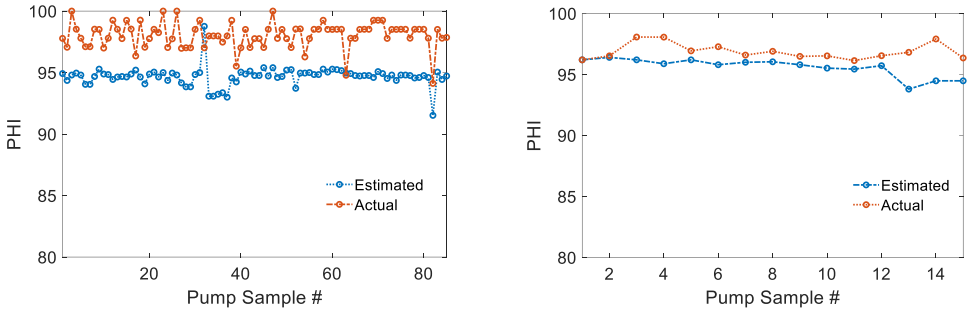


Figure 7: End-of-line validation of the PHI with Hydrocraft pumps. Left: PVX-90, Right: PVX-250

4.2. Validation with different speeds and temperature in laboratory condition:

A total of 12 different data sets have been captured with different testing scenarios of speed and temperature. The error between the estimated and actual PHI is less than 2%, as shown in Table 2 to demonstrate the robustness of the developed algorithms.

Table 2: Estimation performance of 28 cc pump

Dataset	Speed (RPM)	Temperature (°C)	Estimation Error (%)
1	1500	50	1.13
2	2200	50	0.33
3	2900	50	0.67
4	1500	50	1.33
5	2200	50	0.53
6	2900	50	0.87
7	1500	93	**1.53**
8	2900	93	0.27
9	1500	75	1.33
10	2900	75	0.27
11	1500	60	1.33
12	2200	60	0.27

4.3. Validation with Degraded Pump under Laboratory Conditions:

The robustness of the solution also needs to be validated for new pumps, as well as for the old pumps. Pump degradation is emulated in the lab environment by deliberately removing the material from the valve plate. The artificial wearing of the valve plate is carried out using diamond paste & sandpaper. The valve plate has been weighed and a pump performance test has been carried out after each iteration. Validation of the algorithm is carried out at different levels of pump degradation (D1 – D5) with some instances where the pump was run under corner horsepower (CHP) ratings as shown in

Figure 8. The validation results with various levels of pump degradation show an estimation error of less than 2%. This confirms the solution validity for new as well as degraded pumps.

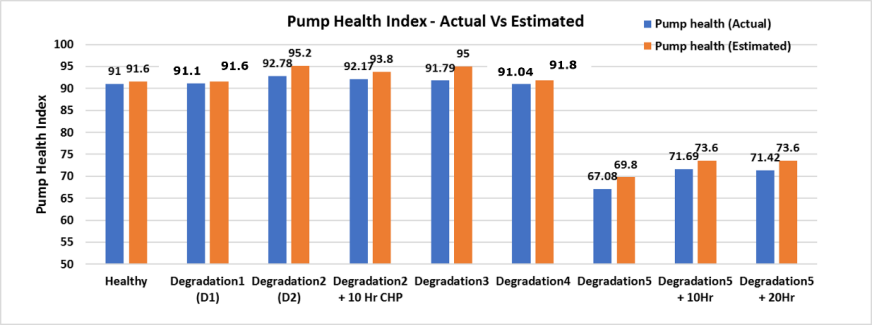


Figure 8: The PHI estimation performance for different test cases

4.4. Hydrocraft Piston Pump Health Prediction

Once the PHI is estimated, the values are stored and are used for predicting the next 'm' PHIs. The stored PHIs are modeled parametrically, and the next 'm' values are predicted using this model. With every incremental instance, a new PHI is obtained. The predicted PHIs are updated accordingly. The remaining useful life is estimated for the predicted PHI, once it approaches the set threshold.

The PHI is calculated every second. As a result, the model fitting, and the prediction are accomplished at the same rate at which the PHI is calculated. However, the horizon of the forecasting algorithm is dependent on the number of samples to be predicted in the future. The accuracy of the forecast is also dependent on the predicted terms. The accuracy decreases with increasing predicted samples. The predictive algorithms are evaluated using historical data. The scheme to test the algorithms is demonstrated in Figure 9.



Figure 9: Predictive algorithm evaluation schema

The performance of the predictive algorithm was validated for two horizons (35 days and 150 days) using the data from a 66 cc Hydrocraft pump which was recorded for a long duration. The prediction accuracies for two horizons are shown in Figure 10 and Figure 11, respectively. The mean errors are 0.69% and 1.14%, respectively. The error increases with increasing the prediction horizon.

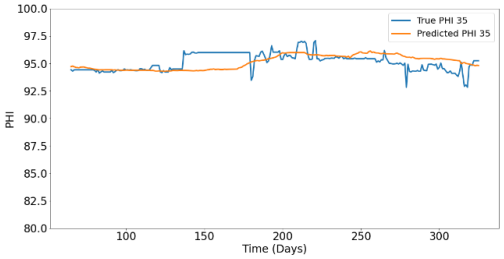


Figure 10: Prediction of PHI for 35 days compared to actual PHI (Mean percentage error = 0.69%)

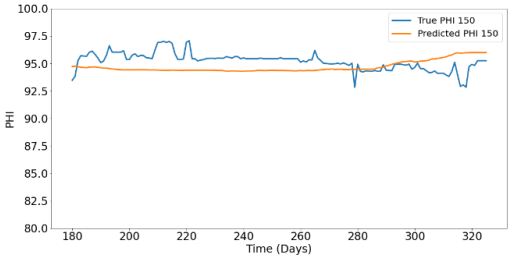


Figure 11: Prediction of PHI for 150 days compared to actual PHI (Mean percentage error = 1.14%)

5. EMBEDDED SYSTEM DEVELOPMENT AND CLOUD ARCHITECTURES

The advent of Industry 4.0 is enabling industries to migrate to IoT-based cloud infrastructures. Deriving actionable inferences from sensor data with estimation and prediction capabilities drive the success of Industry 4.0. However, these solutions require cloud connectivity to enable easier implementation of algorithms and hence add complexity, cost, data insecurity, and latency. For a real-time application, edge hardware is still the preferred mode due to the lower latency of data transfer and the possibility of real-time computation. However, edge computation limits the usage of a plethora of advanced algorithms that can be implemented on cloud-based servers.

To leverage the real-time and on-premises computation capability of edge hardware, we use classical algorithms. However, the PLC coders support the IEC 61131 international standard. As a result, the available predictive model was not usable. Furthermore, due to the limited available computational resources of such devices, the developed algorithms needed to be fast and efficient. The developed algorithms were implemented on different embedded controllers. The implementation framework with the Danfoss Plus+1® MC024™ controller is shown in Figure 12. For this implementation, we acquired data from the PLC of an existing test stand at our facility using a controller area network (CAN). The controller computes the current value of the PHI, along with its forecast after 30 days and 150 days. The current value along with the predicted values are then sent to Danfoss Plus+1® DP730 display for visualization.

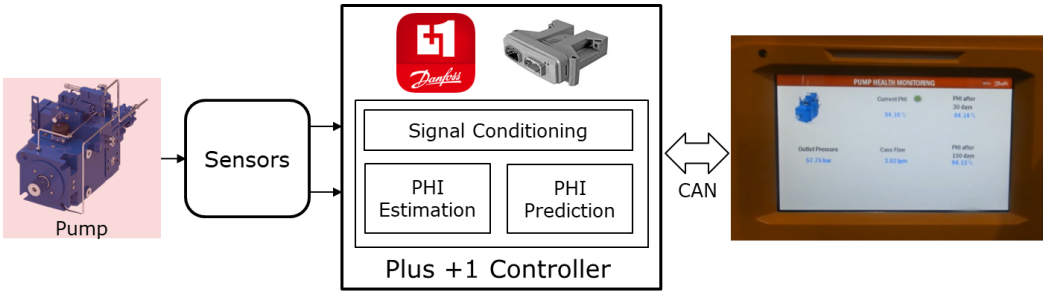


Figure 12: Implementation of the algorithm on Danfoss controller

6. CONCLUSIONS AND FUTURE SCOPES

In this article, we present a system for estimating and predicting the health of open-circuit piston pumps. The developed algorithms are suited for embedded applications that have constraints in terms of both space and computational capability. The concept of PHI is introduced in this work, which can indicate the comprehensive health of the pump. The verification of the proposed algorithms has been

carried out in the lab by recreating the real-field scenarios. Exhaustive validation testing was carried out to demonstrate the efficacy of the method. The error of the algorithms is below the 2% limit for all the tests.

With the recent technological advancements, there is scope for improvement of the accuracy further using advanced time-series and machine learning tools. The PHI is unique and can be computed in real-time. It can also predict the future health state of the pump. Additionally, it is also envisaged to develop algorithms that can identify the defects of piston pumps. This solution would play a critical role in reducing the total cost of ownership and enable benefits from predictive maintenance.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support from members involved in the various stages of this work. Support from Danfoss leadership team is greatly appreciated for providing a conducive environment and encouraging the research and development activities surrounding this technology.

NOTATIONS

Variable	Description	Units
Q_{Th}	Theoretical flow	$[\mathrm{m^3 min^{-1}}]$
Q_m	Main flow	$[\mathrm{m^3 min^{-1}}]$
Q_c	Leakage flow	$[\mathrm{m^3 min^{-1}}]$
η	PHI	[1]
P	Pressure	[bar]
P^{rated}	Rated pressure	[bar]
ω	Speed	$[\mathrm{min^{-1}}]$
d	Displacement	$[\mathrm{cm^3}]$
β	Coefficient of control flow	[1]

REFERENCES

[1] A. K. Samanta, A. Routray, S. R. Khare, and A. Naha, "Minimum distance-based detection of incipient induction motor faults using Rayleigh quotient spectrum of conditioned vibration signal," *IEEE Trans Instrum Meas*, vol. 70, pp. 1–11, 2021.

[2] S. Tang, Y. Zhu, and S. Yuan, "An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal," *Eng Fail Anal*, vol. 138, p. 106300, 2022.

[3] W. Jiang, Z. Li, S. Zhang, T. Wang, and S. Zhang, "Hydraulic pump fault diagnosis method based on EWT decomposition denoising and deep learning on cloud platform," *Shock and Vibration*, vol. 2021, pp. 1–18, 2021.

[4] A. Schumacher, D. Weiss, and R. Rahmfeld, "Condition Monitoring of an Axial Piston Pump based on Graybox Modelling," in *13th International Fluid Power Conference*, 2023.

[5] I. Baus, R. Rahmfeld, A. Schumacher, and H. C. Pedersen, "Development of Methodology for Lifetime Calculation for Axial Piston Units," in *2018 Global Fluid Power Society PhD Symposium (GFPS)*, 2018, pp. 1–7.

[6] I. Baus, R. Rahmfeld, A. Schumacher, and H. C. Pedersen, "Lifetime Impact Prediction of Component Modifications in Axial Piston Units by the Failure Likelihood Assessment," in *12th International Fluid Power Conference*, 2020, pp. 323–331.

[7] C. Li, S. Zhang, Y. Qin, and E. Estupinan, "A systematic review of deep transfer learning for machinery fault diagnosis," *Neurocomputing*, vol. 407, pp. 121–135, 2020.

[8] S. Kulkarni and A. Guha, "Equation-based domain knowledge utilization into neural network structure and learning," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 232, no. 10, pp. 1275–1291, 2018.

[9] N. J. Keller, "Condition Monitoring Systems for Axial Piston Pumps: Mobile Applications," Purdue University, 2020.

[10] S. Kulkarni, A. Guha, S. Dhakate, and T. R. Milind, "Distributed Computational Architecture for Industrial Motion Control and PHM Implementation," in *2019 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 2019, pp. 1–8.

[11] S. Kulkarni and A. Guha, "Industrial Internet of Things (IoT) Computational Architectures and Novel Neural Network for Smart Computing," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2019, pp. 925–930.

[12] S. Meraghni, L. S. Terrissa, S. Ayad, N. Zerhouni, and C. Varnier, "Post-prognostics decision in cyber-physical systems," in *2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*, 2018, pp. 201–205.

[13] S. Meraghni, L. S. Terrissa, N. Zerhouni, C. Varnier, and S. Ayad, "A post-prognostics decision framework for cell site using cloud computing and internet of things," in *2016 2nd International Conference on Cloud Computing Technologies and Applications (CloudTech)*, 2016, pp. 310–315.

[14] S. Chrétien, N. Herr, J.-M. Nicod, and C. Varnier, "A post-prognostics decision approach to optimize the commitment of fuel cell systems in stationary applications," in *2015 IEEE Conference on Prognostics and Health Management (PHM)*, 2015, pp. 1–7.

[15] Y. Gao, Q. Zhang, and X. Kong, "Comparison of hydraulic pump faults diagnosis methods: Wavelet vs. spectral analyses," in *ASME International Mechanical Engineering Congress and Exposition*, 2005, pp. 73–78.

[16] X. Wang, S. Lin, S. Wang, Z. He, and C. Zhang, "Remaining useful life prediction based on the Wiener process for an aviation axial piston pump," *Chinese Journal of Aeronautics*, vol. 29, no. 3, pp. 779–788, 2016.

[17] T. Shinn, R. Carpenter, and R. C. Fales, "Pump Efficiency Parameter Estimation," in *Fluid Power Systems Technology*, 2016, p. V001T01A030.

[18] L. I. Tongyang, W. Shaoping, S. H. I. Jian, and M. A. Zhonghai, "An adaptive-order particle filter for remaining useful life prediction of aviation piston pumps," *Chinese Journal of Aeronautics*, vol. 31, no. 5, pp. 941–948, 2018.

[19] M. A. Zhonghai, W. Shaoping, S. H. I. Jian, L. I. Tongyang, and W. Xingjian, "Fault diagnosis of an intelligent hydraulic pump based on a nonlinear unknown input observer," *Chinese Journal of Aeronautics*, vol. 31, no. 2, pp. 385–394, 2018.

[20] X. Wang, S. Lin, and S. Wang, "Remaining useful life prediction model based on contaminant sensitivity for aviation hydraulic piston pump," in *2016 IEEE International Conference on Aircraft Utility Systems (AUS)*, 2016, pp. 266–272.

[21] J. Yan, H. Zhu, X. Yang, Y. Cao, and L. Shao, "Research on fault diagnosis of hydraulic pump using convolutional neural network," *Journal of Vibroengineering*, vol. 18, no. 8, pp. 5141–5152, 2016.

[22] R. Guo, J. Zhou, C. Zhang, J. Zhao, and Y. Zhang, "Reliability evaluation of axial piston pump based on degradation failure," in *2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, 2017, pp. 204–209.

[23] W. Jiang, Y. Lei, Z. Ren, Z. Li, and S. Zhang, "Application of compressed sensing in remote fault diagnosis of hydraulic pump," 2018.

[24] X. Huang, Z. Yao, S. Huang, and D. Liu, "Performance Degradation Analysis of Axial Piston Pumps Based on Self-Organizing Map," in *2019 Prognostics and System Health Management Conference (PHM-Qingdao)*, 2019, pp. 1–6.

[25] "PLUS+1 Controllers: MC024-110 and MC024-112," no. AI00000011en-US0302. Oct. 2017. [Online]. Available: https://assets.danfoss.com/documents/150228/AI152886481001en-000403.pdf

[26] Danfoss Vickers, "EATON Vickers Hydrokraft Open Loop Piston Pumps Technical Catalog," 2003.

PRACTICAL REVIEW OF RELIABILITY METHODS COMBINED WITH VIRTUAL VALIDATION TECHNIQUES TO SHIFT LIMITS OF TODAY'S HYDROSTATS

Dr. Stefan Haug[1*], Dr. Martin Petzold[2], Dr. Johannes Steinwede[1], Friedemann Nordt[2], Dr. Matthias Blankenberg-Teich[1]

[1]*Entwicklung Funktionsgruppen und Digitalisierung, Bosch Rexroth AG, Glockeraustrasse 4, 89275 Elchingen*
[2] *Entwicklung Punpen und Motoren offener Kreislauf, Bosch Rexroth AG, An den Kelterwiesen 14, 72160 Horb am Neckar*

* Corresponding author: Tel.: +49 7308 82-2050; E-mail address: Stefan.haug3@boschrexroth.de

ABSTRACT

Ensuring sufficient technical reliability is a key factor in the market success of industrial products. However, for the off-highway industry, especially for battery-powered vehicles, the constant trend towards higher power densities of fluid-mechatronic components means an increase in component loads and a reduction in the necessary reliability reserves. To improve key performance indicators such as speed range and pressure limits, while considering the required performance level, deep technical understanding must be systematically transferred into new designs and software.

Innovative hydraulic pumps with modern electronic controls increase mobile machine productivity and reduce energy consumption simultaneously. By utilizing high-end multiphysics simulation within a design for reliability (dfr) approach, it is possible to achieve a robust design for stability at high dynamics using only electronic feedback from a swivel angle sensor signal, without mechanical feedback. Additionally, it is possible to achieve much smoother operation, especially at the beginning of the actuator movement.

Ensuring the benefits of variable pumps and meeting new system requirements, such as high speed and low speed levels with high torque, is becoming increasingly relevant, especially for electrically driven machines. By employing leading-edge simulation of flow fields, cavitation, and thermo-elastohydrodynamics, it is possible to meet these challenging demands.

The overall goal is to make the best possible use of the machine-specific operating limits by ensuring reliability and considering the load/stress on the hydraulic components. This holistic approach has been proven through the first serial system validations.

Keywords: hydraulic, pumps, hydrostats, reliability, virtual validation, electronic control

1. INTRODUCTION

Modern mobile machines are undergoing a continuous technological transformation [1,2,3]. The Off-Highway sector is experiencing increasing demands, growing electronic control, and the introduction of new primary drive systems. The integration of electric motors in mobile hydraulic applications, in

particular, presents significant changes and opportunities [4]. This raises questions for the hydraulic components industry regarding the requirements for their products and the future role of hydrostatics. It is expected that a portion of drive systems will be purely electric, while a substantial portion will remain hydrostatic. In the case of work functions, hydraulic systems will continue to be essential due to the need for compact linear movements and decentralized users.

Discussions on alternative displacement principles have gained momentum in recent years. New displacement principles are being proposed as alternatives to current axial piston units for specific applications [5,6]. However, due to higher costs and technical challenges, none of these designs are suitable for the broad Off-Highway machine market.

So, what requirements are changing or need to be met for successful placement in a competitive displacement market? According to the authors of this paper, these requirements include:

1. Providing optimal solutions that meet classical key performance indicators for current and future requirements.
2. Addressing increasing demands, particularly in terms of speed variability and acoustics.
3. Offering customers maximum flexibility and individuality in their load cycles.
4. Ensuring reliability and robustness.

2. ANSWERING TODAY'S DEMANDS

Expectations regarding power, power density, efficiency, and especially reliability of hydraulic components are continuously increasing [7]. However, the remaining technological potentials of current hydraulic pumps and motors require more advanced tools. Improvements through newer materials, enhanced manufacturing processes, and design changes are gradually diminishing.

In some cases, market expectations even contradict each other. For example, there are desires for higher pressure limits alongside desires for higher rotational speed limits. Without a design change, an increase in one value would require limiting the other.

The answer to these demands is "design smarter"

Our understanding of smarter design includes:

1. Increased granularity in design and development: "From product view to detail view"
2. From overlayed total load-collective to application specific demands

2.1. Solving contradictions by increasing granularity

Defining operating limits for higher complex machines like axial piston pumps and motors on machine level necessarily leads to a safe but minimal operating space. The logical reason behind lies in the multitude of damage mechanisms and failure modes, which limit the possible operating space in partially contradicting manners (Figure1).

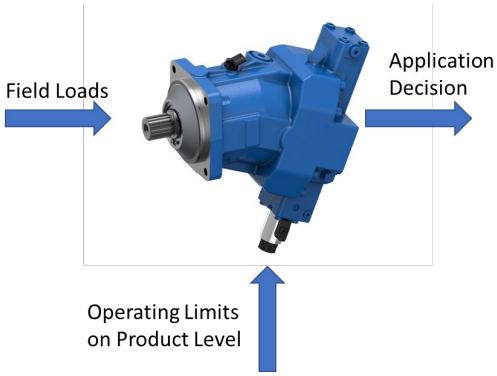


Figure 1: Classical Design Approach on Machine Level

By increasing the granularity during design and application, the operating space can be increased but still kept safe. The approach is shown in Figure 2.

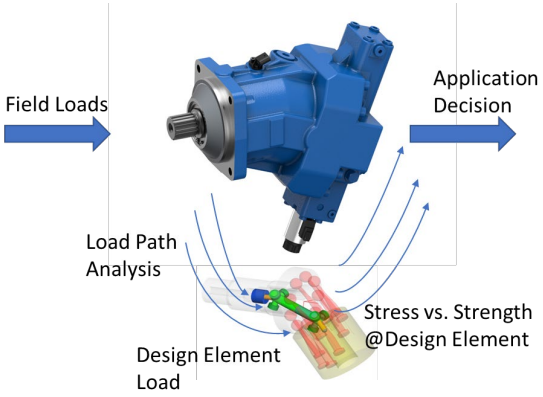


Figure 2: Improved Approach on Design Element Level

To give a simplified example: Many tribological design elements fail roughly proportional to a friction-power called property:

$$P_{friction} = p_{contact} * v_{contact}$$

Friction-power is defined as product of local contact pressure (usually driven by operating pressure) and contact speed (usually driven by rotational speed). It can be understood as the amount of energy per time that triggers many failure modes.

By defining operating limits in the local domain “critical friction-power at design element”, operating limits extend without increasing the risk of product failures. This is caused simply by allowing higher maximal values in operating points, where only one of the two influencing stresses reaches high values.

Exemplary from real live: Pressure operating limits for some high pressure units in closed loop were extended by 5,7% for specific breaking operations without increasing the risk of failures.

This increased granularity during the design process creates benefits on many levels:

- Increased cross-product knowledge management: Knowledge on design elements can be used for multiple products.

- Deeper know-how per element can be achieved without increasing design and research efforts by increasing knowledge reuse due to element view.
- “Elephant cut into slices”: Design element view leads to easier understanding and research capabilities, e.g., speeding up problem solving cases or specific product improvements.

2.2. From overlayed sum-collectives to fitting solutions

As stated in the previous example: Simply by shifting the focus from machine level to design element level, operating limits can be increased under the condition of deeper knowledge of field loads.

Therefor the second pillar of “design smarter” is the improved use of field load information. By creating load path models, design element loads can be derived from field loads. This enables a closed loop from application demand (stress) and product capability (strength) during product application. In combination with a modular design approach, application specific solutions create smartly designed products. Practically, the approach leads to a shift of the application process: From comparing simplified field loads with product catalogue data to a data driven application decisions.

- Reliability can be approached as design domain and therefore be optimized.
- Operating limits can be increased without increasing field failures.
- Enabling system optimizations by comparing element loads.
- The design process opens to enable the use of increasingly available knowledge on field loads.

3. USE OPPOSING OPTIMIZATION APPROACHES BY ELECTRONIC CONTROL

In addition to approaches within the product development process, electronically controlled units for mobile applications offer opportunities to optimize performance and meet customer requirements. Electronic control systems open up two crucial development paths:

- Expansion of the solution space by introducing additional degrees of freedom in components and systems. Shown in Figure 3.
- The ability to selectively influence a larger number of these parameters.

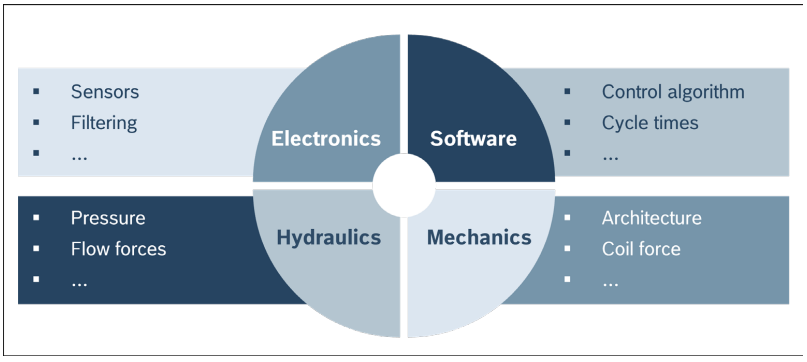


Figure 3: Involved domains for the electronic open circle pump control [10]

By implementing this approach, it becomes feasible to integrate an additional opportunity: by reducing the swivel angle, it is possible to increase the speed limit of a pump (in this case for mobile applications) by 35 %. This adjustment enables an increase in speed within the overall system and allows for an increase in power in other areas.

Traditional hydraulic-mechanical solutions either prove to be highly complex or influence the system in a way that corresponds to continuous parameter shifts across all operating ranges. These hydro-mechanical design elements are either optimized for a specific operating point or represent a

compromise for the entire machine.

The implementation of electronic open circuit pump control (eOC) offers a partial solution to these compromises. One common conflict, for instance, is finding a suitable compromise between dynamics and stability. Additionally, dynamics and cavitation erosion phenomena present two distinct problems that require optimization approaches in opposite directions.[7]

The implementation of electronic pump control in open circuit greatly simplifies the management of these compromises, as optimal parameters only need to be determined or selected for each individual operating point. Customers benefit in particular from the possibility of tailoring parameters to their needs, especially by operating point-dependent parameter fields, which allow optimization for finally many operating points. This optimization can be applied to both control strategies and dynamics (Figure 4).

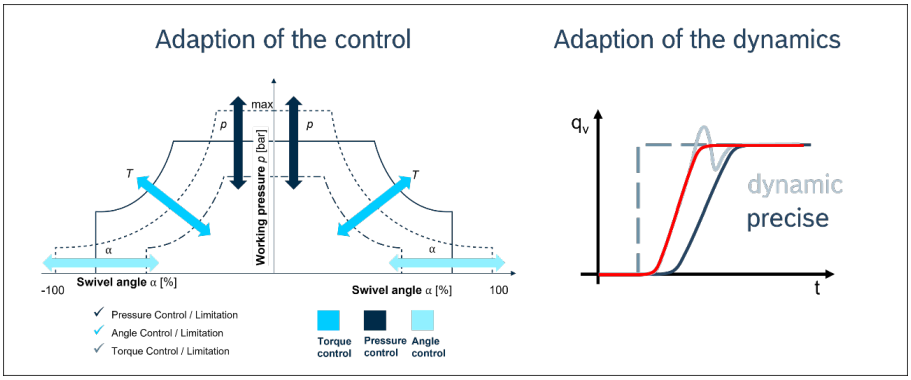


Figure 4: Adaptive application specific control

Moreover, it is possible to achieve specific parameterizations in dynamics, allowing for discrete parameterization in milliseconds or adaptive dynamics. This results in a system that is better suited to meet customer requirements [8]. In contrast, conventional nozzles in a system tend to reduce dynamics across the entire operating range [9]. In comparison, an eOC-controlled pump can achieve a 30 % lower swivel out time depending on the operating point, leading to a reduced swivel time of up to 40 ms in certain applications.

Furthermore, the installed capacity can be utilized more effectively. With better knowledge of the operating conditions, the required reserves can be determined more accurately. This enables a reduction in the overall amount of reserves, resulting in a more efficient use of resources. Additionally, the installed capacity can be accessed as needed, providing flexibility in meeting demand.

4. USE VIRTUAL VALIDATION TO SHIFT TODAY'S LIMITS

To ensure the reliability and robustness of hydraulic pumps, a profound technical understanding of the hydraulic system and its components is crucial. In this regard, advanced simulation techniques serve as a foundation for the virtual validation of products during the development process and for exploring the possibilities of shifting operating limits. These techniques enable thorough investigations and analysis of the system's behavior.

4.1. Set up of the virtual validation process

In order to reach possible potentials in main described requirements, three major points should be

noted.

1. It is essential to have reliable tools and methodologies that can be used to verify the design against the specified requirements.
2. These tools must be cleverly combined and connected in a way that generates an added value.
3. The process must be created in a “modular and self-improving” way.

Rexroth recognized the importance of this early on and has implemented a multi-stage, systematic validation process that combines testing and simulation tools of varying quality in an appropriate manner. In the field of simulation, there are different approaches to address specific questions, as depicted in Figure 5. In many cases, the simulation domains interact with each other, requiring a multidomain approach to solve them. Depending on the problem at hand, the user can perform quick and simple computational fluid dynamics (CFD) calculations to estimate the suction limit, or conduct highly detailed cavitation erosion calculations that accurately map the timing reversal process when necessary.

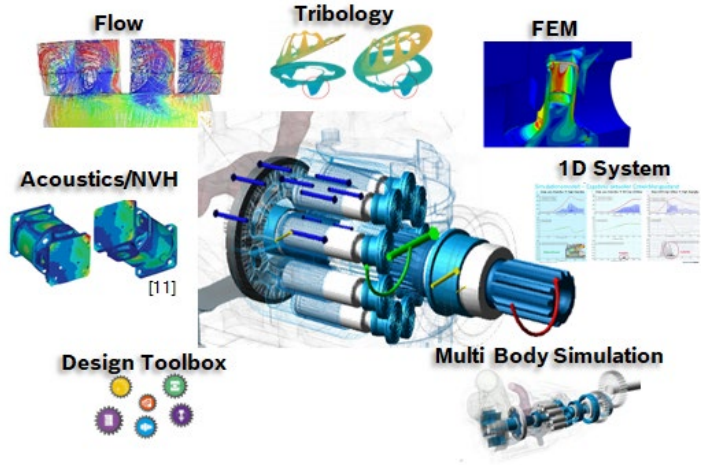


Figure 5: Simulation overview

With the assistance of a tool, users can generate automated validation plans instantly and easily submit and track test and simulation tasks, as illustrated in Figure 6.

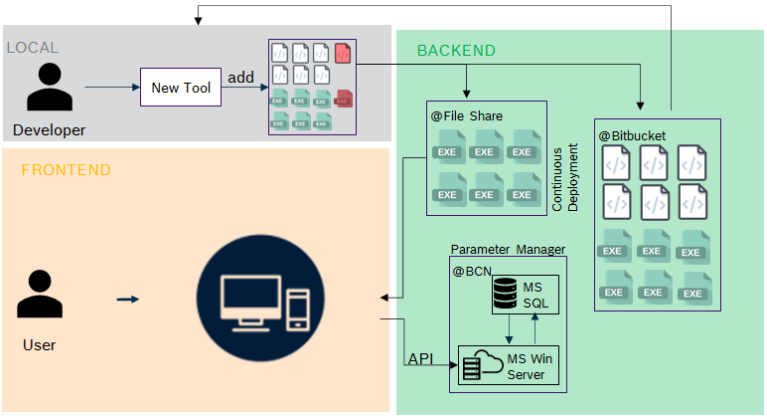


Figure 6: Simulation approach

This streamlined approach allows for increased work efficiency and minimizes unnecessary steps.

Importantly, this tool enables the entire validation process to be measured, allowing for the identification of potential areas for improvement and facilitating continuous enhancements to the overall process.

4.2. Usage in series products and results

This validation process has already been successfully applied to products such as the new axial piston unit for the speed-variable industrial market and the new axial piston unit for the mobile market. With the aid of this new validation process, the limits of these products could be explored in much greater detail within a significantly shorter timeframe. Additionally, this approach has allowed for the identification of areas where defining simulation standards would be beneficial in improving cycle time and quality.

To put these improvements in figures:

The overall efficiency of the new industrial unit could be improved by 3 % by using our simulation standards. Another significant result is the improvement of power density by 88 % as well as nearly doubling the maximum speed. At the same time the NVH level regarding pressure pulsation could be reduced by 66 %.

Similar improvements were especially made in the new mobile unit. Here, the overall efficiency was increased by remarkable 7 %. Additionally, the power density improved by 49% and the maximum speed increased by 15%.

5. CONCLUSION

Acknowledging the enduring significance of hydraulics and the emergence of new technologies like electrified solutions, some of which may even replace hydraulic systems, it is imperative to propel their evolution based on the outlined aspects. The primary requirements are growing more intricate and demand comprehensive solutions. Reliability, customization for individual customers, and the enhancement of key performance indicators form the bedrock of economic success. There continue to be opportunities for performance improvements in established hydrostatic designs (Figure 7). However, the path to achieving these improvements is becoming more complex. The initial products within this holistic framework have been implemented and will validate this approach in the forthcoming years.

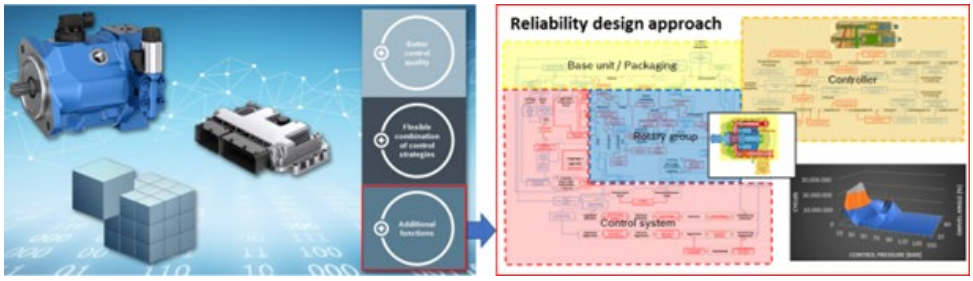


Figure 7: Shifting Limits of today's hydrostats by reliability methods and virtual validation

NOMENCLATUR

$P_{friction}$	Frictional power	W
$p_{contact}$	Contact pressure	N/mm²
$v_{contact}$	contact speed	m/s
dfr	design for reliability	

eOC	electronic open circuit pump control
CFD	computational fluid dynamics
NVH	Noise vibration harshness

REFERENCES

[1] Sprink J, Duensing Y, Figge F, Haas C, Holzer A, Hucko S, Opgenoorth A, Tappeiner Z, Schmitz K (2023) O+P Fluidtechnik 1-2/2023, 12-16 https://digital.oup-fluidtechnik.de/o-p-fluidtechnik-1-2-2023/67492430/13

[2] Frerichs, Ludger (ed.) (2023): Yearbook Agricultural Engineering 2022. Braunschweig: Institute of Mobile Machines and Commercial Vehicles. https://leopard.tu-braunschweig.de/servlets/MCRFileNodeServlet/dbbs_derivate_00051226/jahrbuchagrartechnik2022_jahrbuch2022.pdf

[3] Weber J (2022): Maha Fluid Power Conference Purdue University, May 17-19,2022, Purdue. https://vbp-b40.webspace.tu-dresden.de/downloads/blog-veroeffentlichungen-des-verbundprojektes_dateien_tud/2022-05-19_Maha-FPC_B40_Weber.pdf

[4] Fassbender D (2023): Towards Energy-Efficient Electrified Mobile Hydraulics Considering Varying Application Conditions, Tampere. https://trepo.tuni.fi/bitstream/handle/10024/149626/978-952-03-2952-5.pdf

[5] Bucher Hydraulics GmbH (2024): Neue AX-Kolbenpumpen und –Motoren beenden bisherige Probleme hydraulischer Antriebe, https://www.bucherhydraulics.com/datacat/files/Katalog/Fachartikel/AX/Fachartikel_AX%20Kolbenpumpen%20und%20-Motoren_100-de.pdf

[6] Danfoss A/S (2024): Digital Displacement® pumps, https://www.danfoss.com/en/products/dps/pumps/digital-displacement-pumps/digital-displacement-single-and-multiple-outlet-pumps/#tab-overview

[7] Krüßmann M, Mohn G (2022): New Powerful hydraulic platforms for digitalized mobile machines, 13th International Fluid Power Conference, 13. IFK, March 21-23, 2022, Aachen 911-920

[8] M. Brand, „Electronification of Work Hydraulics", ATZ heavy duty worldwide, Bd. 12, S. 34–37, 2019.

[9] Mohn G, Tetik S (2023): Electronic Open Circuit (eOC) – Elektronische Regelung von Axialkolbenpumpen im offenen Kreislauf, 9. Fachtagung hybride und energieeffiziente Antriebe für Mobile Arbeitsmaschinen; Karlsruhe 2023

[10] Bosch Rexroth AG (2017), Mobile 2017, Ulm. https://www.boschrexroth.com/de/de/transforming-mobile-machines/

[11] Bittner U. (2013), "Strukturakustische Optimierung von Axialkolbeneinheiten", Dissertation, Karlsruhe.

A STUDY ON THE EFFECTS OF BODY DEFORMATION ON THE PERFORMANCE OF EXTERNAL GEAR MACHINES

peer reviewed

Ajinkya Pawar*, Andrea Vacca[1], Manuel Rigosi[2]

[1]*Maha Fluid Power Research Center, Purdue University, Lafayette, IN, USA*
[2]*Casappa S.p.A., Parma, Italy*

* Corresponding author: Tel.: +1 765-426-1654; E-mail address: pawar10@purdue.edu

ABSTRACT

The energy efficiency of external gear pumps (EGPs), similar to all positive displacement machines used for high-pressure applications, is significantly influenced by the power losses occurring in the lubricating interfaces that seal the internal displacement chambers. Therefore, it is crucial to account for these interfaces accurately, when developing predictive simulation tools. However, the literature has suggested various modelling approaches for EGPs, with different assumptions regarding the analysis of these interfaces. This makes it challenging for a designer or a researcher to determine what are the essential physical domains needed for properly modelling the lubricating interfaces and assess the EGP performance.

This paper addresses the above research question by leveraging a comprehensive simulation tool (Multics-HYGESim) developed by the authors' research team to compare the effect of different modelling assumptions. HYGESim includes tribological considerations pertaining to the meshing of the gears, the lubricating films at the tooth tip interfaces, at the journal bearings, and at the lateral interfaces. It also considers realistic fluid properties, including the effects of cavitation and aeration, mixed lubrication effects, as well as material deformation effects for the gears, lateral bushings and the EGP housing. Deformation of the internal parts of an EGP is related to the internal balancing features and it is strongly coupled with the instantaneous pressurization of the pumping volumes. For this reason, a realistic quantification of these effects is difficult in simulation.

Using a commercial EGP design as a reference, with known experimental volumetric and hydromechanical efficiency, this paper demonstrates how predictions can vary based on different simulation assumptions regarding body and lubricating film behaviours. Results are discussed starting from a basic rigid-body assumption that considers only body motion and analytical formulations of lubricating interfaces, to simulation model cases of progressively increasing in complexity to account for deformations of gears, bushings and housing. The results show that consideration of deformation effects allow more accurate prediction of power losses and efficiencies of the pump while simulations carried out without deformation considerations approximate the leakages and the power losses at the lateral lubricating interface though can predict the fluid dynamic performance. These findings will offer valuable insights to EGP designers, enabling them to understand the strengths and limitations of different modeling assumptions on the prediction of EGP behavior, especially regarding the effects of body deformation.

Keywords: External gear pumps, Pump Efficiency, Simulation, Power loss, Body deformation

1. INTRODUCTION

Despite the basic concept for external gear pump (EGP) dates back centuries, EGP technology has been constantly evolving to follow technical trends. In fact, the EGP design is among the most cost-effective ones for positive displacement machines [1]. EGPs also have good performance features, in terms of efficiency, durability, resistance to different forms of contamination. It is therefore natural for an engineer or a researcher to consider application of an EGP to address new applications. For this reason, EGPs can be seen in automotive, aerospace, industrial, mining, construction and agricultural application. And their design is constantly adapting to new application requirement. A significant example is the recent electrification trend in mobile applications, which is bringing to new EGP designs that can better cope with electric prime movers [2]. Therefore, there is a clear need of providing the technical community with the most accurate possible simulation tools, that are able to properly account for the key physical aspects affecting the operation and the energy efficiency performance of EGPs.

Figure 1 shows a typical bearing-block type pressure-compensated EGP design used for high pressure operation. Other EGP designs, with or without lateral compensation, do exist, as mentioned in [1,3], but still involving the same fundamental principles. The fluid is displaced across a pressure difference using rotation of externally engaging spur gears. The radial loads acting on gears are supported using journal bearings while lateral compensating bearing blocks are pressure against the lateral surface of gears to minimize the leakages in axial direction. From a physical perspective, the operation of an EGP can be divided into three domains. First, the fluid domain comprising of the volumes inside the machine such as the inlet and outlet volume, spaces between the gear teeth through which the main displacing action occurs. Second, the solid domain, which comprises of the floating bodies such as gears, the lateral bushings and the housing. Third, the lubricating interface domain, which comprises of thin fluid films between floating bodies which function as load support mechanisms. Additionally, these domains interact with each other leading to a multi-domain coupled operation of an EGP which is challenging to model using simulation techniques.

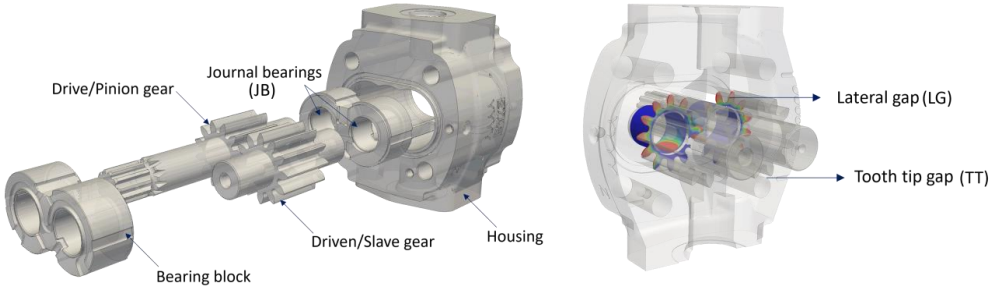


Figure 1: Illustration of a typical pressure compensated EGP and its lubricating interfaces

Over the past few decades, various simulation methodologies of varying complexity have been proposed, that aim to analyse the physical phenomena in one or more operating domains of an EGP as well as the interactions between them. The related papers, as it will be further discussed, shows good correlation with experiments for all these approaches. Therefore, it can be challenging for a designer to understand the strengths and the limitation of each approach.

Models to analyse of fluid displacing action of EGPs in the fluid domain can be classified into three main categories namely analytical models, lumped parameter models and computational fluid dynamics (CFD) models. *Analytical models*, similar to the work by Manring and Kasaragadda [4], Ivantysyn and Ivantysynova [1], provide a theoretical description of fluid

flow inside EGPs and can estimate the kinematic flow ripple using the geometrical information, but do not consider the effects of compressibility of the fluid. The lumped parameter models, such as the works by Vacca and Guidetti [5], Borghi et. al [6], divide the fluid domain into a number of control volumes, and solve mass conservation as well as fluid transport equations to determine the fluid flow and pressurization behavior inside the pump. *Lumped parameter models* are very powerful in simulation of EGPs as they are computationally inexpensive and can help estimating physical phenomena such as outlet pressure ripple, loads acting on the floating components due to fluid pressure, effects of cavitation and aeration etc. The estimation of loads on gears also allows coupling of the fluid domain pressures with the micromotion of gears which can affect the fluid domain predictions significantly. One of the major drawbacks of these models is that they significantly approximate the behavior of lubricating interfaces inside the pump and use analytical approximations as well as lubricating film gap assumptions while estimating the power losses arising from these interfaces. Therefore, when it comes to torque efficiency prediction, these models cannot be considered most accurate. *CFD models*, as seen in the works of Castilla et. al [7], Frosina et al. [8], divide the fluid domain into infinitesimal meshes and solve the partial differential equations in a distributed mesh domain using numerical techniques and can accurately estimate the fluid flow and pressure behavior inside an EGP. These studies allow accurate estimation of fluid-related phenomena such as cavitation and incomplete filling of the machine, local effects of fluid inertia as well as bubble collapse and damage at the cost of higher computational resources. 3D CFD models can also approximate the lubricating interface behaviour with fixed geometry film gap to obtain better estimation of the volumetric losses in the EGPs. Although 3D CFD models can show capability of coupling the micromotions and deformation effects with fluid dynamic evaluation, the CFD studies on EGPs do not consider these effects.

The modelling of lubricating interface domain in EGPs including the above aspects is crucial to accurately predict both the volumetric and the torque losses, as well as to assess the durability of a given design. The behavior of the lubricating interfaces is significantly affected by the film gap height distribution which is a function of the motion and the deformation of the floating bodies forming these interfaces and thus is highly coupled with the solid body domain behavior of the pump. For this reason, models addressing only the fluid domain might not be sufficient, and *tribological models* should be introduced. Different models of this kind have been developed with different assumptions while analysing the behavior of various lubricating interfaces. Taking an example of the lateral gap lubricating interface in EGPs, studies such as the ones by Borghi et al. [9] assume a predefined gap height distribution to determine the film pressure distribution from solution of Reynolds equation. Dhar and Vacca [10] showed the effect of coupling the axial motion of the lateral bushing on the lateral film behavior assuming that the lateral bushing is always under the state of force balance. The same authors (Dhar and Vacca [11]) extended the model to include the effects of pressure and thermal deformation of lateral bushing in the lubricating film analysis. Thiagarajan and Vacca [12] extended this work to include the mixed lubrication regime modelling and effect of surface roughness on the lateral lubricating film power losses. A recent work from the authors' team [13] introduced the multi-domain simulation tool Multics-HYGESim, which allows the simultaneous/coupled analysis of different domains of an EGP. The tool, which will be further described in Section 2, has a modular structure, and can accommodate new modules such as the housing deformation and the gear meshing ones presented respectively in [14,15].

The current study aims to leverage the capabilities of Multics-HYGESim to analyse effects of different physical aspects associated with the operation of an EGP in terms of simulations. Simulation options of Multics-HYGESim, purposely introduced in this research, allows changing the complexity of the assumptions taken to carry out the analysis of the EGP on various levels.

For example, the model can analyse only the fluid domain along with consideration of gear micromotion, whilst simplifying the lubricating domain with analytical solutions of the journal bearing and using a constant gap height laminar equation to model lateral gap interface. While, the most physically and computationally complex simulation possible involves consideration of lubricating interfaces using solution of the Reynolds equation, along with effects of linear and tilting motions of gears and bushings, deformations of gears, bushings and the housing.

Using this flexibility of the simulation tool, four simulation cases with increasing complexity will be considered. The overall behavior of the reference pump will be compared across these four cases in terms of various parameters that are important for the EGP designers and manufacturers to prototype new high performing units. These parameters include comparison of differences in housing wear, overall Tooth Space Volume (TSV) pressurization, outlet flow/pressure ripple, volumetric and frictional losses from lubricating interfaces across four cases. Based on the comparison, the authors aim to establish a correlation between the effects of motion and deformation of different bodies on the performance characteristics of the machine. Finally, the results from the simulation tool in all four cases will be compared with experimental data of volumetric and hydromechanical efficiency, outlet pressure ripple and housing wear to understand the importance of consideration different physical effects during simulation, on the EGP performance prediction. This should help a designer to make proper decisions on the assumptions to make when simulating an EGP.

The remaining part of the paper is divided into three sections. Section 2, describes the simulation tool and the details of analysis of different domains in brief, followed by description of cases considered along with the underlined assumptions and level of complexity of physical effects that are evaluated during simulation. Section 3 describes the simulation operating conditions and gives detailed comparison of reference machine performance parameters and provides insights into correlation of physical aspects considered in simulation with the results. Section 4 talks about important conclusions and provides recommendations regarding physical aspects to consider during simulation to predict the different components of the EGP performance.

2. METHODOLOGY

2.1. Multics – HYGESim overview

Figure 2 shows the schematic of different solvers of the simulation tool and the domains of the machine they model.

Fluid Domain Modeling

The evaluation of fluid dynamic behavior takes place using the fluid dynamic solver where the pump domain is divided into multiple control volumes or tooth space volumes. Using a lumped parameter approach, pressure build-up equation (eq. 1) is solved for each control volume to predict the pressurization inside the pump. The flow between the control volumes through various geometrical connections such as the ones due to grooves on the bushings is modelled using orifice equation (eq. 2). The fluid dynamic solver also models the leakages at the gear tip – housing interface using a Couette-Poiseuille equation (eq. 3). A geometrical pre-processor is run to determine the variation of TSV volume, the time derivative of TSV volume, connection areas and diameters between different control volumes, and other geometrical parameters required by the simulation model as a function of rotation angle of the shaft gear. More details regarding the approach can be found in [5, 13].

$$\frac{dP_i}{dt} = \frac{K_T}{V_i}\left(Q_{in} - Q_{out} - \frac{dV_i}{dt}\right) \tag{1}$$

$$Q_{i,j} = sign(p_i - p_j)C_f\Omega\sqrt{\frac{2(p_i - p_j)}{\rho}} \tag{2}$$

$$Q_{i,j} = \left(-\frac{h_{i,j}^3(p_i - p_j)}{12\mu L} + \frac{v_{i,j}h_{i,j}}{2}\right)b \tag{3}$$

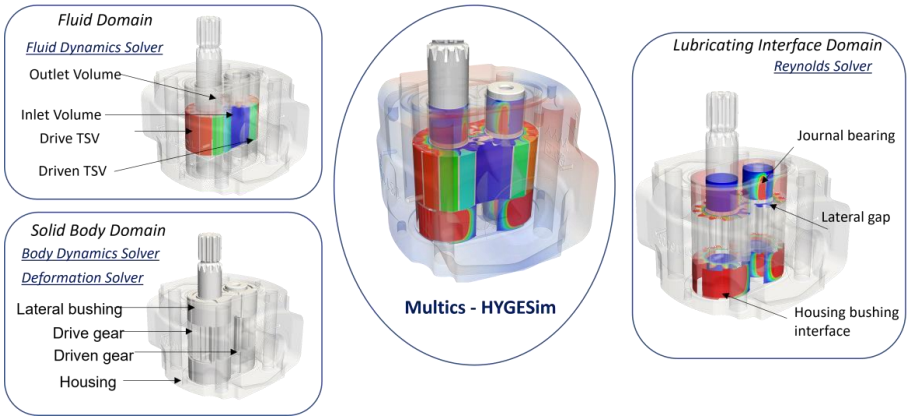


Figure 2: Different solvers in Multics-HYGESim simulation tool

Lubrication Domain Modeling

The behavior of lubricating films at journal bearing, lateral gap and casing-bushing interface is modelled by the Reynolds solver, which solves the universal mixed Reynolds equation (eq. 4). To estimate the contact forces based on the roughness profile of the bodies, an approach proposed by Lee and Ren as described in [13] is used which relates the gap height information of the film to the contact pressure in the regime of asperity contact. The mixed lubrication modelling allows evaluation of viscous as well as asperity friction (eq. 5) and accurate evaluation of power losses from the lubricating interfaces. To evaluate the meshing losses, a curve-fit relation proposed by Manne et al. [15] is used, to obtain which the authors simulated the EHL contact considering mixed lubrication effects.

$$\nabla\cdot\left(\phi_p\left(\frac{K_T h^3}{12\mu}\nabla\rho\right)\right) = \nabla\cdot\left(\rho\overrightarrow{v_m}\left(\phi_R R_q + \phi_c h\right)\right) + \nabla\cdot\left(\rho\phi_s R_q\left(\frac{\overrightarrow{v_t} - \overrightarrow{v_b}}{2}\right)\right) + \frac{\partial\left(\rho\left(\phi_R R_q + \phi_c h\right)\right)}{\partial t} \tag{4}$$

$$\overrightarrow{F_{fric}} = \int_{\Omega_{JB}}\left(\frac{\mu\overrightarrow{V_D}}{h}\left(\phi_f + \phi_{f,s}\right) + \frac{\phi_{f,p}\,h\overrightarrow{\nabla P}}{2}\right)dA + \mu_{Asp}p_c dA\overrightarrow{v_D} \tag{5}$$

Solid Domain Modeling

The solid domain modelling includes the body dynamics solver and the deformation solver. Body dynamics solver computes the linear and angular rigid body motion of floating bodies, i.e. the gears and the lateral bushings, by solving Newton's second law. The loads acting on bodies from TSV pressures, lubricating interfaces, contact forces as well as frictional forces are considered while evaluating the motion of the bodies. The deformation solver uses the influence matrix approach, which is based on finite element analysis under reference loads and

scaling the obtained deformation based on actual loads as described in [10, 13], to determine the elastic deformation of the gears, bushings as well as the housing. The deformation of the gears can also be determined analytically by using Euler-Bernoulli beam theory and is chosen for this work as it is more computationally inexpensive.

The next section describes the simulation cases that are considered for the purpose of this study and the assumptions as well as physical phenomena considered for each case and the method of evaluation. It will also try to provide reasoning behind choosing these particular cases.

2.2. Simulation cases analysed

Figure 3 gives an overview of the simulation cases considered for the proposed study. For each case, the reference EGP will be simulated at corner operating conditions encompassing the overall operating region of the machine.

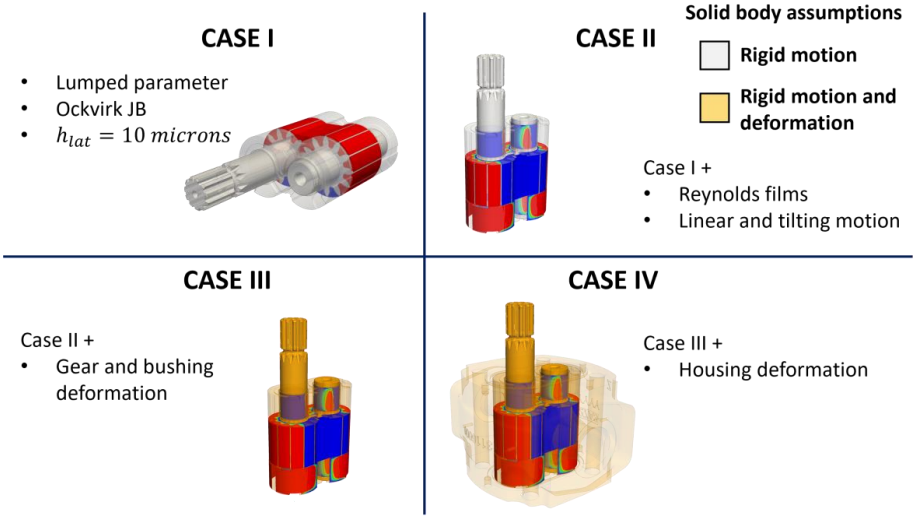


Figure 3: Overview of simulation cases considered

Case I: Lumped parameter simulation with analytical films

This case considers only the evaluation of fluid domain with rigid body micromotion of gears. The assumptions under this case indicate simulation framework used for multiple previous studies in analysis of EGMs using lumped parameter models [5, 6], which has been shown to predict the performance behaviour of the machine including the hydromechanical [16] and volumetric efficiency, housing wear [5] etc. As the simulation framework in this case uses only 0D equations, the analysis is computationally inexpensive and can be used for quick performance prediction of the machine with considerable accuracy and therefore is considered as one of the cases analysed in this study. Important assumptions involve the maximum discharge coefficient of orifices, which is considered as 0.7, and the constant lateral gap height of 10 microns, which is equivalent to average gap height at the lateral gap interface considering deformation and tilting effects. All other parameters are evaluated either based on geometry (such as areas and hydraulic diameters of orifice connections) or analytical equations (friction force evaluation at lubricating interfaces).

Case II: Rigid simulation considering lubricating domain

This case considers Reynolds films to model the film behavior at journal bearing, lateral gap and housing bushing interface. The simulation can estimate linear and tilting motion of the

lateral bushing in addition to the motion of gears due to accurate pressure force and moment evaluations from the lubricating interfaces. This simulation case does not consider deformation effects, but can allow comparison with case I, with more accurate consideration of leakages and viscous losses, especially from the lateral gap interface without assumption of any rigid gap height. One of the advantages of this case is accurate estimation of forces and moments on the lateral bushing from the gear side, allowing accurate design of balancing features.

Case III: Simulation considering deformation of gears and bushings

This case estimates the pressure deformation of gears and lateral bushings and their effects of the lubricating interface behavior. Consideration of deformation of these bodies allow accurate estimation of the power losses from journal bearing and lateral gap interfaces as it does not involve any assumptions pertaining to lubricating film behavior except consideration of symmetric behavior of top side and bottom side films.

Case IV: Simulation considering deformation of gears, bushings and housing

In addition to gears and bushings, this case estimates the pressure deformation of the housing body of the machine. In the current study, this case aims to analyse the effect of housing deformation on pump performance parameters such as housing wear, frictional losses and leakages at lubricating interfaces. Compared to case III, this case removes the symmetric assumption of top and bottom side films, and simulates all 12 lubricating interfaces shown in **Figure 2**, providing the complete picture of EGP operating under isothermal conditions, but results in the most computationally expensive simulation.

3. RESULTS AND DISCUSSIONS

The reference unit, PHP20QW20.20 is simulated at 5 different operating conditions shown in **Table 1**. ISOVG-46 is considered as the operating fluid while the temperature is assumed constant at 50^{o} C. Nominal dimensions of the unit are considered for simulation. **Table 2** presents the EGP parameters. The simulations consider an initial wear-in simulation (at operating conditions suggested by the manufacturer), to derive a wear-in housing profile that is be used for the subsequent simulations of **Table 1**.

Table 1: Parameters of the reference EGP

EGP parameters	Value
Displacement	22.38 cc/rev
Type	Spur involute
Maximum operating speed	3500 RPM
Maximum operating pressure	250 bar, 300 bar (intermittent)

Table 2: Operating Conditions

Terminology	Operating speed and pressure
Low Speed Low Pressure (LSLP)	500 RPM, 50 bar
Low Speed High Pressure (LSHP)	500 RPM, 250 bar
Medium Speed Medium Pressure (MSMP)	1500 RPM, 150 bar
High Speed Low Pressure (HSLP)	2500 RPM, 50 bar

High Speed High Pressure (HSHP)	2500 RPM, 250 bar

3.1. Housing wear-in and fluid dynamic comparison

Figure 4a indicates the housing wear predicted by the simulation model and comparison with experimentally measured wear profile. Consideration of deformation is important as seen from case III and case IV results to predict the amount of wear on the housing surface. Case IV considers deformation of internal housing surface while determining the wear, allowing estimation of worn region at different axial sections as shown in **Figure 4b**, leading to a better prediction of the trend of housing wear with angle. Both cases III and IV overestimate the wear, though the magnitude of the wear lies within manufacturing tolerance region. There is a possibility that model is overpredicting the deformation of the lateral bushing and journal bearings which dictate the positions of gears inside the housing and the magnitude of the wear, as the influence matrix approach assumes a linear dependence of deformation of applied pressure. The contact zone between lateral bushing and the housing, as well as the journal bearing liner and the bushing can exhibit non-linear deformation effects, which the proposed model does not consider.

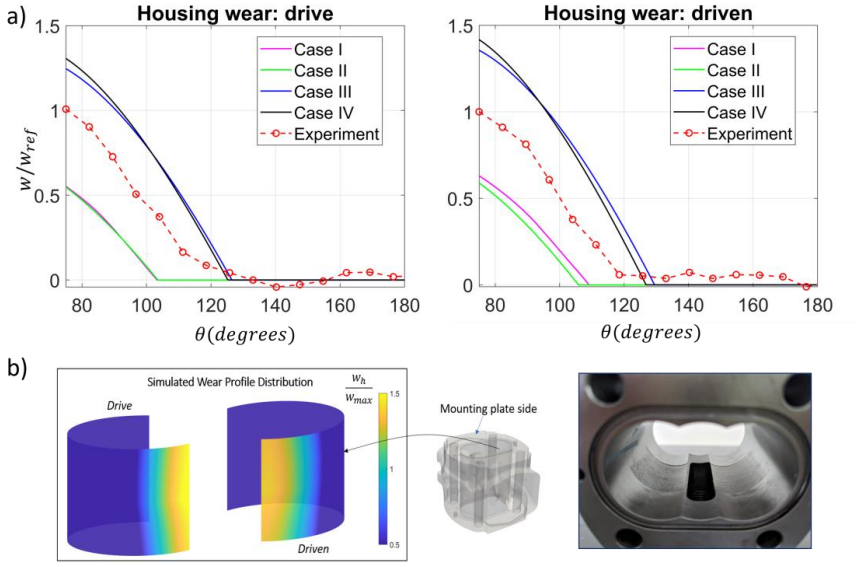


Figure 4: a) Comparison of housing wear for all cases **b)** 3D worn region prediction from case 4

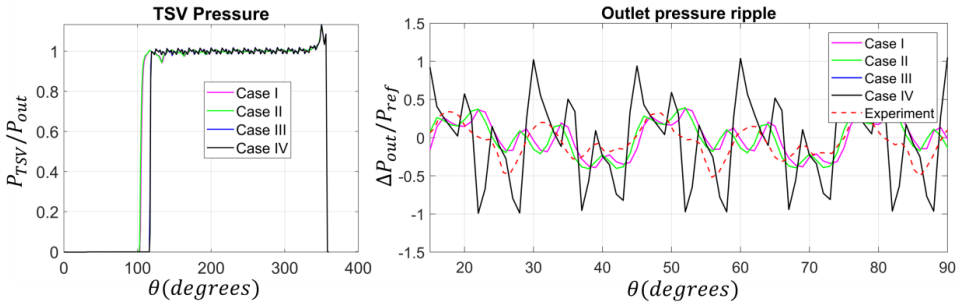


Figure 5: Comparison of TSV pressure and outlet pressure ripple for all cases

The housing wear profiles obtained from simulation are given as input to the model for corresponding cases. Figure 5 shows comparison of TSV pressurization and outlet pressure ripple of the pump at HSHP operating condition. Clearly, cases I & II, and cases III & IV show very similar fluid dynamic behavior. The positions of the gears inside the housing affect the fluid dynamic behavior significantly. Due to deformations of the bushings and the housing, the gears in cases III and IV are pushed towards the suction side by a larger magnitude. Additionally, the deformations of bodies negate the journal bearing effect seen in cases I and II at high speed, which tends to increase the minimum gap at the journal bearings leading to an inefficient sealing at tip-housing interface. Therefore, cases I and II show early TSV pressurization from tip leakages, while cases III and IV show TSV pressurization when the chamber is exposed to backflow groove leading to different fluid dynamics compared to cases I and II. Comparing the outlet pressure ripple, cases I and II show a peak-to-peak magnitude match of ripple compared to experiments while cases III and IV show higher ripple magnitude. The magnitude of deformation also affects the outlet pressure ripple and can be a reason for overprediction of peak-to-peak magnitudes.

3.2. Power loss and leakage prediction

Figure 6a indicates distribution of power loss predicted from different lubricating interfaces of the EGP at HSHP operation by the simulation model. Case I only involves lumped assumptions but can estimate the tooth tip as well as meshing losses accurately but fails to estimate losses at JB and LG interfaces due to not capturing the mixed lubrication effects and assumptions of constant gap at LG interface. Case II estimates the JB losses on approximately the same level as cases III and IV indicating mixed lubrication effects are more important in JB loss prediction as compared to deformation. Case IV can capture the losses arising from both top and bottom side films including the effects of tilting of the gear and shows a lower JB loss compared to case III, indicating importance of considering the asymmetric effects at this interface. Case II overestimates losses from LG interface. This can be explained from the lateral gap height comparison between cases II and III shown in **Figure 7.** As case II does not consider deformation, the value of gap height is seen to be very small throughout the film leading to higher frictional losses, while case III indicates that the minimum gap height region will be concentrated in the zone towards the suction groove. As case III does not consider tilting of gears, it predicts a higher zone of contact between gears and bushings, while case IV indicates that the losses from the LG interface are on the same order of meshing interface losses at HSHP operating condition.

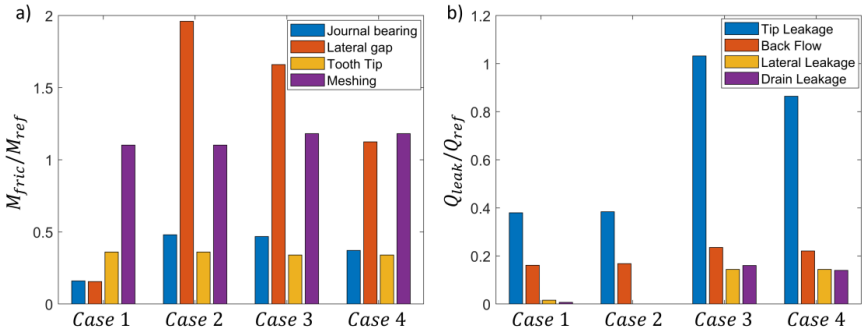


Figure 6: a) Comparison of frictional moment and **b)** leakages for all cases

Figure 6b presents comparison of leakages predicted by the simulation model for all cases at LSHP operating condition. It can be seen that all cases indicate radial or tooth-tip leakages to be the major contributor to volumetric losses in this type of EGPs. As tip leakages are gear position dependent, cases I & II and cases III & IV indicate similar values of radial leakages.

Case IV considers axial variation of housing wear as well as tilting of gears, providing a higher sealing towards the suction side while predicting tip leakages showing a smaller value compared to case III. Case I approximates the lateral and drain leakages using analytical expressions. Case II predicts negligible lateral and drain leakages due to very small gap heights as shown in **Figure 7** while cases III and IV show presence of leakages at the lateral film due to higher gap heights that are resulting due to deformation of the bushings. All cases show similar magnitude of backflow. As seen from **Figure 8a**, for LSHP operating condition, cases III and IV show better prediction the volumetric efficiency of the pump indicating that leakages are captured with higher accuracy. As the leakages at lateral gap interface depend highly on the gap height at these interfaces, the magnitude of deformation can significantly influence the leakages predicted by the simulation model. The same is true for tooth tip/radial leakages.

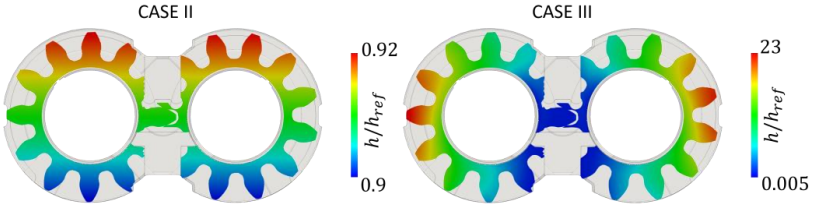


Figure 7: a) Comparison of gap height at lateral gap interface for cases II and III

3.3. Efficiency comparison

Figure 8a shows the comparison of volumetric efficiency with experimental values. Cases I and II overpredict the efficiency at lower speeds and underpredict the efficiency at higher speeds. One of the reasons for underpredicted efficiency by these cases would be smaller amount of movement of gears towards the suction resulting from no deformation considerations can reduce the volume of the fluid the pump actually displaces. The predicted volumetric efficiency from Case I is also a function the value of the lateral gap assumed. Consideration of deformations of all bodies can best help in predicting the volumetric efficiency trends as well as magnitudes. **Figure 8b** indicates importance of considerations of lubricating interface losses while predicting hydromechanical efficiency.

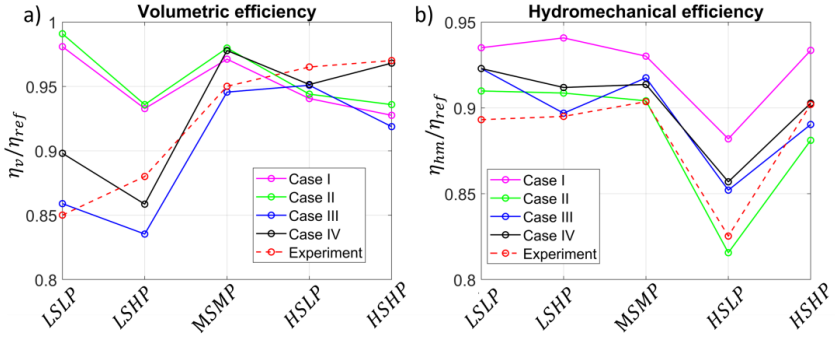


Figure 8: Comparison of EGP efficiencies predicted by simulation model with experiments for each case

4. CONCLUSION AND OUTLOOK

The current study uses the Multics.HYGESim simulation model developed by the authors to drive a study assessing the importance of different modeling assumptions to simulate External Gear Pumps (EGPs). The research intends to provide the technical community insights on the

consequences of including or neglecting physical aspects that determine EGP operation. The particular focus is on the different ways the lubricating gaps (i.e. the tooth tip gap and the lateral gap) of an EGP can be modeled. Four different cases with increasing simulation complexity are considered:

Case I: Lumped parameter simulation with analytical films (Average t_{sim} = 2 minutes)
Case II: Rigid simulation considering lubricating domain (Average t_{sim} = 2 hours)
Case III: Simulation considering deformation of gears and bushings (Average t_{sim} = 2 days)
Case IV: Simulation considering deformation of gears, bushings and housing (Average t_{sim} = 5 days)

Although the simulation parameters for each Case could be further refined/adjusted for a better matching with experimental data, in this research all simulations were performed with the same modeling setting with respect to geometrical data, and coefficients to be used in modeling equations (orifice coefficient, transition between laminar and turbulent flow, etc). The main finding is that all the models have more or less pronounced limitations in matching available experimental data. It has to be remarked that the experiments could be reproduced more accurately by using actual measured geometrical clearances, rather than the using the nominal geometry of the reference machine as performed in this study. In particular, case 1 underestimates the losses from both leakages and frictional losses. Case 2, underpredicts the gap height at lateral gap leading to lower leakage and higher frictional losses. Case 3 shows a good prediction of gap height at the lateral gap interface, but considers symmetric assumptions and ignores gear tilting leading to higher tooth tip leakage prediction. Case 4 considers housing deformation which allows simulating the axially asymmetric nature of the EGP with inclusion of gear tilting in the model. These deformation considerations help in getting a good match of volumetric as well as hydromechanical performace of the EGPs.

What is most important to point out is that the designer should select the simulation assumptions based on the phenomena he/she would like to observe. In case the focus is on pressure ripple or average volumetric efficiency, a simplified simulation such as in Case I can be sufficient. However, it must be understood that there will not be any capability of detecting features such as chances of wear due to material contacts and accurate frictional loss as well as leakage through lubricating interfaces. Increasing the model complexity does not necessarily increase model accuracy in parameters that can be predicted by a simpler model, such as volumetric efficiency, but it can be applicable especially towards accurate modeling and design of lubricating interfaces in EGPs.

NOMENCLATURE

P_i	Pressure of i^{th} TSV
t	Time
K_T	Isothermal bulk modulus of fluid
V_i	Volume of i^{th} TSV
$Q_{in,out}$	Flow entering/exiting i^{th} TSV
$Q_{i,j}$	Flow through flow connections between i^{th} and j^{th} TSV
C_f	Flow coefficient of the orifice
Ω	Area of the orifice
ρ	Density of the fluid
μ	Viscosity of the fluid
L	Width of the tooth tip
$h_{i,j}$	Tooth tip gap height at tooth separating i^{th} and j^{th} TSV

$v_{i,j}$ Tooth tip velocity at tooth separating i^{th} and j^{th} TSV
$\varphi_{P,R,C,S}$ Flow factors for mixed lubrication modeling
h Gap height at lubricating interface
$\overrightarrow{v_{m,t,b}}$ Velocities of surfaces bounding the lubricating interface

REFERENCES

[1] Ivantysyn, J., and Ivantysynova, M., 2003, “Hydrostatic pumps and motors”. New Dehli, India: Tech Books Int.

[2] Zappaterra F., Vacca A., Sudhoff S.D., "A Compact Design for an Electric Driven Hydraulic Gear Machine Capable of Multiple Quadrant Operation," Mechanism and Machine Theory, 177, 10504, 2022.

[3] Stryczek J, Fundamentals of designing hydraulic gear machines, PWN, 2021

[4] Manring, N., D., and Kasaragadda, S., B. 2003, “The Theoretical Flow Ripple of an External Gear Pump," Journal of Dynamic Systems, Measurement, and Control, vol. 125, no. 3, p. 396.

[5] Vacca, A., Guidetti, M., 2011, "Modelling and Experimental Validation of External Spur Gear Machines for Fluid Power Applications," Elsevier Simulation Modelling Practice and Theory, 19 (2011) 2007–2031.

[6] Borghi, M., Zardin, B., Specchia, E. 2009, “External Gear Pump Volumetric Efficiency: Numerical and Experimental Analysis”, SAE Technical Papers, doi: 10. 4271/2009- 01- 2844.

[7] Castilla, R., Gamez-Montero, P-J., Ertrk, N., Vernet, A., Coussirat, M., Codina, E., 2010, “Numerical simulation of turbulent flow in the suction chamber of a gearpump using deforming mesh and mesh replacement”, Int. J. Mech. Sci. 52 (10) 1334–1342.

[8] Frosina, E., Senatore, A., Rigosi, M., 2017, “Study of a high-pressure external gear pump with a computational fluid dynamic modeling approach”, Energies 10 (8) 1113.

[9] Borghi, M, & Zardin, B. "Axial Balance of External Gear Pumps and Motors: Modelling and Discussing the Influence of Elastohydrodynamic Lubrication in the Axial Gap." Proceedings of the ASME 2015 International Mechanical Engineering Congress and Exposition. Volume 15: Advances in Multidisciplinary Engineering. Houston, Texas, USA. November 13–19, 2015.

[10] Dhar S., Vacca A., A novel CFD – Axial motion coupled model for the axial balance of lateral bushings in external gear machines, Simulation Modelling Practice and Theory, Volume 26, 2012, Pages 60-76, ISSN 1569-190X

[11] Dhar S., Vacca A.,, A novel FSI–thermal coupled TEHD model and experimental validation through indirect film thickness measurements for the lubricating interface in external gear machines, Tribology International, Volume 82, Part A, 2015, Pages 162-175

[12] Thiagarajan, D., Vacca, A., 2017, "Mixed Lubrication Effects in the Lateral Lubricating Interfaces of External Gear Machines: Modelling and Experimental Validation," Energies, 10(1), 111.

[13] Ransegnola, T., Zappaterra, F., Vacca, A., 2022, "A Strongly Coupled Simulation Model for External Gear Machines Considering Fluid-Structure Induced Cavitation and Mixed Lubrication," Applied Mathematical Modelling, 104, 721-749.

[14] Pawar A., Vacca A., and Rigosi M., "Prediction of housing wear-in in external gear machines considering deformation effects, ASME Symposium of Fluid Power and Motion Control, 2023

[15] Manne VHB, Vacca A. And Singh K., 2023 A Curve-fit Traction Coefficient Relation of Mixed EHL Line Contact with Hydraulic Fluid and Steel Surfaces, Tribology Transactions, 66:2, 364-380

[16] Rituraj R., Vacca A., Rigosi M., Modeling and validation of hydro-mechanical losses in pressure compensated external gear machines, Mechanism and Machine Theory, Volume 161, 2021

A NOVEL PULSATION COMPENSATOR FOR DISPLACEMENT MACHINES

peer reviewed

Gudrun Mikota[1]*, Bernhard Manhartsgruber[1]

[1]*Institute of Machine Design and Hydraulic Drives, Johannes Kepler Universität Linz, Altenbergerstrasse 69, A-4040 Linz*

* Corresponding author: Tel.: +43 732 2468-6521; E-mail address: gudrun.mikota@jku.at

ABSTRACT

This paper presents a mechanical compensator for hydraulic pressure pulsations induced by a pump or motor. The compensator is based on the combination of pump or motor mass moment of inertia and the torsional compliance of the adjacent coupling. According to the displacement volume, angular deflections of the resulting oscillator correspond to volumes which can compensate for geometric or dynamic pulsations at the natural frequency of the oscillator. Two design concepts are suggested for the coupling. Three models are set up for various configurations of the overall system, accounting for limited inertia of an electric motor and the influence of an outlet pipeline. Model parameters are determined for a 32 ccm radial piston pump and a 250 ccm axial piston pump. Torsional damping from viscous friction is roughly estimated and considered by different levels. At the pump outlet, frequency response functions between flow rate excitation and pressure response are calculated. The results show that the compensation effect is relevant for both pumps and particularly robust with respect to resonances in the hydraulic system.

Keywords: pump, pressure pulsations, compensator, torsional oscillations, coupling

1. INTRODUCTION

The periodic operation of hydraulic pumps and motors usually leads to a pulsating flow which depends on the kinematics of the working principle. Moreover, the precompression of displacement chambers need not match the pressure at the outlet and periodic impulsive flow rates can result from this effect. The overall flow rate pulsation of the machine excites pressure pulsations in the attached hydraulic system. This results in structural vibrations and noise, and may even lead to cavitation. Various countermeasures have been suggested, at the cost of efficiency or additional components [1].

In structural mechanics, vibration compensators are supposed to counteract harmonic forces by natural vibrations at the same frequency. In hydraulic systems, such compensators can be understood to counteract harmonic flow rates by their natural vibrations. A Helmholtz resonator [2], for instance, extracts harmonic flow rates in its attachment point. Its natural frequency is defined by the inertia of the attachment pipe and the compliance of the resonator chamber.

For practical inclusion in hydraulic systems, more compact compensators are required [3]. A special solution is described and investigated in [4]. Compared to a Helmholtz resonator, the inertia of the fluid in the attachment pipe is replaced by a lumped mechanical mass, while the resonator chamber provides a hydraulic spring.

In this paper a flow rate pulsation compensator realised by the inertia of the machine rotor and the torsional compliance of its coupling is proposed [5]. The compensator constitutes a torsional oscillator which counteracts the pulsating flow from the machine into the hydraulic system. This

requires a lightly damped coupling with a torsional stiffness selected to match the natural frequency of rotor and coupling with the pulsation frequency. The ratio between flow rate pulsation and angular velocity amplitudes is given by the displacement volume of the machine. By appropriate angular movements, the machine is supposed to compensate its own pulsation. Contrary to separate devices like Helmholtz resonators, pulsation compensation takes place in the displacement chamber itself, which is the source of flow rate pulsation. Moreover, maximum compactness is achieved.

The novel compensator concept promises a reduction of pressure pulsations in hydraulic systems at the expense of a tailored torsional stiffness of the coupling attached to pump or motor. Alternatively, the concept offers more freedom in the design of the machine itself as it may be exchanged against other measures for pulsation reduction (e.g. the number of pistons).

Nevertheless, the practical realisation of the novel compensator faces several challenges. Hydraulic damping is always present and reduces the effectiveness of pulsation compensation. A broadband effect will only be achieved for certain relations between machine and coupling properties. In order to investigate the feasibility of the concept, a frequency domain study is carried out for two pumps of different type and size. In particular, viscous friction in sealing gaps is quantified, torsional vibrations of the entire pump drivetrain are modelled, and the dynamic behaviour of a transmission line is taken into account.

2. SYSTEM CONFIGURATION

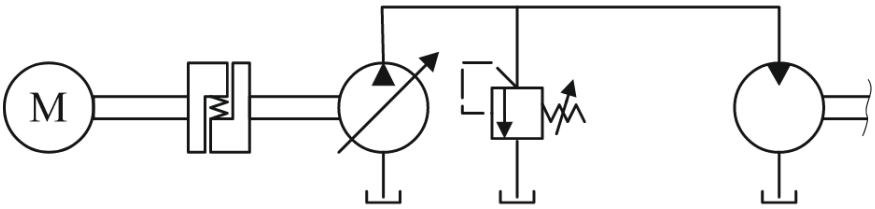


Figure 1: Schematic of overall system

As one out of many application examples, **Figure 1** shows the schematic of an open circuit transmission with a variable displacement pump and a fixed displacement motor. The pump is driven by an electric motor at constant speed, with which it is connected by a coupling. Compared to the pump, the electric motor's mass moment of inertia will be several times higher so that the first torsional natural frequency of the drivetrain can be approximated from the inertia of the pump rotor and the torsional stiffness of the coupling.

In the hydraulic system, the pump excites pressure pulsations at a fundamental frequency depending on rotational speed and number of displacement chambers, which are usually accompanied by several higher harmonics. In principle, all torsional natural modes of the pump drivetrain could be used for pulsation compensation. The corresponding natural frequencies must match the pump harmonics to be compensated. In order to compensate a single pulsation frequency, the torsional stiffness of the coupling is chosen for an appropriate match with the first natural frequency of the pump drivetrain.

Coupling design should avoid any torsional damping since the torsional oscillator includes unavoidable viscous damping at the pump rotor. Depending on the type of pump, the rotor's mass moment of inertia may vary. In particular, this is the case for radial piston pumps, where a mean value will be used for compensator design. Of course, the system which is driven by the hydraulic motor could also be equipped with an appropriate coupling to realise another compensator.

3. COUPLING CONFIGURATONS

The compensator coupling requires a relatively low torsional stiffness in the absence of damping, as opposed to conventional couplings of comparable size. Therefore, two different concepts are suggested with individual advantages of each.

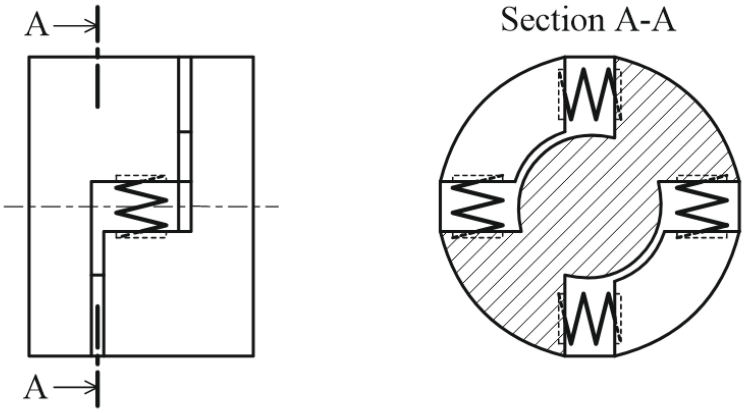


Figure 2: Coil spring concept

Figure 2 shows a concept where two coupling halves engage in such a way that four tangential coil springs can be arranged in between, which define the torsional stiffness of the coupling. This concept is favoured by the fact that it covers a wide range of stiffness while broken springs are easily exchanged.

The concept in **Figure 3** uses a radial arrangement of flat rectangular beams, whose tangential deflection under torque defines the torsional stiffness. The beams are connected to an inner and an outer ring, all of which can be manufactured as an integral part. The outer ring is held in the carrier flange by a shrink fit, while the inner ring may be connected to the pump shaft by a key. This concept seems compact enough to be integrated in a pump housing and could also be cascaded for the compensation of several pulsation frequencies. An example for two pulsation frequencies is shown in **Figure 4**. For such an arrangement, it may be necessary to adapt the intermediate flange inertia to obtain feasible values of both torsional stiffnesses.

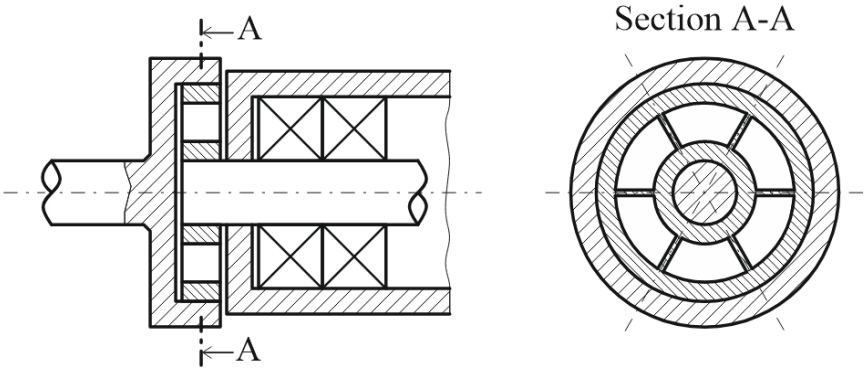


Figure 3: Beam concept

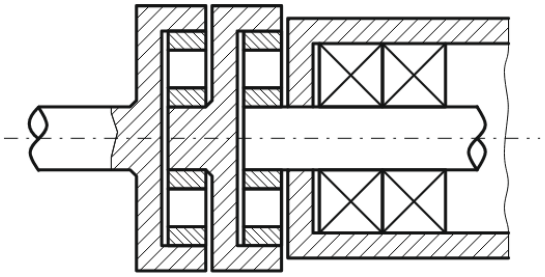


Figure 4: Cascaded beam concept

4. FREQUENCY DOMAIN MODEL

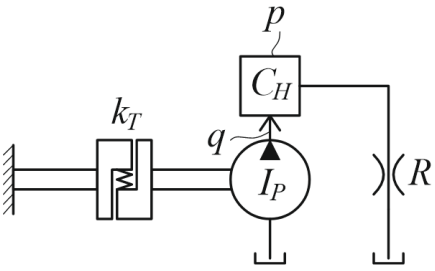


Figure 5: Schematic of system with fixed boundary condition at electric motor end and hydraulic system with throttle towards tank

A first numeric assessment of the novel compensator concept is made by frequency domain models of the systems in **Figure 5**, **Figure 6**, and **Figure 7**. To demonstrate the compensation effect with a minimum number of components, **Figure 5** includes a coupling with a fixed boundary condition at the electric motor end, a pump rotor, the inevitable hydraulic capacity at the pump outlet, and a throttle towards tank. Due to viscous fluid friction, torsional damping acts on the pump rotor. With the torsional stiffness k_T of the coupling, the mass moment of inertia I_P of the pump rotor, the torsional damping constant c_T, the hydraulic capacity C_H, the hydraulic resistance R, and the displacement volume V_P of the pump, the time domain model reads

$$\begin{bmatrix} I_P & 0 \\ -\frac{V_P}{2\pi} & C_H \end{bmatrix} \begin{bmatrix} \ddot{\varphi}_P \\ \ddot{p} \end{bmatrix} + \begin{bmatrix} c_T & 0 \\ 0 & \frac{1}{R} \end{bmatrix} \begin{bmatrix} \dot{\varphi}_P \\ \dot{p} \end{bmatrix} + \begin{bmatrix} k_T & \frac{V_P}{2\pi} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varphi_P \\ p \end{bmatrix} = \begin{bmatrix} 0 \\ \dot{q} \end{bmatrix}, \tag{1}$$

where q is the flow rate excitation into the hydraulic capacity C_H, p is the pressure in this capacity, and φ_P is the rotation angle of the pump. Equation (1) represents a linear two-degree-of-freedom model with asymmetric system matrices due to hydraulic-mechanical interaction. The second order mechanical subsystem describes an oscillator, while the hydraulic subsystem is only first order. The corresponding frequency domain model reads

$$\left(-\omega^2 \begin{bmatrix} I_P & 0 \\ -\frac{V_P}{2\pi} & C_H \end{bmatrix} + \mathrm{i}\omega \begin{bmatrix} c_T & 0 \\ 0 & \frac{1}{R} \end{bmatrix} + \begin{bmatrix} k_T & \frac{V_P}{2\pi} \\ 0 & 0 \end{bmatrix} \right) \begin{bmatrix} \Phi_P(\mathrm{i}\omega) \\ P(\mathrm{i}\omega) \end{bmatrix} = \begin{bmatrix} 0 \\ \mathrm{i}\omega Q(\mathrm{i}\omega) \end{bmatrix}, \tag{2}$$

where capital letters for excitation and states denote the respective Laplace transforms, and ω is the angular frequency.

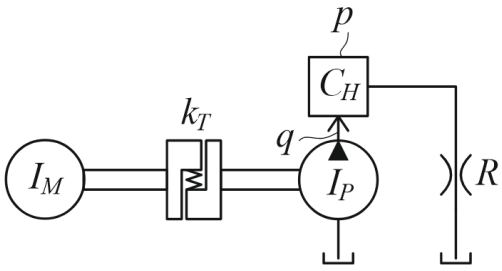


Figure 6: Schematic of system with electric motor and hydraulic system with throttle towards tank

Compared to **Figure 5**, **Figure 6** also accounts for the electric motor's limited mass moment of inertia I_M. The corresponding model also includes the Laplace transformed rotation angle Φ_M of the electric motor and reads

$$\left(-\omega^2\begin{bmatrix} I_M & 0 & 0 \\ 0 & I_P & 0 \\ 0 & -\frac{V_P}{2\pi} & C_H \end{bmatrix} + \mathrm{i}\omega\begin{bmatrix} 0 & 0 & 0 \\ 0 & c_T & 0 \\ 0 & 0 & \frac{1}{R} \end{bmatrix} + \begin{bmatrix} k_T & -k_T & 0 \\ -k_T & k_T & \frac{V_P}{2\pi} \\ 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} \Phi_M(\mathrm{i}\omega) \\ \Phi_P(\mathrm{i}\omega) \\ P(\mathrm{i}\omega) \end{bmatrix} = \begin{bmatrix} T_M(\mathrm{i}\omega) \\ 0 \\ \mathrm{i}\omega Q(\mathrm{i}\omega) \end{bmatrix}, \quad (3)$$

where T_M denotes the Laplace transform of any torque from the electric motor.

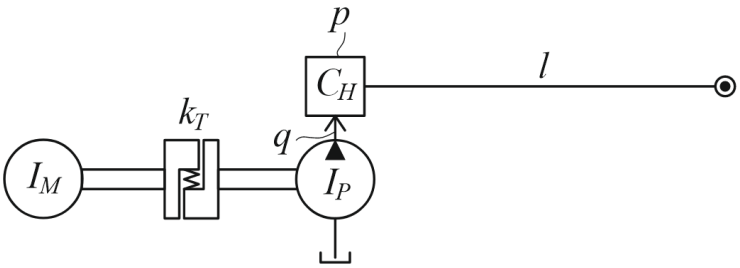


Figure 7: Schematic of system with electric motor and hydraulic system with pipeline towards constant pressure boundary condition

To account for resonances in the hydraulic subsystem, the throttle is replaced by a pipeline from the hydraulic capacity to a constant pressure boundary condition. This configuration is shown in **Figure 7**. Using the pipeline description for dynamic laminar flow as developed in [6], the system model reads

$$\left(-\omega^2\begin{bmatrix} I_M & 0 & 0 \\ 0 & I_P & 0 \\ 0 & -\frac{V_P}{2\pi} & C_H \end{bmatrix} + \mathrm{i}\omega\begin{bmatrix} 0 & 0 & 0 \\ 0 & c_T & 0 \\ 0 & 0 & S_{12} \end{bmatrix} + \begin{bmatrix} k_T & -k_T & 0 \\ -k_T & k_T & \frac{V_P}{2\pi} \\ 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} \Phi_M(\mathrm{i}\omega) \\ \Phi_P(\mathrm{i}\omega) \\ P(\mathrm{i}\omega) \end{bmatrix} = \begin{bmatrix} T_M(\mathrm{i}\omega) \\ 0 \\ \mathrm{i}\omega Q(\mathrm{i}\omega) \end{bmatrix}, \quad (4)$$

in which

$$S_{12} = -\frac{\mathrm{i}A}{\sqrt{E\rho}f(\mathrm{i}\omega)}\cot\left(\frac{f(\mathrm{i}\omega)\omega l}{\sqrt{E/\rho}}\right) \tag{5}$$

with

$$f(\mathrm{i}\omega) = \sqrt{-\frac{\mathrm{J}_0\left(\mathrm{i}r\sqrt{\mathrm{i}\omega/\nu}\right)}{\mathrm{J}_2\left(\mathrm{i}r\sqrt{\mathrm{i}\omega/\nu}\right)}}, \tag{6}$$

where J_0 and J_2 are Bessel functions of first kind, l, r, and A denote pipeline length, radius, and cross-sectional area, respectively, E is the bulk modulus, ρ the mass density, and ν denotes the kinematic viscosity of the fluid.

The torsional stiffness of the coupling is calculated in order to match the pulsation frequency f_P with the torsional natural frequency of the pump drivetrain. This results in

$$k_T = \frac{(2\pi f_P)^2}{\frac{1}{I_M} + \frac{1}{I_P}}. \tag{7}$$

A torsional damping constant c_{TG} is estimated for different types of sealing gap with gap height h. A cylindrical sealing gap appears at the inner radius r_i of the radial piston pump rotor. The relative motion of gap walls due to the angular velocity $\dot{\varphi}_P$ of the rotor causes a uniform shear stress

$$\tau_R = \nu\rho\frac{r_i\dot{\varphi}_P}{h}, \tag{8}$$

resulting in a shear torque

$$T_{\tau R} = 2\pi r_i^2 l_P\, \tau_R \tag{9}$$

over the rotor length l_P and a torsional damping constant

$$c_{TGR} = \frac{T_{\tau R}}{\dot{\varphi}_P} = 2\pi\nu\rho\frac{r_i^3 l_P}{h}. \tag{10}$$

Between axial piston pump rotor and valve plate, the sealing gap assumes the form of a disc with inner radius r_i and outer radius r_a. The shear stress

$$\tau_A = \nu\rho\frac{x\dot{\varphi}_P}{h} \tag{11}$$

depends on the intermediate radius x and causes the shear torque

$$T_{\tau A} = \int_{r_i}^{r_a} 2\pi x^2 \tau_A dx. \tag{12}$$

In this case, the torsional damping constant becomes

$$c_{TGA} = \frac{T_{\tau A}}{\dot{\varphi}_P} = \frac{\pi}{2}\nu\rho\frac{r_a^4 - r_i^4}{h}. \tag{13}$$

Accounting for comparable viscous damping from the slippers, the overall torsional damping constant becomes $c_T = 2c_{TG}$.

For a medium size radial piston pump driven by a servo motor and a large axial piston pump driven by an induction motor, model parameters are shown in **Table 1**.

Table 1: Model parameters

	Radial piston pump	Axial piston pump
I_M	0.018 kg m^2	2.55 kg m^2
I_P	0.0061 kg m^2	0.0959 kg m^2
V_P	32 cm^3	250 cm^3
l_P	40 mm	
r_i	20 mm	30 mm
r_a		70 mm
h	0.04 mm	0.03 mm
z	7	9
n	25 Hz	25 Hz
f_P	zn	zn
V_d	V_P	$V_P/2$
E	1400 MPa	1400 MPa
C_H	V_d/E	V_d/E
R	10^{11} Pa m^{-3} s	10^{10} Pa m^{-3} s
ρ	860 kg m^{-3}	860 kg m^{-3}
ν	40 mm^2 s^{-1}	40 mm^2 s^{-1}
l	1.1 m … 2.1 m	0.7 m … 1.7 m
r	7.5 mm	16 mm
A	$r^2\pi$	$r^2\pi$

5. RESULTS AND DISCUSSION

The following figures show frequency response functions between flow rate excitation and pressure response in the hydraulic capacity at the pump outlet. Different curves are shown for systems with and without compensator. To account for parameter uncertainty and damping mechanisms not included in the model, a damping constant of $10\ c_T$ is used for a compensator with light damping, while $100\ c_T$ is used for a compensator with high damping.

From the model in **Figure 5**, the frequency response functions in **Figure 8** and **Figure 9** have been obtained for radial piston pump and axial piston pump, respectively. The compensator causes a local antiresonance (151 Hz for the radial piston pump and 221 Hz for the axial piston pump), which is followed by a resonance (166 Hz for the radial piston pump and 232 Hz for the axial piston pump). Both antiresonance and resonance are more pronounced for light damping. At excitation frequencies around the antiresonance, pressure response is reduced by the compensator. The contrary happens around the resonance, whose excitation should be avoided. To use the compensation effect, the pulsation frequency of the pump should match the antiresonance. Since the torsional stiffness of the coupling is determined from Eq. (7) and the model in **Figure 5** does not account for the electric motor's limited mass moment of inertia, the model must be refined to obtain the intended match.

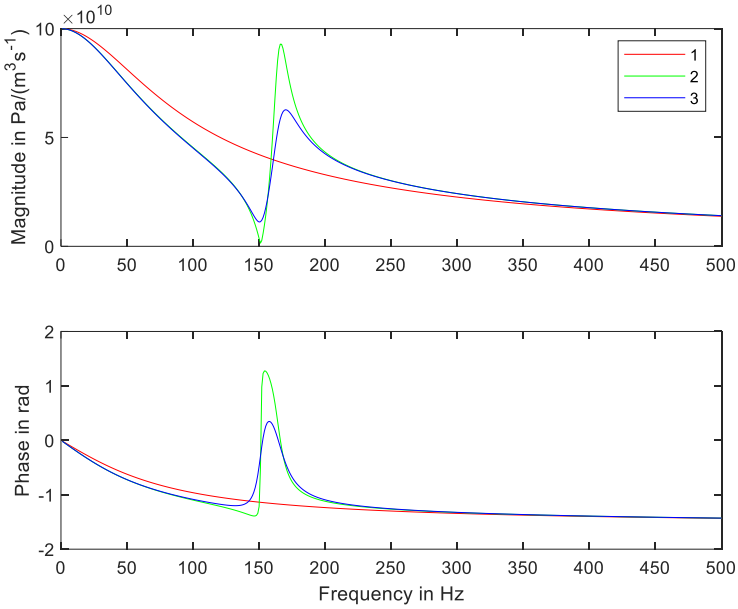


Figure 8: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, radial piston pump, pump inertia, capacity and resistance at the outlet, 1: without compensation, 2: light damping compensator, 3: high damping compensator

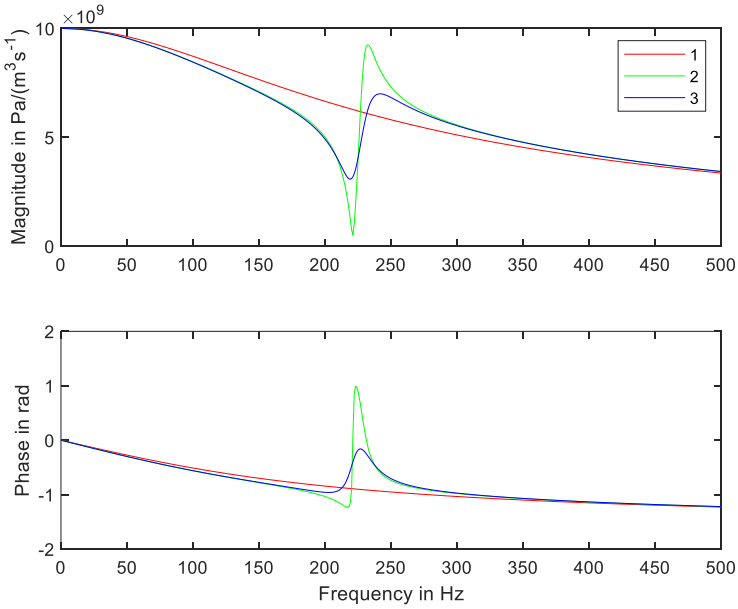


Figure 9: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, axial piston pump, pump inertia, capacity and resistance at the outlet, 1: without compensation, 2: light damping compensator, 3: high damping compensator

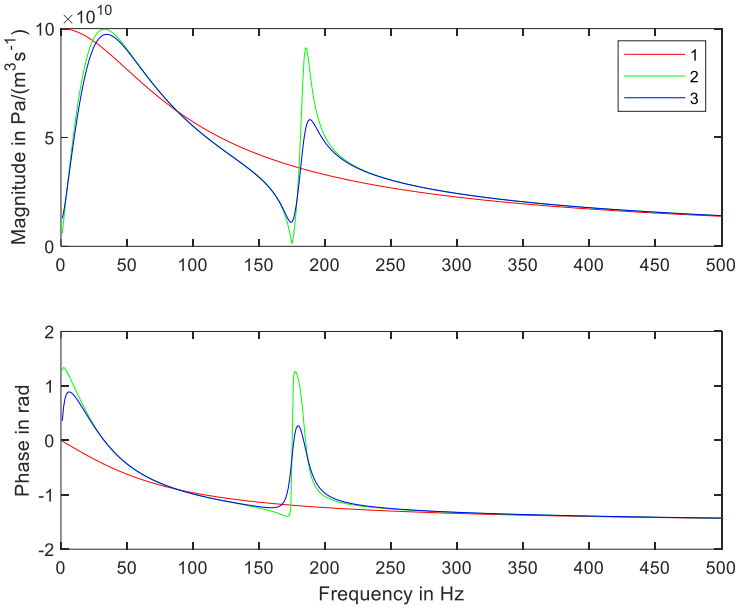


Figure 10: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, radial piston pump, motor inertia, pump inertia, capacity and resistance at the outlet, 1: without compensation, 2: light damping compensator, 3: high damping compensator

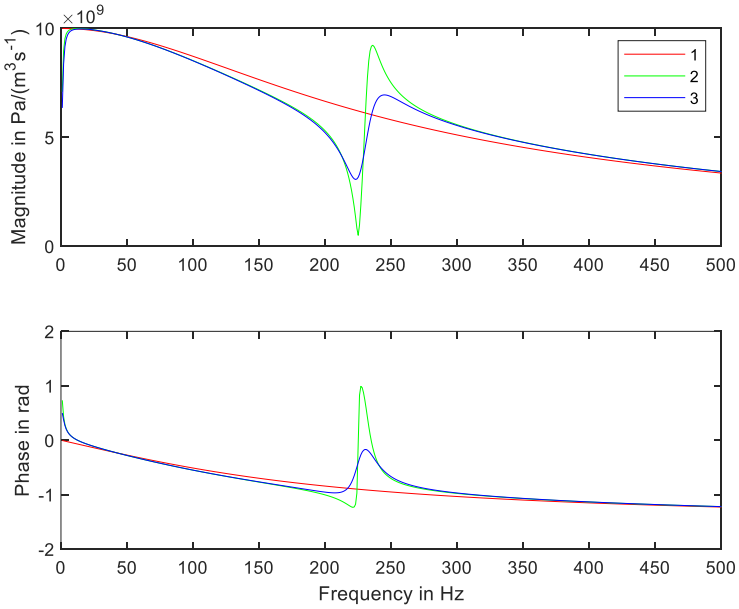


Figure 11: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, axial piston pump, motor inertia, pump inertia, capacity and resistance at the outlet, 1: without compensation, 2: light damping compensator, 3: high damping compensator

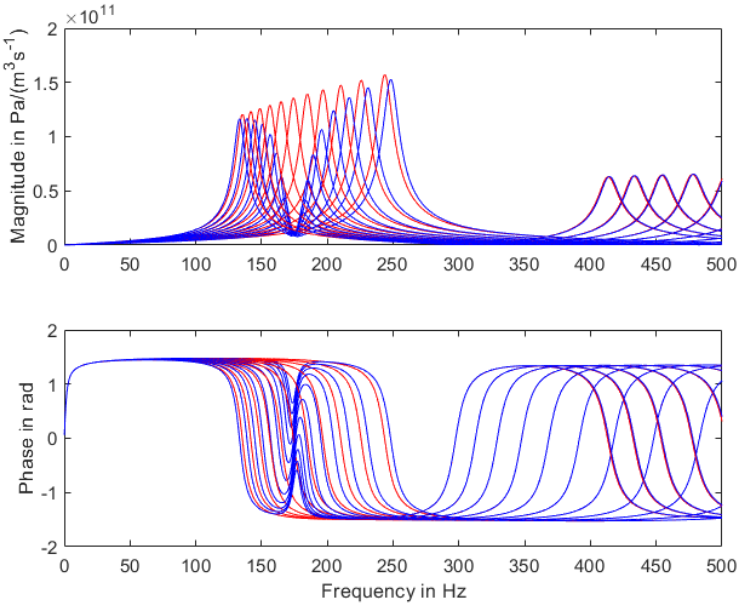


Figure 12: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, radial piston pump, motor inertia, pump inertia, capacity and pipeline at the outlet, red curves: different pipeline lengths without compensation, blue curves: different pipeline lengths with high damping compensator

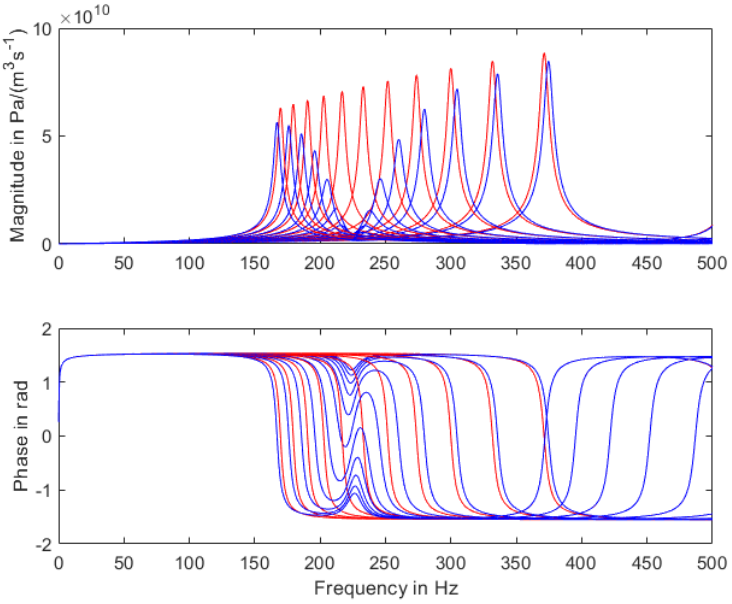


Figure 13: $P(\mathrm{i}\omega)/Q(\mathrm{i}\omega)$, axial piston pump, motor inertia, pump inertia, capacity and pipeline at the outlet, red curves: different pipeline lengths without compensation, blue curves: different pipeline lengths with high damping compensator

Figure 10 and **Figure 11** result from the model in **Figure 6**. Accounting for the motor inertia, another resonance emerges (33 Hz for the radial piston pump and 13 Hz for the axial piston pump). The radial piston pump antiresonance shifts up to 175 Hz. For the axial piston pump, the electric motor's mass moment of inertia is higher in relation, resulting in an antiresonance shift up to 225 Hz. In both cases, one can see that the entire pump drivetrain dynamics is essential for the antiresonance frequency, which now matches the pulsation frequency of the respective pump.

Figure 12 and **Figure 13** refer to the model in **Figure 7**. For radial piston pump outlet pipeline lengths from 1.1 m to 2.1 m in steps of 0.1 m, **Figure 12** shows frequency response functions with and without the high damping compensator. In **Figure 13** such frequency response functions are shown for axial piston pump outlet pipeline lengths from 0.7 m to 1.7 m in steps of 0.1 m. One can see that the compensation effect is largely independent of the outlet pipeline resonance frequency. The benefit of the compensator becomes most apparent where the intended antiresonance coincides with the original outlet pipeline resonance. At the cost of two side resonances, the intended antiresonance remains.

6. CONCLUSION AND OUTLOOK

For hydraulic pumps and motors, a novel pulsation compensation device has been suggested. With a suitable coupling, the drivetrain of the hydraulic machine forms a lightly damped torsional oscillator whose natural vibrations compensate flow rate excitations occurring in the displacement chamber. The device can be realized by merely defining the torsional properties of the coupling. Two possible coupling design concepts have been indicated. For a medium size radial piston pump and a large axial piston pump, frequency domain models show significant attenuation of pressure pulsation at a fixed excitation frequency and high robustness with respect to response characteristics of the hydraulic system at the pump outlet.

In a next step, laboratory tests shall be carried out with a compensator tailored to the radial piston pump. Coupling design shall be detailed to allow for sufficient torsional compliance despite strength requirements. Under excitation from the pump itself, pressure response shall be compared between setups with compensator and standard coupling, respectively. In both cases, different hydraulic networks will be installed at the pump outlet. As the flow rate excitation from the pump cannot be measured, similar experiments with external valve excitation shall be carried out to enable direct comparisons with frequency domain models. If experimental results come near the predicted behaviour, the novel compensator offers effective pulsation reduction at little cost for many displacement machine applications at constant rotational speed.

NOMENCLATURE

A	Pipeline cross-sectional area	m^2
C_H	Hydraulic capacity	$m^3\ Pa^{-1}$
c_T	Torsional damping constant	$N\ m\ s\ rad^{-1}$
E	Bulk modulus	Pa
f_P	Pulsation frequency	Hz
h	Gap height	m
I_M	Electric motor mass moment of inertia	$kg\ m^2$
I_P	Pump mass moment of inertia	$kg\ m^2$
k_T	Torsional stiffness	$N\ m\ rad^{-1}$
l	Length of pipeline	m
l_P	Length of pump rotor	m

n	Rotation frequency	Hz
p	Pressure at pump outlet	Pa
P	Laplace transformed pressure at pump outlet	Pa
q	Flow rate at pump outlet	$m^3 s^{-1}$
Q	Laplace transformed flow rate at pump outlet	$m^3 s^{-1}$
r	Radius of pipeline	m
R	Resistance	$Pa m^{-3} s$
r_a	Outer radius of pump rotor	m
r_i	Inner radius of pump rotor	m
T_M	Laplace transformed electric motor torque	N m
T_τ	Shear torque	N m
V_d	Dead volume at pump outlet	m^3
V_P	Displacement volume	m^3
x	Intermediate radius	m
z	Number of pistons	-
ν	Kinematic viscosity	$m^2 s^{-1}$
ρ	Fluid density	$kg m^{-3}$
τ	Shear stress	Pa
Φ_M	Laplace transformed rotation angle of electric motor	rad
$\dot{\varphi}_P$	Angular velocity of pump	$rad s^{-1}$
φ_P	Rotation angle of pump	rad
Φ_P	Laplace transformed rotation angle of pump	rad
ω	angular frequency	$rad s^{-1}$

REFERENCES

[1] Harrison AM, Edge KA (2000) Reduction of axial piston pump pressure ripple. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering 214(1):53-63

[2] Ingard U (1953) On the theory and design of acoustic resonators. Journal of the Acoustical Society of America 25(6):1037-1061

[3] Mikota J (2002) Contributions to the development of compact and tuneable vibration compensatros for hydraulic systems. Dissertation, Johannes Kepler University Linz

[4] Mikota J, Reiter H (2003) Development of a compact and tuneable vibration compensator for hydraulic systems. International Journal of Fluid Power 4(1):17-31

[5] Mikota G (2022) Vorrichtung zur Reduktion von Druckpulsationen in einem Hydrauliksystem. Austrian Patent Application A 50695/2022, 9 Sept 2022

[6] D'Souza AF, Oldenburger R (1964) Dynamic Response of Fluid Lines. Transactions of the ASME – Journal of Basic Engineering 86:589-598

Chapter 9

New and Special Applications-I

DEVELOPMENT OF RECIPROCATING AIR EXPANDER FOR μ - CAES TECHNOLOGY

Jan Markowski*, Dominik Gryboś[1], Jacek Leszczyński[1]

AGH University of Science and Technology, Faculty of Energy and Fuels, Department of Department of Thermal and Fluid Flow Machines, Av. Mickiewicza 30, 30-059 Krakow, Poland

* Corresponding author: Tel.: +48 506152247 ; E-mail address: jmarko@agh.edu.pl

ABSTRACT

As renewable energy sources (RES), such as solar and wind, continue to grow in the energy mix, it becomes crucial to address their inflexibility and daily demand-production gap. Energy storage is vital, allowing storage during low demand and supplying power when needed. The authors investigated compressed air energy storage (CAES) as a mechanical solution, focusing on micro-CAES (μ-CAES) for smaller industries or housing estates. This research aimed at optimizing μ-CAES efficiency through modifications to the air expander's construction (piston parameters) and process parameters (air pressure, temperature). Initial results with a 3-cylinder expander showed an average power of 2.2 kW and 0.45 efficiency. Further enhancements, such as compressed air supply control and waste heat utilization during expansion expects to affect on increase of efficiency to 0.8.

Keywords: Compressed air, Air expander, Pneumatic system, Energy storage, Integrated design

1. INTRODUCTION

With the continuous increase of renewable energy sources in the energy mix, it becomes necessary to address the issue associated with their intermittency (solar, wind) and the mismatch between daily energy demand and production [1, 2]. The way to overcome this challenge is energy storage. Storage systems can be divided into mechanical, chemical, electrochemical, and electrical systems [3].One of the mechanical energy storage systems is Compressed Air Energy Storage (CAES), which is dynamically evolving and highly promising [4]. Micro-CAES is an energy installation with a rated power of up to 1 MW, typically fluctuating in the range of tens of kW [5]. Its role involves effective storage and delivery of electrical energy, contributing to the stability and efficiency of the entire energy infrastructure. This technology also finds its use in independent energy systems for individual homes [6]. CAES can also directly collaborate with renewable energy sources such as photovoltaics to optimize the energy efficiency of a building [7]. As there are issues with efficiency or start-up time of those systems a lot of hybrid combinations can be seen in literature. Authors [8] connects CAES with super capacitors creating CAES-SC to achieve efficiency of 57,9 %. Another solutions requires additional heat storage to optimize thermodynamic processes during compression and expansion in systems operation [9].

The design process of entire Micro-CAES system can be divided into specific stages that include analyses of individual subsystems. Generally, we can categorize the design into Conceptual, Pneumatic, Mechanical and Electrical. Each of these steps is continuously reanalysed throughout the entire design process and reconsidered after each modification introduced in another stage. **Figure 1** show integrated design patch scheme.

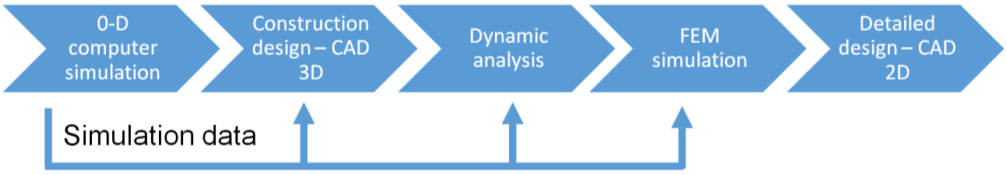


Figure 1: Integrated design patch scheme.

Conceptual Design involves analysing various aspects, starting from determining the overall desired parameters of the system that will impact its functionality. At this stage, system components are being preliminarily selected. The choice of optimal solutions, especially the type of expander, plays a significant role in achieving the highest efficiency [5]. The focus then shifts to Pneumatic System Design, where analysis is conducted on elements such as the air tank, piston, and valves, as well as on process parameters like supply pressure, storage pressure, and air control system, to meet the goals set in the conceptual phase. The subsequent step is Mechanical Construction Design, where elements related to mechanical drive and transmission are analysed. This includes evaluating the strength of each component to ensure the durability, reliability, and efficient operation of the device. Those stages allow for the creation of a coherent CAD design of the mechanical drive of the piston expander, shown in **Figure 2**

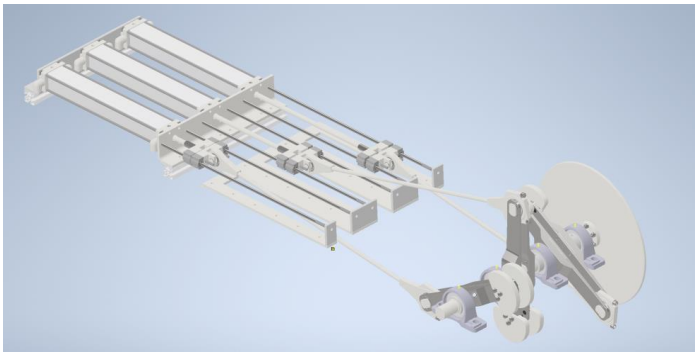


Figure 2: Mechanical construction of multi-cylinder air expander

As a result, the system design process involves a gradual transition from general parameters, through construction specifications, to a detailed consideration of individual elements, ensuring consistency and efficiency throughout the design process.

2. MATHEMATICAL MODEL

Our μ-CAES solution involves the utilization of a compressor, a storage tank, and a reciprocating engine. Actuators driven by compressed air initiate rotary motion of the shaft through a crankshaft. The shaft is connected to a high-speed generator via a mechanical gearbox, ensuring the required rotational output speed. Considering micro-scale installations, potential issues with excessively long start-up times or high resistance during intermittent operation are avoided. Therefore, the decision was made not to install a gas turbine, which is not resistant to varying mass flow rates of the supplying air and would also significantly increase installation costs. Instead, a reciprocating engine was installed, capable of operating with variable loads, pressures, and mass flow rates during a single tank discharge. **Figure 3** below illustrates the micro-CAES scheme.

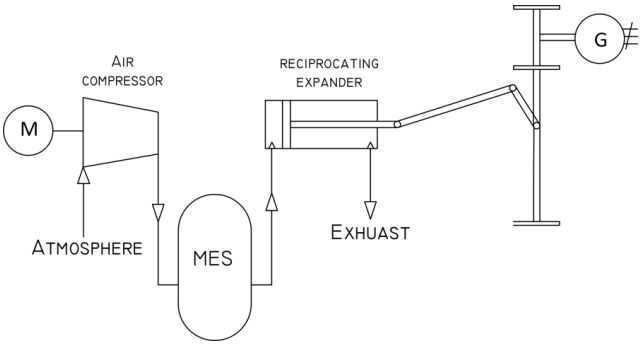


Figure 3: Mechanical construction of micro-CAES

The mathematical model is highly important in the real system design process as it provides force necessary data for conducting dynamic strength analysis of the system components. Model was based on differential equations described by the authors [10].

$$\begin{cases} J\frac{d^2\alpha}{dt} = \sum T \\ \omega = \frac{d\alpha}{dt} \\ L\frac{dI}{dt} = K\omega - R_G I \end{cases} \tag{1}$$

Driving torque:

$$T_D = r \cdot F_1 \sin(\alpha)\left(\left(\frac{r}{l}\right)^2 \sin^2(\alpha) + \frac{r}{l}\cos(\alpha)\sqrt{1-\left(\frac{r}{l}\right)^2 \sin^2(\alpha)} - 1\right) \tag{2}$$

The braking torque arising from the friction in the bearings:

$$T_F = f_1 \cdot Pz \cdot d_b + f_0(\nu \cdot \omega)^{\frac{2}{3}} \cdot d_b \cdot 10^{-7} \tag{3}$$

Generator breaking torque:

$$T_G = \frac{3}{2} \cdot I \cdot \Psi \cdot N_{pb} \tag{4}$$

The angular speed of the shaft:

$$\omega(t) = \frac{d\alpha}{dt} \tag{5}$$

The mass flow is calculated by formulas presented by Leszczyński and Gryboś [11]:

$$\dot{m} = \begin{cases} A\sqrt{\frac{2\kappa}{\kappa-1} \cdot \rho p_1 \left(\left(\frac{p_2}{p_1}\right)^{\frac{2}{\kappa}} - \left(\frac{p_2}{p_1}\right)^{\frac{\kappa+1}{\kappa}}\right)} \; for \;\; \frac{p_2}{p_1} > \beta \\ 0 \qquad for \quad p_1 < p_2 \\ A\sqrt{\kappa\rho p_1 \left(\frac{2}{\kappa+1}\right)^{\frac{\kappa+1}{\kappa-1}}} \qquad for \;\; \frac{p_2}{p_1} \leq \beta \end{cases} \tag{6}$$

As noted in [11] the value of β equals:

$$\beta = \zeta(\left(\frac{2}{\kappa+1}\right)^{\frac{\kappa}{\kappa-1}} \quad (7)$$

Where ζ takes values from 0 to 1 depending cross section edges smoothness.

The power output is calculated as follows:

$$P_{out} = I^2 R_G \quad (5)$$

Estimated power of compressed air [12]:

$$P_{ca} = p_a \dot{V}_a \ln \frac{p_a}{p_b} \quad (6)$$

Round trip efficiency compares the output power of the expander and power of used compressed air :

$$\eta = \frac{P_{out}}{P_{air}} \quad (7)$$

To verify the accuracy of the mathematical model, experimental data from a single-cylinder expander were compared in previous study by authors [10]. Obtained data from the simulation was cross-referenced and compared with real-world observations or measurements. This step is crucial to confirm the reliability of the simulation data used for designing a multi-cylinder prototype. The validation process substantiated and affirmed the reliability and accuracy of the simulation results. The consistency between the simulated outcomes and the actual performance provided a robust validation of the mathematical model.

3. RESULTS

To select components such as a generator or mechanical gear, computer simulations were performed to illustrate the unit's characteristics. Proper selection of parameter values will allow the expander to operate in the highest efficiency ranges. **Table 1** below shows desired parameters of three piston air expander.

Table 1: Desired parameters of three piston air expander

Parameter	Value
Mechanical power of Expander	1-2 kW
Energy storage	0,9 kWh (2x1m^3)
Roundtrip efficiency	Over 50%
Discharge time	20-30 min @ 7 bar
Storage pressure	11 bar
Expansion pressure	6-11 bar

In the first step, in order to select the appropriate control strategy characteristics of expander with implemented PWM (Pulse Width Modulation) function were carried out. PWM control function regulate the injection of compressed air into the actuator chamber. Depending on the specified condition, the supply of compressed air is somewhat restricted, resulting in reduced consumption with a slight decrease in output power. Due to the expansion of air in the chamber, even when the supply is cut off, its velocity does not change significantly, and it is capable of performing work. **Figure 4 and 5** shows characteristics of expander with implemented PWM function and three actuators with 50mm of piston diameter.

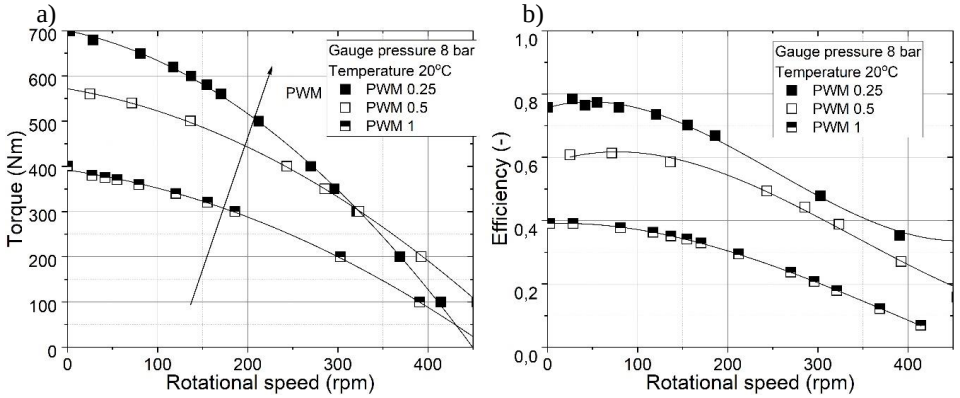


Figure 4: Characteristics of multi-piston expander with PWM function: a) torque and b) efficiency.

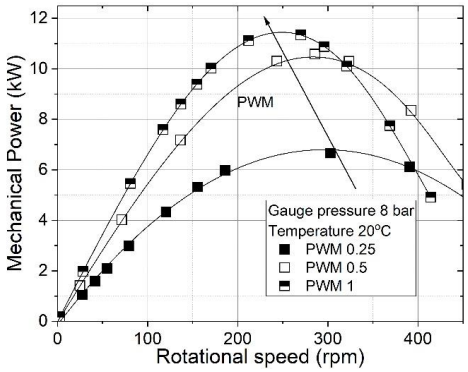


Figure 5: Characteristics of multi-piston expander with PWM function: mechanical power.

For each subsequent lower PWM coefficient, the mechanical power decreases. With gauge pressure of 8 bar and PWM 0.25 efficiency of 0.8 was obtained with generated mechanical power of 1.2 kW. In further stages it was decided to analyze system with multi-piston expander with implemented PWM 0.25 function to seek the highest possible efficiency. Characteristics shown on **Figure 6** were determined for three different supply gauge pressure values: 6, 8, and 10 bars , in order to select the appropriate supply pressure.

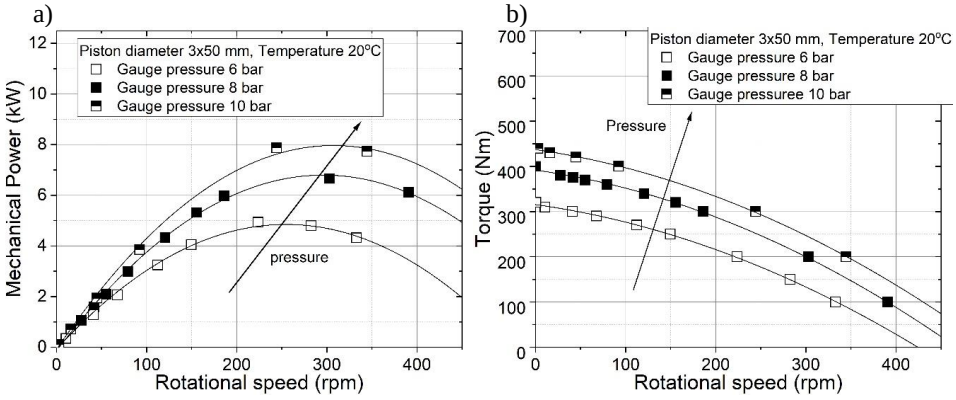


Figure 6: Characteristics of multi-piston expander: a) mechanical power and b) torque.

The maximum mechanical power obtained increases with the supply pressure, reaching from 5 kW for gauge pressure of 6 bar to 8kW for gauge pressure of 8 bar. Since the anticipated highest efficiency is expected to occur at low rotational speeds, it has been decided that further considerations will focus on a system powered by a gauge pressure of 8 bars in order to meet the intended output power. The characteristics (**Figure 7 and Figure 8**) were conducted for a supply pressure of 8 bar for three different sizes of actuators.

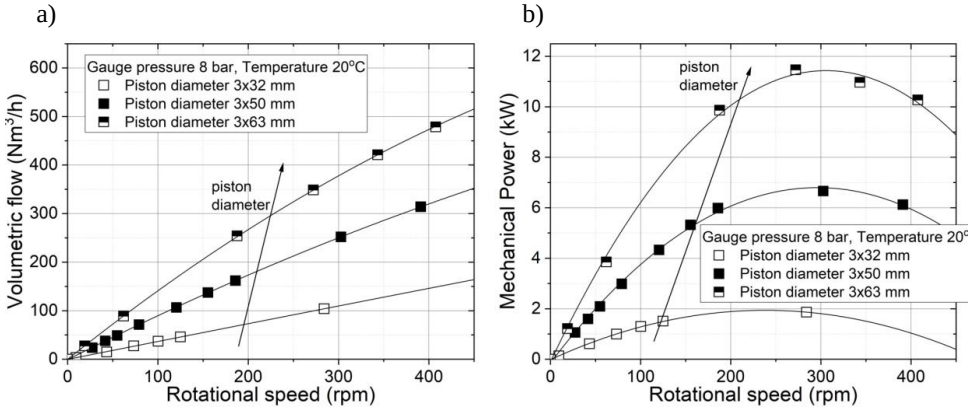


Figure 7: Characteristics of multi-piston expander: a) volumetric flow and b) mechanical power.

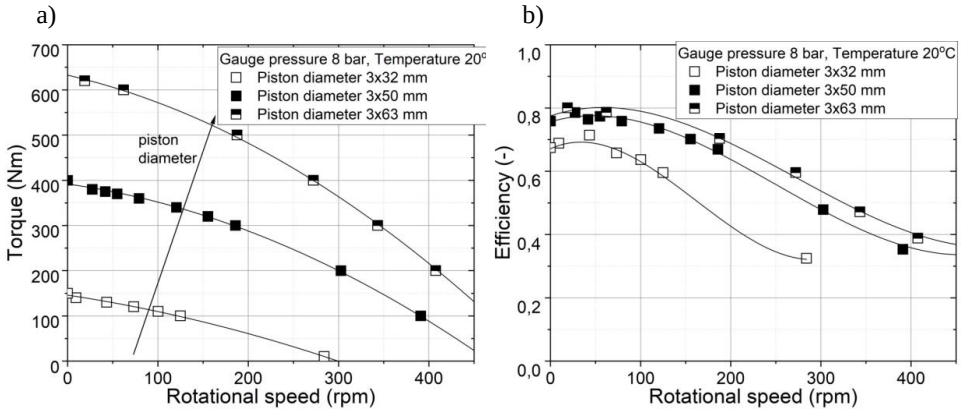


Figure 8: Characteristics of multi-piston expander: a) torque and b) efficiency.

The expander's characteristics are intricately linked to its ability to function optimally at speeds below 60 revolutions per minute. This feature underscores the importance of precision and careful consideration during the design process. In order to achieve the highest efficiency of the expander, it is critical to maintain a rotational speed around 50 revolutions per minute. In the subsequent stages of designing the entire system, the construction planning will be centred on ensuring optimal operation within this particular speed range. It has been emphasized that the power generated by the device is satisfactory, serving as a positive indicator for the overall efficiency and performance of the entire project. **Figure 9** below shows the linear velocity of the piston for three different sizes of actuators powered by a gauge pressure of 8 bars.

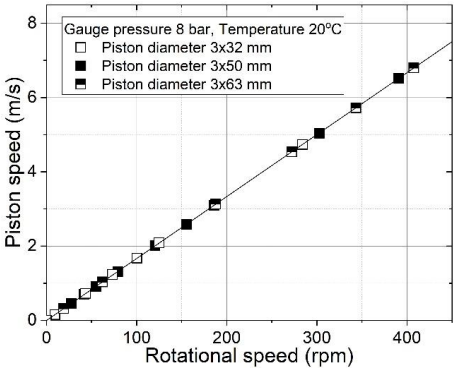


Figure 9: Linear speed of piston for three different actuators sizes.

At a constant supply pressure, the piston velocities remain unchanged regardless of the actuator size increasing with rotational speed. Irrespective of whether the actuator is larger or smaller, it receives the same piston velocity at the same level of supply pressure. The last considered aspect was examining the impact of temperature on the performance of the expander. Characteristics shown on **Figure 10** below were obtained for three different values of the supplied air temperature.

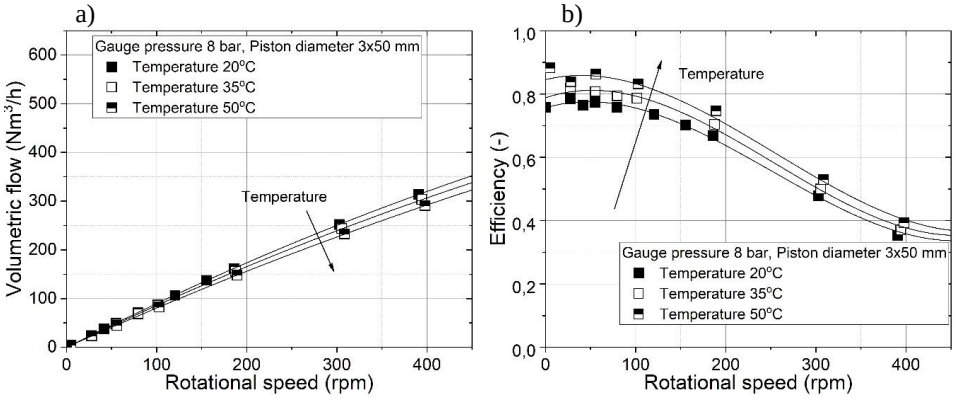


Figure 10: The influence of temperature on the characteristics of the air expander: a)volumetric flow, b) efficiency.

As the temperature increases, the changing air density leads to reduced volumetric consumption, resulting in higher achieved overall efficiency. Heating the air to a temperature of 50 degrees allows to achieve efficiency of up to 0.9 The overarching goal of this project is to optimize the design and functionality of the pneumatic system to seamlessly match the specified power output, ensuring both efficiency and reliability. Implementing proper control of the entire system, using Pulse Width Modulation (PWM) function and introducing additional air heating aim to further enhance of systems round trip efficiency. It is important to emphasize that the pursuit of higher efficiency is not a singular effort but rather a combination of various strategies. Only the integration of all these efficiency-improving methods and an analysis of their optimal collaboration will allow for the best performance outcomes.

4. CONCLUSIONS

The paper introduces a mathematical model essential for establishing design and process parameters in micro-CAES piston expanders. Computer-simulated data from this model serve as inputs for

strength calculations, FEM analysis, and the construction of the mechanical drive. A specific case study is showcased, focusing on the design of a 1-2 kW piston expander operating at a gauge pressure of 6-10 bar. The anticipated mechanical efficiency of the air expander is targeted at 0.7. Simulations showed that compressed air supply control and waste heat utilization during expansion can affect on increase of efficiency to 0.8. The synergy between process parameters, PWM control system, and supplementary heating plays a key role in elevating the overall effectiveness of the expander within the CAES system. Only through the proper integration of these efficiency-enhancing methods and a thorough analysis of their optimal collaboration can the system achieve its highest potential.

5. ACKNOWLEDGMENTS

The work was carried out as part of Research project supported by program „Excellence initiative – research university" for the AGH University of Krakow, research subvention under contract no. 16.16.210.476 supported by the Polish Ministry of Science and Higher Education and by the Polish National Agency for Academic Exchange under the Strategic Partnerships Program, Project No. BPI/PST/2021/1/00023.

NOMENCLATURE

A	Cross-section	m^2
I	Current	A
J	Moment of inertia	Nm
K	Generator constant	Wb
L	Inductance	H
P_z	Equivalent load	N
P_{out}	Power output	W
P_{ca}	Compressed air power	W
R_G	Resistance	Ohm
T	Torque	Nm
$\dot{V}$	Volumetric flow	$\frac{m^3}{min}$
d_b	Bearing pitch diameter	m
f_0, f_1	Bearing factors	-
l	Crank arm length	m
$\dot{m}$	Mass flow	kg/s
p_1, p_2	Pressures in actuator chambers	
p_a	Storage pressure	Pa
p_b	Pressure after the expansion process	Pa
r	Crank length	Pa
t	Time	m
α	Angular position of the crank	s
β	Critical pressure ratio	rad
ω	Angular speed	$\frac{rad}{s}$
ν	Kinematic viscosity of oil	mm^2/s
κ	Adiabatic exponent for air	-
ρ	Density	kg/m^3
ζ	Numerical factor	-
CAES	Compressed Air Energy Storage	
PWM	Pulse Width Modulation	

REFERENCES

[1] Mahmoud M, Ramadan M, Olabi A-G, Pullen K, Naher S (2020) A review of mechanical energy storage systems combined with wind and solar applications. Energy Convers Manag 210:1–14

[2] Yu Q, Wang Q, Tan X, Fang G, Meng J (2019) A review of compressed-air energy storage. J Renew Sustain Energy 11:1–12

[3] Mitali J, Dhinakaran S, Mohamad AA (2022) Energy storage systems: a review. Energy Storage and Saving 1:166–216

[4] Barbour ER, Pottie DL, Eames P (2021) Why is adiabatic compressed air energy storage yet to become a viable energy storage option? iScience 24:1–26

[5] He W, Wang J (2018) Optimal selection of air expansion machine in Compressed Air Energy Storage: A review. Renewable and Sustainable Energy Reviews 87:77–95

[6] De Lieto Vollaro R, Faga F, Tallini A, Cedola L, Vallati A (2015) Energy and thermodynamical study of a small innovative compressed air energy storage system (micro-CAES). Energy Procedia 82:645–651

[7] Simpore S, Garde F, David M, Marc O, Castaing-Lasvignottes J (2019) Sensitivity analysis and optimization of a compressed air energy storage (CAES) system powered by a photovoltaic plant to supply a building. Procedia Manuf 35:137–142

[8] Zhewu C, Zheming T, Shuiguang T, Qinguo Z, Liang L (2023) CAES-SC hybrid energy storage: Dynamic characteristics and control via discharge process. J Energy Storage 72:108561

[9] Wang S, Zhang X, Yang L, Zhou Y, Wang J (2016) Experimental study of compressed air energy storage system with thermal energy storage. Energy 103:182–191

[10] Leszczyński JS, Gryboś D, Markowski J (2023) Analysis of optimal expansion dynamics in a reciprocating drive for a micro-CAES production system. Appl Energy 350:121742

[11] Leszczyński JS, Gryboś D (2019) Compensation for the complexity and over-scaling in industrial pneumatic systems by the accumulation and reuse of exhaust air. Appl Energy 239:1130–1141

[12] Cai M, Kagawa T, Kawashima K (2002) Energy Conversion Mechanics and Power Evaluation of Compressible Fluid in Pneumatic Actuator Systems. 37th Intersociety Energy Conversion Engineering Conference (IECEC) 471–474

COMPARATIVE ANALYSIS OF PERFORMANCES OF NON-METAL PRESSURIZED RESERVOIRS WITH VARIABLE VOLUME

Jing Yao[1,3], Dingyu Wang[1,3]*, Dong Liang[1,3], Jinlu Hao[1,3], Pei Wang[2,3]

[1] *State Key Laboratory of Crane Technology, Yanshan University, West of Hebei Avenue 438, 066004 Qinhuangdao, China*
[2] *Hebei Innovation Center for Equipment Lightweight Design and Manufacturing, Yanshan University, West of Hebei Avenue 438, 066004 Qinhuangdao, China*
[3] *School of Mechanical Engineering, Yanshan University, West of Hebei Avenue 438, 066004 Qinhuangdao, China*

* Corresponding author: Tel.: +86 15931165526; E-mail address: dingyu_wang@stumail.ysu.edu.cn

ABSTRACT

The lighter weight of hydraulic reservoirs can improve the power-to-weight ratio of hydraulic equipment and reduce power consumption, achieving energy conservation and emission reduction. In this study, three types of non-metal pressurized reservoirs with variable volume (VVPR) are designed and manufactured. Among them, a single-layer configuration is the basic functional prototype. Firstly, the structure of VVPR is designed to meet the differential volume compensation by its variable volume in an asymmetric hydraulic cylinder system. Then, series and parallel configurations are proposed based on the basic prototype. Finally, a test bench is built to test the step and sine response performances of the VVPR. The response time of the reservoir is not affected by the step amplitude of the hydraulic cylinder changes, with an average response time of 0.028 s. In addition, the lag time of the single-layer reservoir is the smallest in the sine test, which is 0.11 s. Through comparative analysis, the configuration of the series reservoir is more compact and can get more pressure. Parallel reservoir can get greater variable volume with the same displacement change. It provides an alternative solution for lightweight hydraulic systems in the future.

Keywords: Hydraulic Reservoir, Lightweight, Pressurized Reservoir, Variable Volume

1. INTRODUCTION

Hydraulic transmission is widely used in high-level mobile equipment such as aerospace, robotics, and engineering machinery due to its high power-to-weight ratio, fast response, and stepless speed regulation. Hydraulic systems often rely on reservoirs that are considerably larger and weigh in mobile hydraulic machinery equipment. The lightweight of hydraulic components and systems can improve the maneuverability, navigating ability, and carrying capacity of mobile equipment. The hydraulic oil reservoir is the most promising component [1], and the hydraulic oil reservoir is lightened mainly through two ways: reducing its volume and using non-metallic materials for its shell [2]. With increasing demands for reducing emissions and energy consumption, the design of hydraulic reservoirs needs to be optimized for weight and volume [3]. Traditional open reservoirs are not only bulky, but their internal oil is also prone to contamination, posing a threat to the hydraulic system. To solve the problem, ARGO-HYTOS has introduced a solution known as the “Hybrid Integrated Tank” [4]. This design leverages polyamide material to achieve exceptional mechanical strength and thermal stability, which intelligently combines the two manufacturing technologies of rotational molding and injection molding. In contrast, closed reservoirs provide smaller volume and lighter weight, and their internal oil is not susceptible to contamination. Currently, the main directions for lightweight closed

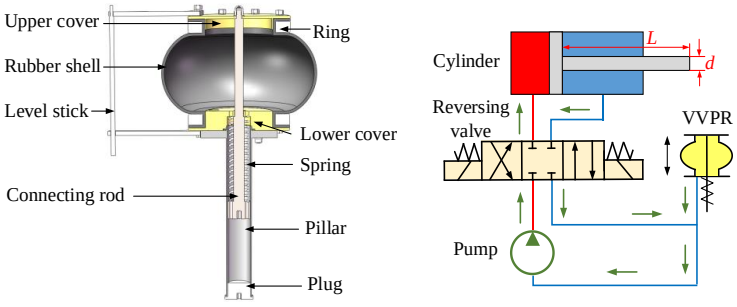
hydraulic reservoirs are pressurized and variable capacity. For example, a spring-pressurized reservoir with a metal cylinder as its shell. When fluid flows into or out of the container, the spring contracts or expands in response to volume changes. At the same time, the spring generates fluid pressure by its force [5]. However, the metal construction of the reservoir results in a large volume and mass. Presently, there are emerging innovations in closed reservoirs, with a focus on novel structural designs and advanced materials for smaller volumes and lower weight [6]. To further reduce volume, the concept of variable volume reservoirs is proposed. Variable volume reservoirs are made of non-metallic materials, significantly reducing their volume and weight compared to pressurized reservoirs, providing a new method to solve the lightweight problem of reservoirs. Smart Reservoir Inc. [7] has produced a variable volume reservoir characterized by a housing crafted from a dual polymer material that exhibits remarkable chemical inertness towards hydraulic fluids. This innovation has advantages such as reduced weight, compact dimensions, and a linear output profile. It has already found application across diverse industries.

In this study, three types of non-metal pressurized reservoirs with variable volume (VVPR) are designed and manufactured. The single-layer configuration is the basic functional prototype, which is designed to meet the differential volume compensation by its variable volume in an asymmetric hydraulic cylinder system. In addition, series and parallel configurations are proposed based on the basic prototype, adapting to the compensation of more oil volume in the system.

2. CONFIGURATION DESIGN

2.1. Principle of the VVPR

The structure of VVPR is designed to meet the differential volume compensation in an asymmetric hydraulic cylinder system and increase the suction pressure of the hydraulic pump. The single-layer configuration is the basic functional prototype of the VVPR and mainly includes a rubber shell, spring, upper and lower covers, connecting rod, rings, and other components, which is shown in **Figure 1(a)**. Among them, the elastic shell made of rubber is used to store hydraulic oil. The upper cover and the connecting rod are fixedly connected and move together. The upper and lower covers are in contact and the spring is in free length during initial installation. The spring is located in the pillar and is always compressed in the working cycle of the reservoir. The compression of the spring is the same as the upward displacement of the upper cover during the working process of the VVPR. In addition, the function of absorbing and discharging oil and increasing the suction pressure of the pump is achieved through the interaction of spring force and internal pressure acting on the upper cover. The VVPR is airless, sealed, and slightly pressurized, with a small volume and lightweight.



(a) Structure of the single-layer reservoir (b) Application in the system

Figure 1: Working principle diagram of the VVPR

In the system application, the main function of the reservoir is to compensate for the volume

difference of the hydraulic cylinder, which is generated by the asymmetric structure of the hydraulic cylinder, shown in **Figure 1(b)**. The valve control hydraulic system comprises a hydraulic pump, reversing valve, single-rod hydraulic cylinder, and VVPR, forming a closed loop. In the operation of the system, the hydraulic cylinder is extended and retracted by switching the reversing valve, with discharging and absorbing oil of the VVPR at the same time. In the initial application stage, air bleeding will be achieved through the exhaust valve of the upper cover. For example, when the hydraulic cylinder is extended, the oil from the outlet of the pump to the rodless chamber of the hydraulic cylinder through the reversing valve. The oil in the rod chamber enters the suction port of the pump. However, due to the different volumes of the rod and rodless chambers, the oil in the VVPR is replenished into the inlet of the pump. The hydraulic cylinder retraction process is the opposite.

In the case of not considering the system leakage and oil compression, the volume difference between all hydraulic cylinders with rod and rodless chambers in the hydraulic system is shown in Equation (1).

$$\Delta V = \sum_{i=1}^{n} \frac{\pi {d_i}^2}{4} L \tag{1}$$

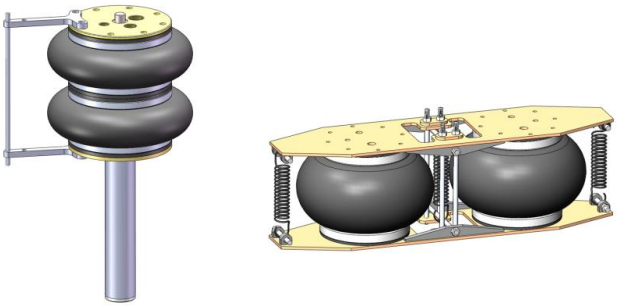
Where ΔV is the volume difference, d is the cylinder rod diameter, and L is the cylinder stroke.

Therefore, the volume of the VVPR should be designed to meet the requirements of the volume difference required by the hydraulic system.

2.2. Structural iteration

The volume and pressure of VVPR are two main performance parameters in the design stage. To adapt to the compensation of more oil volume in the system, series and parallel configurations are proposed based on the basic prototype. The series configuration is shown in **Figure 2(a)**, and it is different from a single-layer reservoir in that it comprises two single-layer rubber shells, which are connected by embedded threads in the rings. Furthermore, it can also be connected in series with multiple shells, such as three or four.

Series configuration of VVPR provides higher pressure than a single-layer reservoir, under the same length of pillar. Because the pressure of the reservoir is generated by the deformation of the spring, the higher the total shell of VVPR, the bigger the pressure provided. However, this configuration increases the vertical height of VVPR. In addition, under the same volume, the maximum outer diameter of the series configuration is smaller than that of the single-layer reservoir. The maximum outer diameter refers to the maximum radial distance of the rubber shell when the VVPR is at its maximum compression position, which reflects the compactness of the horizontal spatial layout.



(a) Series reservoir (b) Parallel reservoir

Figure 2: The configuration iteration of VVPR

The parallel configuration of VVPR is shown in **Figure 2(b)**, in which the two single-layer rubber shells are arranged in parallel, and the springs are connected between the upper and lower covers. It can also be arranged three or four rubber shells in a horizontal circumferential array. Compared with the other two configurations, it can discharge or absorb more oil under the same height change of the VVPR.

However, compared with the series configuration of VVPR, the parallel configuration reduces the vertical height but increases the size of the horizontal space occupied by the VVPR in the working state. The spring used in the parallel configuration is the tension spring rather than a compression spring. Moreover, to obtain the same pressure, more springs need to be connected in parallel to improve spring stiffness, due to the increase in the pressed area of the upper cover.

Advantages of the VVPR:

- Adapt to the change of oil volume in the system.
- Lightweight, small volume and low cost.
- Compared with the bladder accumulator, temperature variation does not affect the pressure variation by use of a mechanical spring. In addition, no gas is required.

When the VVPR is used in a hydraulic system, the configuration design can be reasonably selected according to the spatial layout of the system.

3. MODEL DEVELOPMENT

Take the single-layer configuration of VVPR as an example. The variable volume function of the reservoir is realized by the compression deformation of the rubber shell of VVPR, shown in **Figure 3**. Besides, the spring force generated by spring deformation is transmitted to the upper cover through the connecting rod to realize oil pressurization of the VVPR. The relationship between volume, pressure, and structural parameters can be obtained through geometric and force analysis.

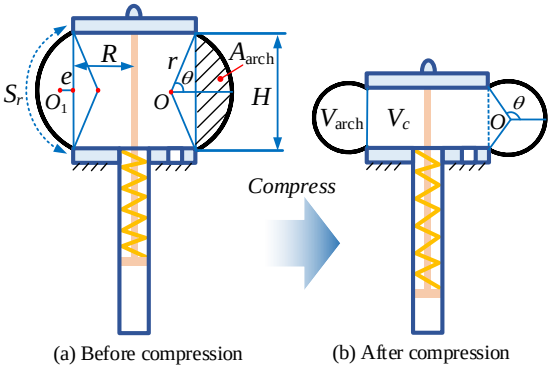


Figure 3: Compression deformation of the rubber shell of VVPR

3.1. Geometrical analysis

In this section, the relationship between the volume of the reservoir and the height of the rubber shell will be studied. According to the structure of the reservoir, the rubber shell is a body of revolution. The longitudinal cross-section of the rubber shell and structural parameters are illustrated in **Figure 3(a)**.

The following two hypotheses are proposed to simplify the analytical process of the geometric model:

- The length of the curved longitudinal cross-section is constant while the height of the reservoir changes.

- The curve is always arc-shaped in the process of reservoir work.

The height of the reservoir is calculated as shown in Equation (2).

$$H = 2r\sin\theta \tag{2}$$

where r is the radius of the arc, θ is the half-central angle of the arc.

The relationship between arc length and half-central angle is as follows.

$$S_r = 2r\theta \tag{3}$$

Combining Equation (2) and Equation (3), the expression of reservoir height can be calculated as:

$$H = \frac{S_r \sin\theta}{\theta} \tag{4}$$

According to **Figure 3(b)**, the structural volume of the single-layer reservoir is given by:

$$V_1 = V_c + V_{\text{arch}} \tag{5}$$

where V_c is the volume of the cylinder between the upper cover and lower cover, V_{arch} is toroidal volume which is formed by an arch rotating around the central axis. V_c is calculated as:

$$V_c = \pi R^2 H \tag{6}$$

where R is the radius of the upper and lower cover. V_{arch} is calculated as:

$$V_{\text{arch}} = 2\pi(R + e)A_{\text{arch}} \tag{7}$$

where A_{arch} is the arch area and e is the distance from the centroid O_1 of the arch to the arc chord. According to the geometrical relationship, A_{arch} and e can be derived as follows [8].

$$A_{\text{arch}} = \frac{r^2(2\theta - \sin 2\theta)}{2} \tag{8}$$

And

$$e = r\left[\frac{4\sin^3\theta}{3(2\theta - \sin 2\theta)} - \cos\theta\right] \tag{9}$$

Therefore, the structural volume V_1 of the single-layer reservoir can be written as:

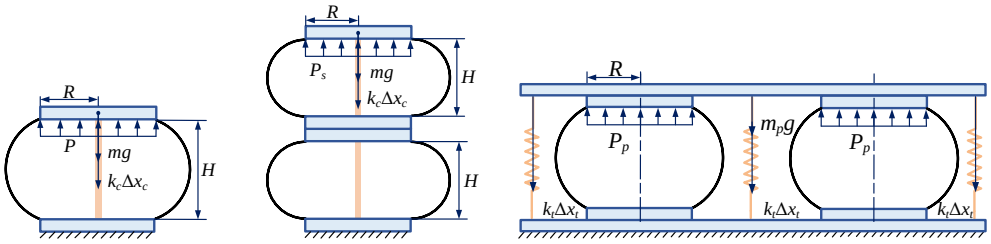
$$V_1 = \pi R^2 H + \frac{\pi {S_r}^2}{4\theta^2}(2\theta - \sin 2\theta)\left\{R + \frac{S_r}{2\theta}\left[\frac{4\sin^3\theta}{3(2\theta - \sin 2\theta)} - \cos\theta\right]\right\} \tag{10}$$

The volume expressions of the combined reservoirs can also be derived by using Equation (10). Assuming that the shape of the two rubber shells has the same change when the height of VVPR changes. The volume of both series and parallel reservoirs is two times that of the single-layer reservoir. During the volume changes, the maximum height of the parallel reservoir is the same as the height of the single-layer reservoir, while the maximum height of the series reservoir is twice the height of the single-layer reservoir.

3.2. Force analysis

During the operation of absorbing oil of the VVPR, the volume change of oil drives the upper cover of the reservoir upward, which causes the spring to compress, and the spring force acts on the oil

through the upper cover to achieve the purpose of pressurization. Ignore the slight influence of the elastic rubber shell on the upper cover. In the state of force balance, the force analysis of the upper cover of the single-layer reservoir is shown in **Figure 4(a)**.



(a) single-layer reservoir (b) series reservoir (c) parallel reservoir

Figure 4: Force analysis of the VVPR

The upper cover stops moving when the maximum volume of oil is reached, at which point the maximum pressure of oil can be obtained from the reservoir.

$$P = \frac{mg + k_c\Delta x_c}{\pi R^2} \tag{11}$$

where m is the mass of the single upper cover, P is the pressure of the single-layer reservoir, k_c is the stiffness of the compression spring, Δx_c is the compression of the spring, $\Delta x_c=H$.

It is clear to see that the maximum output pressure of the reservoir is determined by the height of the reservoir and the area of the upper cover.

Similarly, the maximum output pressure of the corresponding configuration can be obtained according to the force analysis of the series reservoir and the parallel reservoir. The force analysis of the series reservoir is shown in **Figure 4(b)**.

According to the force analysis, when the displacement of the upper cover of the series reservoir is maximum, where $\Delta x_c=2H$. The pressure of a series reservoir can be obtained.

$$P_s = \frac{mg + 2k_cH}{\pi R^2} \tag{12}$$

The force analysis of the parallel reservoir is shown in **Figure 4(c)**. According to the force analysis, the maximum output pressure of the parallel reservoir is as follows.

$$P_p = \frac{m_pg + nk_t\Delta x_t}{2\pi R^2} \tag{13}$$

where P_p is the output pressure of the parallel reservoir, k_t is the stiffness of the tension spring, n is the number of the tension springs, m_p is the mass of the upper cover of the parallel reservoir, Δx_t is the stretching length of the tension spring, $\Delta x_t=H-l_0$, l_0 is the free length of tension spring.

By comparison, among the three configurations of VVPR, the series reservoir can provide the maximum output pressure.

3.3. Comparison of configurations

By comparing and analyzing the characteristics of the three configurations of VVPR, the results are as follows:

- Single-layer configuration of VVPR

As a basic configuration, the VVPR is convenient to process and install, but when the demand of volume difference is large, it needs a plurality of integral single-layer reservoirs to be used in parallel and connected to the oil suction port through oil pipes.

- Series configuration of VVPR
  This configuration can have higher pressure and good space compactness, but it increases the difficulty of installation adjustment and vertical height of the VVPR.
- Parallel configuration of VVPR
  The variable volume is the largest under the same displacement. In addition, it is easy to maintain and replace components, because the spring is installed externally. However, it occupies more horizontal space and requires multiple springs.

4. TEST AND ANALYSIS

4.1. Prototype and test principle

Based on the mathematical model, three types of reservoir prototypes are designed and manufactured, including single-layer, series, and parallel configurations, as shown in **Figure 5**. The primary parameters of prototypes are shown in **Table 1**.



Figure 5: Prototype of the VVPR: (a) Single-layer reservoir (b) Series reservoir (c) Parallel reservoir

Table 1: Parameters of three configurations

Parameter	Single-layer reservoir	Series reservoir	Parallel reservoir
Structural volume [l]	1	2	2
Variable volume [l]	0.8	1.6	0.8
Pressure [bar]	0.2-0.51	0.15-0.65	0.15-0.62
Total weight of VVPR [kg]	6.3	7.2	13.9
Radius of upper cover [mm]	68	68	68

A test bench is established for the static and dynamic performance of three reservoirs. The test principle is shown in **Figure 6**. The parameters of the hydraulic system are shown in **Table 2**. By changing the state of the cut-off valves, static testing of the VVPR can be carried out. During dynamic testing, the right hydraulic cylinder is connected to the VVPR, and step and sine response tests are conducted on the VVPR through a directional valve controlled hydraulic cylinder.

Table 2: Parameters of hydraulic system

Parameter	Value
System pressure [MPa]	10
System flow [l/min]	20
Cylinder diameter [mm]	100
Cylinder rod diameter [mm]	70
Hydraulic cylinder stroke [mm]	250

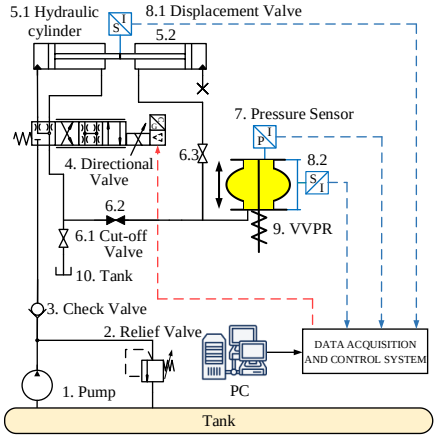


Figure 6: The VVPR test principle

4.2. Static performance analysis

The static performance test is to determine the relationship between pressure, displacement and variable volume during the oil charge and discharge of the three reservoirs, and the results are shown in **Figure 7**.

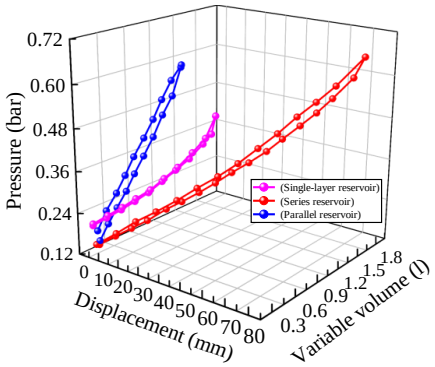


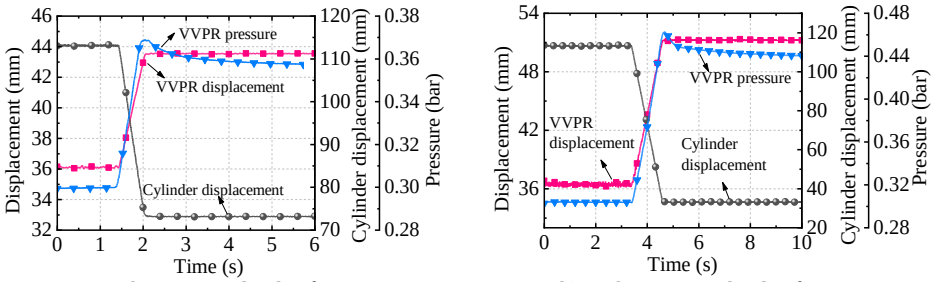
Figure 7: Static performance test results of the VVPR

The pressure and variable volume of the reservoir increase with the increase of displacement. Taking the oil charge process of the series reservoir as an example, when the displacement is 2.13-71.73 mm, the oil charge volume increases from 0.058 l to 1.624 l, and the pressure increases from 0.157 bar to 0.653 bar. In addition, it can be seen that the static test curves of the VVPR do not coincide during the oil charge and discharge, showing obvious hysteresis, in which the hysteresis of the single-layer reservoir is the smallest. Through comparative analysis, the series reservoir can get more pressure. Parallel reservoir can get greater variable volume with the same displacement change.

4.3. Dynamic performance analysis

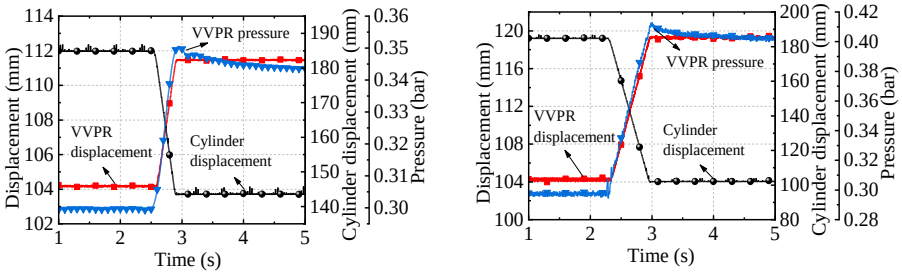
To compare the step response performance of different reservoir configurations, the initial pressure of the reservoir is set at 0.3 bar, and the stroke change of the hydraulic cylinder is 40 mm/80 mm. The step response test curves are shown in **Figure 8-10**. The hydraulic cylinder retracts at 1.428 s, and after a response time of 0.028 s, the pressure of the VVPR increases. After a rise time of 0.72 s, the pressure increases by 0.057 bar and reaches stability, as shown in **Figure 8(a)**. In the experiment,

due to the large instantaneous flow rate caused by the hydraulic cylinder, the pressure of the VVPR is overshooted, and the pressure overshoot of the parallel reservoir is the smallest.



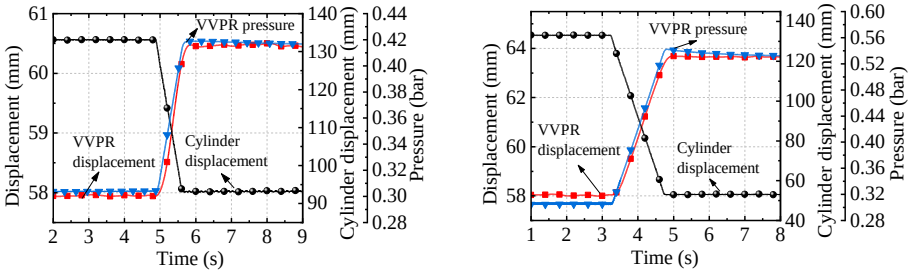
(a) 0.3 bar step amplitude of 40 mm (b) 0.3 bar step amplitude of 80 mm

Figure 8: Step response test curve of the single-layer reservoir



(a) 0.3 bar step amplitude of 40 mm (b) 0.3 bar step amplitude of 80 mm

Figure 9: Step response test curve of the series reservoir



(a) 0.3 bar step amplitude of 40 mm (b) 0.3 bar step amplitude of 80 mm

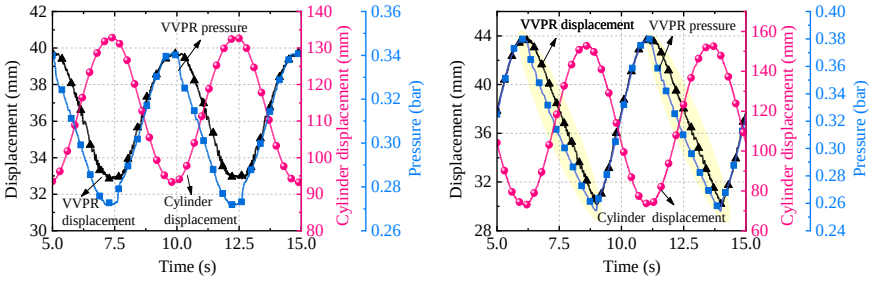
Figure 10: Step response test curve of the parallel reservoir

Taking reservoir pressure as an example, the rise time and pressure variation amplitude are shown in **Table 3**. As the step amplitude of the hydraulic cylinder increases, the rise time and change amplitude of the reservoir pressure curve gradually increase. Among them, the rise time of the parallel reservoir is the longest. The response time of the reservoir is not affected by the step amplitude of the hydraulic cylinder changes, with an average response time of 0.028 s.

Table 3: Step response parameters of the VVPR

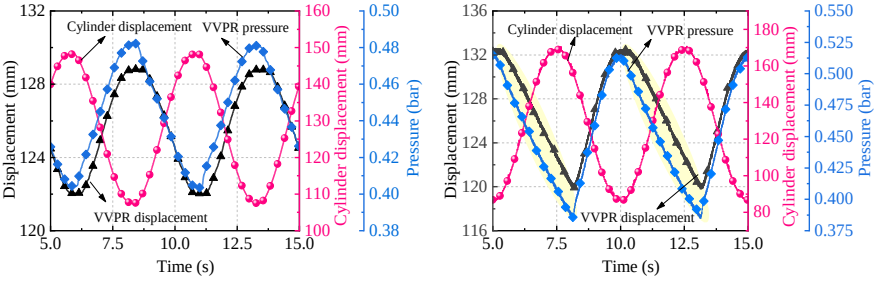
Parameter	Single-layer reservoir		Series reservoir		Parallel reservoir	
Step amplitude [mm]	40	80	40	80	40	80
Response time [s]	0.028	0.026	0.03	0.035	0.02	0.03
Rise time [s]	0.72	1.26	0.44	0.71	0.85	1.54
Pressure variation amplitude [bar]	0.057	0.141	0.044	0.104	0.117	0.229

The sine response test curves with 0.2 Hz ±20 mm/40 mm are shown in **Figure 11-13**, which are the results of single-layer reservoir, series reservoir and parallel reservoir respectively. It is clear that the corresponding pressure and displacement of the VVPR also change with the response of the extension and retraction of the hydraulic cylinder. For example, when the hydraulic cylinder is extended, the VVPR discharges oil, and the displacement and pressure of the VVPR gradually decrease.



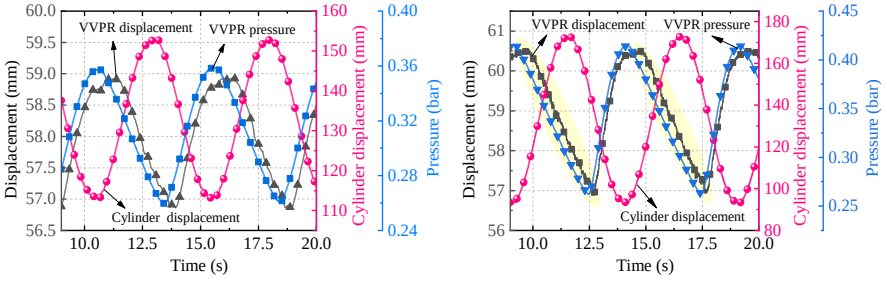
(a) Sine test with 0.2 Hz ±20 mm (b) Sine test with 0.2 Hz ±40 mm

Figure 11: Sine response test curve of the single-layer reservoir



(a) Sine test with 0.2 Hz ±20 mm (b) Sine test with 0.2 Hz ±40 mm

Figure 12: Sine response test curve of the series reservoir



(a) Sine test with 0.2 Hz ±20 mm (b) Sine test with 0.2 Hz ±40 mm

Figure 13: Sine response test curve of the parallel reservoir

According to the test results, the sine response parameters of the VVPR are shown in **Table 4**, and the lag time is the time for the reservoir displacement response to lag behind the hydraulic cylinder displacement response. For example, when the working amplitude of the hydraulic cylinder is 20 mm, the single-layer reservoir lags behind the hydraulic cylinder by 0.11 s, with a phase difference of 0.04π. With the increase of the sinusoidal amplitude of the hydraulic cylinder, the lag phase difference of the reservoir also increases. In addition, with the increase of the amplitude, the displacement and pressure of the reservoir are lagging, and there is a "sharp point" (marked by the yellow area).

Compared with the three configurations, the lag time of the single-layer reservoir is the smallest under the same amplitude.

Table 4: Sine response parameters of the VVPR

Parameter	Single-layer reservoir		Series reservoir		Parallel reservoir	
amplitude [mm]	20	40	20	40	20	40
Lag time (s)	0.11	0.43	0.12	0.68	0.79	1.06
Lag phase difference	0.04π	0.17π	0.05π	0.27π	0.32π	0.42π

5. CONCLUSION AND OUTLOOK

In this study, three different configurations of VVPR are designed and tested, and their performance parameter, structural differences and applicable space occasions are compared. The VVPR can compensate for the volume difference in the hydraulic system and provide pressure for the pump oil supply.

Based on the geometric and force analysis, the relationship among volume, pressure and structural parameters can be obtained. Three prototypes of the reservoir are manufactured. Through comparative analysis, the configuration of the series reservoir is more compact and can get more pressure. Parallel reservoir can get greater variable volume with the same displacement change. In addition, a test bench is built to test the step and sine response performances of the VVPR. The experimental results indicate that in the sine test, the reservoir follows the action of the actuator, but there is a slight delay compared to the movement of the hydraulic cylinder due to the influence of oil flow. The response time of the reservoir is not affected by the step amplitude of the hydraulic cylinder changes, with an average response time of 0.028 s. Compared with the three configurations, the lag time of the single-layer reservoir is the smallest, which is 0.11 s. However, there are a large number of charging and discharging cycles in the application of the VVPR, and the fatigue characteristics deserve further study.

ACKNOWLEDGEMENTS

This work was funded by the National Key Research and Development Program of China (No. 2018YFB2000703).

NOMENCLATURE

d	The cylinder rod diameter	mm
L	Cylinder stroke	mm
ΔV	Volume difference	l
r	The radius of the arc	mm
θ	The half-central angle of the arc	rad
H	The height of the reservoir	mm
S_r	Arc length	mm
R	The radius of the upper and lower cover	mm
V_1	The volume of the single-layer reservoir	l
V_c	The volume of the cylinder between the upper cover and the lower cover	l
V_{arch}	The volume of the body of revolution which is formed by the arch rotating around the central axis	l
A_{arch}	The arch area	mm^2

e	The distance from the centroid O_1 of the arch to the arc chord	mm
P	The pressure of the single-layer reservoir	bar
P_s	The pressure of the series reservoir	bar
P_p	The pressure of the parallel reservoir	bar
m	The mass of the single upper cover	kg
m_p	The mass of the upper cover of the parallel reservoir	kg
k_c	The stiffness of the compression spring	N/mm
Δx_c	The compression of the spring	mm
k_t	The stiffness of the tension spring	N/mm
n	The number of the tension springs	-
l_0	The free length of the tension spring	mm
Δx_t	The stretching length of the tension spring	mm

REFERENCES

[1] Kong X D, Zhu Q X, Yao J, Shang Y X, Zhu Y (2021). Basic theory and key technology of "new method for lightweight design and manufacturing of hydraulic components and systems". Journal of Mechanical Engineering, 57(24), 4-12.

[2] Kong X D, Zhu Q X, Yao J, Shang Y X, Zhu Y (2020) Reviews of lightweight development of hydraulic components and systems for high-level mobile equipment. Journal of Yanshan University, 44(03), 203-217.

[3] Qi G F, Zhang J J, Sun J G (2011) Miniaturization trend of the hydraulic fuel tank and the new trend of development. Machine Tool & Hydraulics, 39(24), 66-68, 104

[4] Wartlick K, Baumann M, Veres A, Weidemann R (2020) CFD-aided optimization of customer-specific tank systems using an innovative labyrinth deaerator. 12th International Fluid Power Conference, Dresden, 2020.

[5] Ouyang X P, Fan B Q, Yang H Y, Ding S (2016) A novel multi-objective optimization method for the pressurized reservoir in hydraulic robotics. Journal of Zhejiang University-SCIENCE A, 17(6), 454-467.

[6] Wang P, Yao J, Feng B D, Li M D, Wang D Y (2022) Modelling and dynamic characteristics for a non-metal pressurized reservoir with variable volume. Chinese Journal of Mechanical Engineering, 35(1), 1-15.

[7] Seguin C (2003) Variable volume reservoir. US20030136123A1.

[8] Chen J J, Yin Z H, Rakheja S, He J H, Guo K H (2018) Theoretical modelling and experimental analysis of the vertical stiffness of a convoluted air spring including the effect of the stiffness of the bellows. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 232(4), 547-561.

SHIP BALLASTING PROCESS TIME CALCULATION WITH USE OF SUBMERGED BALLAST PUMPS WITH HYDRAULIC DRIVE SUPPLIED FROM CONSTANT PRESSURE HYDRAULIC CENTRAL LOADING SYSTEM ON MODERN PRODUCT AND CHEMICAL TANKERS

peer reviewed

Andrzej Banaszek [1], Radovan Petrovic [2*], Sasa Botocanin[3]

[1] *West Pomeranian University of Technology Szczecin / Maritime Research and Transport Faculty, Street Al. Piastów 41, 71-065 Szczecin, Poland , E-mail: andrzej.banaszek@pzut.edu.pl*

[2] *The University "Nikola Tesla" in Belgrade, Faculty of Information Technologies and Engineering, 11158 BEOGRAD SERBIA, Cara Dusana 62-64, e-mail: radovan4700@yachoo.com*

[3] *PPT-Namenska AD – Company Wall,Cara Dušana 101, 37240 Trstenik , Serbia, e-mail: sasa.batocanin@ppt-namenska.rs*

* Corresponding author: e-mail address: Tel. +38 1637580104 radovan4700@yachoo.com

ABSTRACT

Ballast systems are among one of the most important installations installed on board modern product and chemical tankers. They have a significant impact on the safety of the ship by determining its stability on the sea wave and the strength of the hull. Logistically, the efficiency of these systems is crucial in planning the loading and unloading times of a given tanker. Due to the explosion-hazardous zone, ballast pumps of the submersible type are often equipped with hydraulic drive. The purpose of the article is to present a methodology for calculating the time of ballasting process of a modern chemical tanker and the flow control of submerged ballast pumps with hydraulic drive, supplied from the hydraulic constant-pressure central loading system. The results of the calculations are important for the correct determination of the liquid cargo loading time of a tanker at the fuel terminal, the organization of the logistics and service system for the ship and for the entire loading port. In addition, the paper presents the construction of a submersible ballast pump with hydraulic drive with a description of the hydraulic system installed on board a modern tanker. The methods of controlling the flow of submerged ballast pumps are described, along with a new concept of using constant torque regulators to control the pump flow. Experimental results and theoretical calculations are presented.

Keywords: Balalst system , ballast pumps, hydraulic drive, hydraulic central loading system, product and chemical tankers

1. INTRODUCTION

Specialized vessels - product and chemical tankers - are used to transport liquid petroleum cargoes by sea. Ballast systems installed on their decks are among the most important service installations.... They have a major impact on the safety level of the ship. The size of the water ballast determines the depth of the ship, the position of the center of gravity and the metacentric height. The above technical parameters have a significant impact on the ship's wave behavior and wave stability. Often, poor ballasting of such a ship can lead to overturning of the ship on a wave in rough sailing conditions and a maritime disaster. The speed of the ship's captain's reaction in case of stability problems is therefore extremely important in the problem of safe navigation. Many times, on modern product and chemical tankers, the technical capability of the ballast system also determines the total handling time in port, which is extremely important from the logistical and business side. Newly built tankers, according to the requirements of the International Maritime Organization (IMO), must have a hull structure of the "Double Hull" type[14]. In a hull design of this type, each cargo tank must be separated from the ship's outer side by ballast tanks or an empty cofferdam. This arrangement protects the ship from the outflow of liquid cargo in the event of a collision or if the ship enters an underwater obstacle. An example of this type of ship is the B573-I/2 class product tanker m/t "Simunye", built at Szczecin Shipyard S.A. for the shipowner Unicorn Tankers Ltd. of South Africa (see Fig.1). The ship's side ballast tanks, created the possibility of mounting directly inside two (2) submerged ballast pumps with hydraulic drive. This allowed the creation of a new type of ballast system, with the ballast pumps fully installed inside the ballast tanks (as opposed to the traditional solution with ballast pumps installed in the engine room). This saves a lot of space inside the ship's engine room, while reducing the total length of ballast pipes and flow resistance in the ship's ballasting operation. This article presents the idea of a new type of

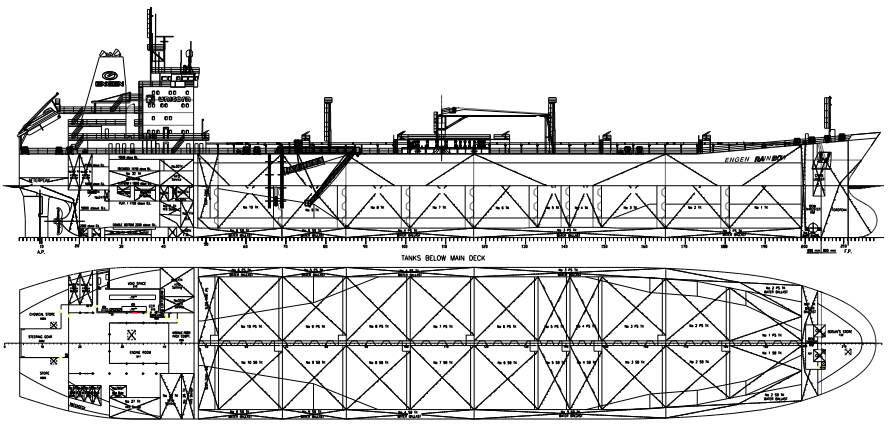


Figure 1. Product tanker B573-I/2 class m/t 'Simunye', built by the Shipyard Szczecinska S.A/Poland for the ship-owner -Unicorn Tankers from Republic of South Africa [16]

ballast system with submerged ballast pumps fed from a constant-pressure hydraulic central loading system. Also described are ways of constant-pressure control of the capacity of submerged ballast pumps using constant Torque controllers and their influence on the time of

ship ballasting operations. Experimental results and theoretical calculations of the flow of a submerged ballast pump for different drop pressure settings in a hydraulic drive motor are presented In the literature, due to the specificity of the maritime subject, there are no studies on the above topic. We can only note the of the paper authors : (Banaszek [1]), Banaszek, Petrovic [2][4], Banaszek,Urbanski [3], describing the use of a hydraulic central loading system to drive submerged ballast pumps, as well as the works of Gorski-Perepeczko [10][9], Kutyrkin-Postnikov [12] and the instructions of pump manufacturers Framo [5], Hyunday [6], Shinko [7] and also other authors [17][18]. In this article, the theoretical calculations carried out are supplemented by the presentation of exemplary solutions of this type of hydraulic systems installed on board product and chemical tankers built in Szczecin Shipyard / EU-Poland.

2. DESCRIPTION OF THE STRUCTURE OF THE BALLAST SYSTEM WITH SUBMERSIBLE HYDRAULICALLY DRIVEN BALLAST PUMPS

The hull design of "Double Hull" types on modern product and chemical tankers and the resulting side ballast tanks made it possible to install submerged ballast pumps directly in them (see Fig.2.). Submersible ballast pumps are generally 1-stage, centrifugal pumps. According to Lloyd's Register regulations [14], this is a hazardous area due to the possibility of explosion. For this reason, the most common way to drive them on modern tankers is to use hydraulics, as safer than the alternative solution with electric motors. Total length of submerged ballast pumps

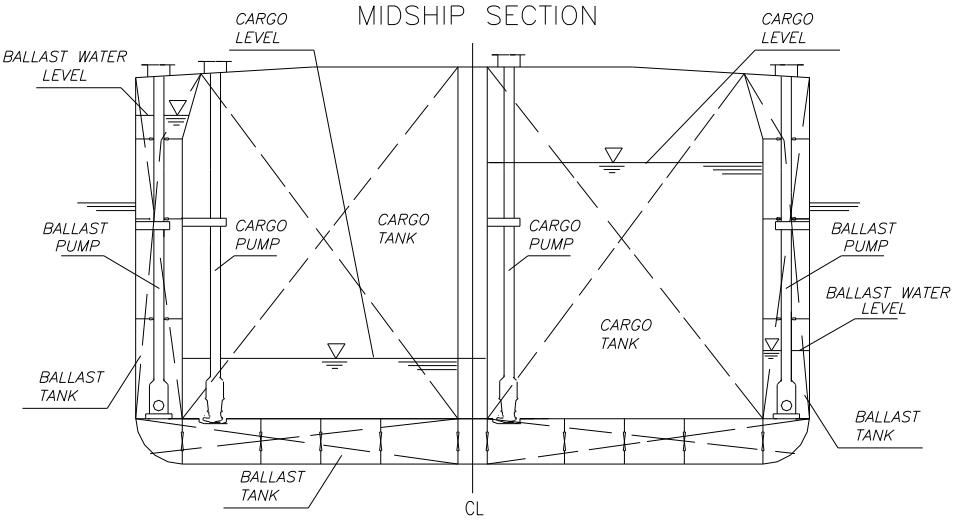


Figure 2. Submerged ballast pumps montage inside of side ballast tanks

must be adjusted each time to the situation on the ship - to the current height of the ballast tanks (see Fig.2). As a result, the height of the submerged pumps reaches 14-20 [m]. Older solutions with a long drive shaft and electric motor often experience vibration problems during the operating period. This is often the result of the unbalance of the long drive shaft. Mounting the hydraulic drive motor at the bottom, directly in the ballast pump head, reduces the length of the shaft and avoids excessive vibration problems to a minimum (see Fig..3a.). Due to the difficulty of service access to the head during normal operation of the pump (see Fig.3b), a fixed-pitch

A2FM-type axial piston hydraulic motor from Bosch-Rexroth/Germany (recognized as one of the most reliable hydraulic motors in its class [8]) was used as the drive motor. On the top plate, located directly on the deck coaming on the open deck, there is a hydraulic block designed to control the flow of oil supplying the hydraulic motor. Connected to the above block are hydraulic branches running from the hydraulic central cargo system. Hydraulic oil flows from the block to the hydraulic engine through a system of concentric lines, where there is a pressure line in the center, around the return line area, insulated on the outside with a glycerine filled cofferdam. In the event of loss of tightness through the hydraulic lines - the leaking

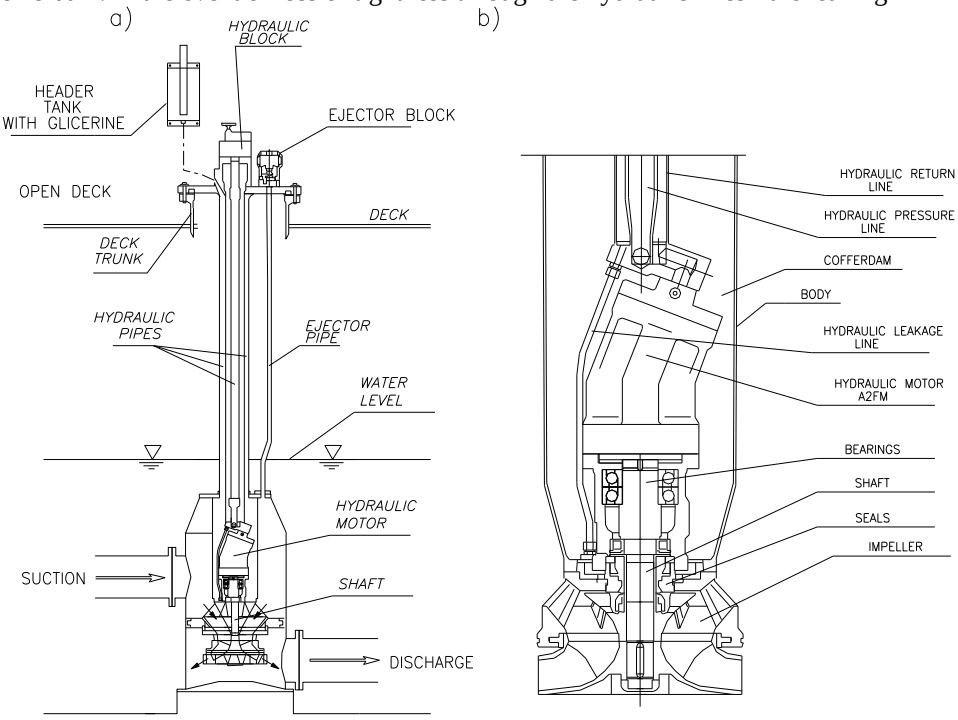


Figure 3. Typical submerged ballast pump with hydraulic drive (Framo/Norway) [5]

oil pushes the glycerin out of the tank to the outside alerting the service to the damage of the pump (see Fig.3a). Glycerin is fully safe for the natural marine environment, as it is fully biodegradable in seawater. Fig.4 shows the construction of a typical new-type water ballast system with submerged ballast pumps, installed on board modern product and chemical tankers. It consists of two ballast buses to which all ballast tanks are connected by separate pipe branches. A system of remotely controlled ballast valves allows any configuration of such a system. In the event of damage, it is possible to cut off the damaged pump and continue pumping operations by a second ready ballast pump. Ballast piping on typical ships is generally long, running along the entire double bottom of the ship. This situation can cause cavitation problems during emptying operations of ballast tanks, especially remote from the pumps. For this reason, each ballast pump is equipped with a semi-draining system with an air separator. Two parallel ballast trunk lines with two separate and independent ballast pumps increase the reliability of this important ship system in the process.

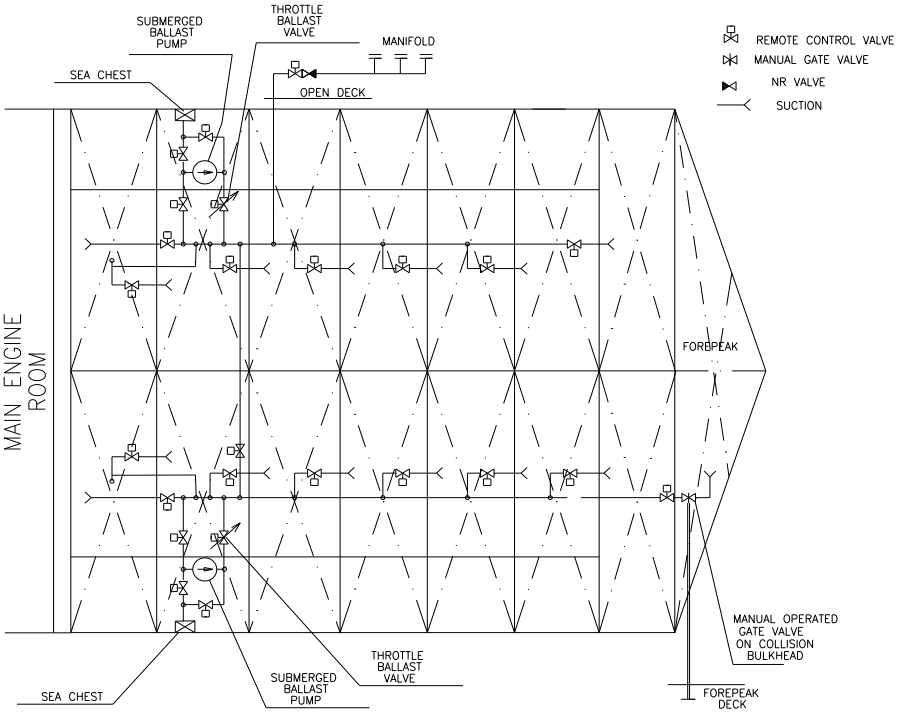


Figure 4. The water ballast system diagram with submerged ballast pumps

3. HYDRAULIC DRIVE SYSTEM WITH CONSTANT-TORQUE FLOW CONTROL

On board the B 573-I/2 class product tanker m/t "Simunye"[16] were 2 (two) submerged ballast pumps of the SB300 (Framo) type, with a nominal discharge flow of Q_{BPO} = 800 each, obtained at discharge pressure H_{BPO} =25 [mlc], When pumping seawater with a density of ρ_W=1,02 and kinematic viscosity ν_W =1,0 [cSt]. A hydraulic controller fulfilling the principle of a hydraulic oil flow controller was installed on the supply hydraulic line running from the hydraulic central main line. The controller set a constant value for the flow of hydraulic oil supplying the hydraulic pump motor, regardless of its load. Hydraulic motors of the A2FM axial-piston type (Bosch Rexroth [8]), installed in ballast pumps, are characterized by high volumetric efficiency. The value of volumetric leakage does not exceed the level of 5% of the total flow demand. Therefore, it can be approximated that the value of the flow demand of a hydraulic drive motor is proportional to the speed of the motor shaft. Figure 5 shows the flow and drive characteristics of the SB300 type ballast pump at nominal speed. The nominal value of the motor flow demand Q_S =240[l/min], which corresponds to the speed of the pump impeller n_S =1164 [rpm].

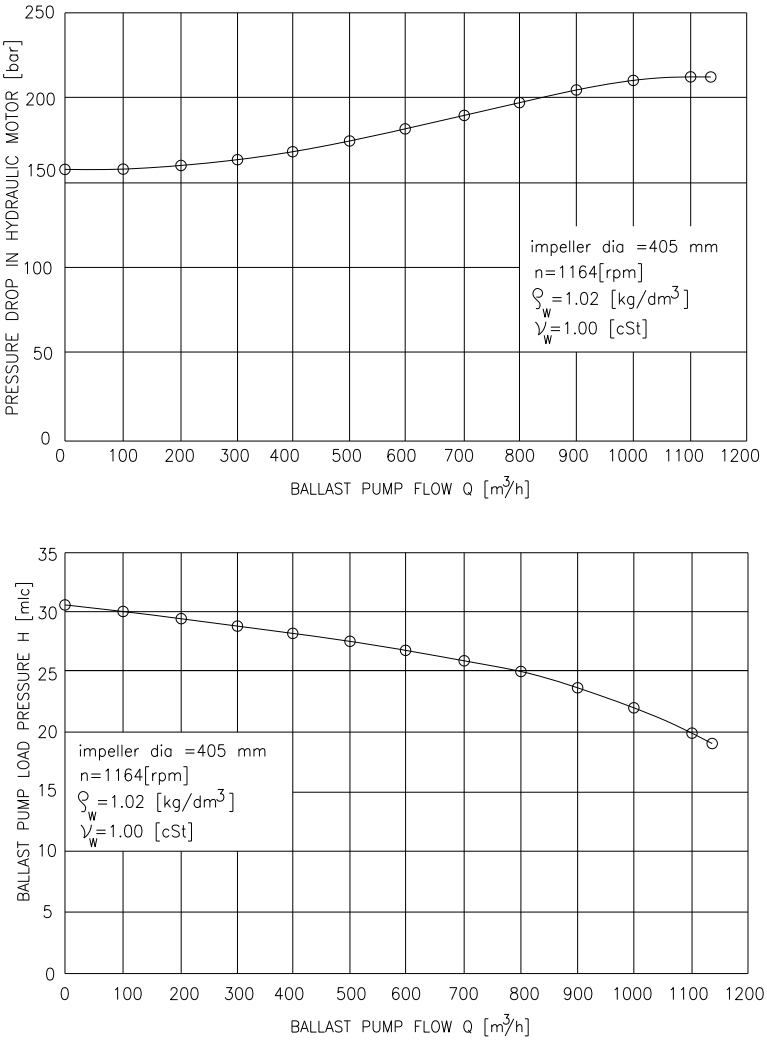


Figure 5. Technical characteristics of submerged ballast pump SB300 type (Framo) [5]

In order to better stabilize the flow value of the ballast pump, constant-torque controllers were used for capacity control. Its task is to maintain a constant oil pressure drop in the hydraulic drive motor. The staff supervising the operation of the hydraulic central loading system on board the described tanker can control the following technical parameters of the hydraulic power system:

p_G - the delivery pressure of the main power pack and hydraulic central loading system

Δp_S - the value of hydraulic oil pressure drop in the hydraulic motor of the investigated device connected to the hydraulic central loading system

The value of the first parameter can be adjusted using the p=const controller installed on the main power pack in the Power Pack Room. In the adjustment procedure, it should be remembered that the value of the discharge pressure in the hydraulic system should not be less than the total pressure drop of the hydraulic oil in the hydraulic motor of the most loaded

hydraulic consumer and the pressure drop in the hydraulic system, including all hydraulic control valves in the path of hydraulic oil flow. In the case of ballast pumps, the value of the pressure drop Δp_S =const is regulated by a hydraulic controller in the pump's hydraulic block. In Fig 6. one showed the structure the pressure drops of hydraulic oil in the supply process of analyzed ballast pump. Hydraulic oil discharged by the central power pack is moved by the main pressure line of hydraulic central loading system and the master valve (constant-torque controller) to the driving motor of ballast pump and then comes back by the main return line to the power pack

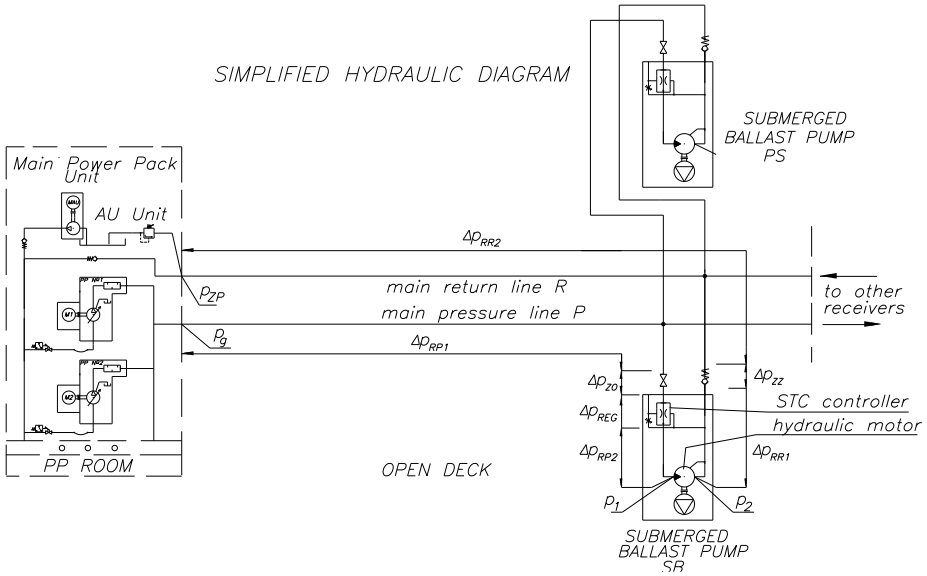


Figure 6. Hydraulic oil pressure drops in supply process of ballast pump from hydraulic central loading system

In compliance with diagram, one can write that pressure drop Δp_S in hydraulic motor of ballast pump is equal:

$$\Delta p_S = p_1 - p_2 \tag{1}$$

where : p_1, p_2 - hydraulic motor inlet and outlet pressure,

The minimum working pressure of the hydraulic power pack p_G, must be not less than:

$$p_G = p_1 + \Delta p_{RP} + \Delta p_{ZO} + \Delta p_{REG} \tag{2}$$

where: Δp_{RP} - pressure drop in hydraulic main pressure line

$$\Delta p_{RP} = \sum_{i=1}^{n} \lambda_i \cdot \frac{l_i}{d_i} \cdot \rho \cdot \frac{\upsilon_i^2}{2} + \sum_{j=1}^{m} \zeta_j \cdot \rho \cdot \frac{\upsilon_j^2}{2} \quad ; \quad \upsilon_{i,j} = \frac{4Q_{i,j}}{\pi \cdot d_{i,j}^2} \tag{3}$$

where: $d_{i,j}$, $Q_{i,j}$, $v_{i,j}$ -pipe diameter , oil flow, oil velocity, Δp_{ZO} - ball valve pressure drop, Δp_{REG} - pressure drop in hydraulic constant-torque controller

Hydraulic motor outlet pressure can be calculated from the equation :

$$p_2 = p_{ZP} + \Delta p_{ZZ} + \Delta p_{RR} \tag{4}$$

where: - Δp_{RR} hydraulic main return line pressure drop

$$\Delta p_{RR} = \sum_{i1=1}^{n1} \lambda_{i1} \cdot \frac{l_{i1}}{d_{i1}} \cdot \rho \cdot \frac{\upsilon_{i1}^2}{2} + \sum_{j1=1}^{m1} \zeta_{j1} \cdot \rho \cdot \frac{\upsilon_{j1}^2}{2} \quad ; \quad \upsilon_{i1,j1} = \frac{4Q_{i1,j1}}{\pi \cdot d_{i1,j1}^2} \tag{5}$$

- Δp_{ZZ} - NR valve pressure drop, p_{ZP} - adjustment pressure in support relief valve of main hydraulic power pack unit (see diagram in Fig.6)

In typical hydraulic systems of central loading, the value of the pressure p_{ZP} carries out 2-10 [bar] in dependences from the size of the system and the auxiliary hydraulic unit. In hydraulic central loading system installed on board product tanker m/t „Simunye" adjusted value of the opening support relief valve carried out: $p_{ZP} = 4$ [bar] (6)

Taking into account equals (2) and (4), the minimum value of the delivery pressure of the central hydraulic power pack should be not less than:

$$p_{G\min} \geq \Delta p_s + \Delta p_{RP} + \Delta p_{ZO} + \Delta p_{REG} + p_{ZP} + \Delta p_{ZZ} + \Delta p_{RR} \tag{7}$$

For the purpose of avoiding of problems with an unstable acting of the p=const. controller, the adjusted value p_{Gmin} in effect is accepted on the higher level:

$$p_G = \Delta p_s + \Delta p_{RP} + \Delta p_{ZO} + \Delta p_{REG} + p_{ZP} + \Delta p_{ZZ} + \Delta p_{RR} + \Delta p_{SS} \tag{8}$$

where Δp_{SS} - pressure drop safety surplus, typically $\Delta p_{SS} = 10$ [bar]

In analyzed case, as was above described, the hydraulic controller STC -90 adapts the value of the pressure drop Δp_{REG} in this way that the size the pressure drop in the hydraulic motor must be constant (control algorithm formula): $\Delta p_S = \text{const}$ (9)

4. MATHEMATICAL MODEL

The drive torque on the impeller drive shaft of ballast pump (is described by means of the Schloesser formula [15]:

$$M_S = M_{ST} - \Delta M_{mf} - \Delta M_{m\mu} - \Delta M_{mh} - \Delta M_C \tag{10}$$

where: - M_S - actual torque on hydraulic motor drive shaft

- theoretical torque on motor drive shaft

$$M_{ST} = \frac{Dm}{2\pi} \cdot \Delta p_S$$

- Dm - hydraulic motor stroke displacement

- ΔM_{mf} - torque losses caused by mechanical friction

$$\Delta M_{mf} = C_{mf} \cdot \frac{Dm}{2\pi} \cdot \Delta p_S \tag{11}$$

C_{mf} - mechanical friction coefficient dependent on pump construction

- $\Delta M_{m\mu}$ - torque losses caused by viscosity friction

$$\Delta M_{m\mu} = C_{m\mu} \cdot \mu \cdot n_S \cdot Dm \tag{12}$$

$C_{m\mu}$ - viscosity friction coefficient dependent on pump construction and oil parameters,

μ- dynamic viscosity of the hydraulic oil, n_S - ballast pump rotation velocity

- ΔM_{mh} - torque losses caused by hydraulic losses in a motor

$$\Delta M_{mh} = C_{mh} \cdot \frac{\rho \cdot {n_S}^2}{4\pi} \cdot \sqrt[3]{{q_S}^5} \tag{13}$$

C_{mh} - hydraulic losses coefficient , ρ- hydraulic oil density, ΔM_C - constant loss torque in hydraulic motor

The dependence (10) can be write in the simplified form as:

$$M_S = \frac{Dm}{2\pi} \cdot \Delta p_S - \Delta M_S = M_{ST} - \Delta M_S \tag{14}$$

where: ΔMs - total torque losses in hydraulic motor

$$M_S \approx M_{ST} = \frac{Dm}{2\pi} \cdot \Delta p_S \tag{15}$$

This means that the constant value of the hydraulic oil pressure drop in the ballast pump driving motor, holding consequently quasi-constant value of the impeller driving torque.

In compliance with the dynamic similarity theory of the centrifugal pumps (Łazarkiewicz-Troskolański [11], Jedral [13]), the change of the rotation speed of the pump impeller with relation to the nominal speed causes the change of main technical parameters of the pump according to following formula:

$$\frac{Q_{BPO}}{Q_{BP}} \cong \frac{n_n}{n_S} \tag{16}$$

$$\frac{H_{BPO}}{H_{BP}} \cong \left(\frac{n_n}{n_S} \right)^2 \tag{17}$$

where: H_{BPO}, Q_{BPO}, n_n - technical parameters of ballast pump at nominal pressure drop in a motor

H_{BP}, Q_{BP}, n_S - technical parameters of ballast pump at nominal pressure drop in a motor

In result ,from equations (17-19), the corrected values of the ballast pump flow characteristics can be calculated from the following simplified equations:

$$H_{BP} = H_{BPO} \cdot \frac{\Delta p_S}{\Delta p_{S0}} \tag{18}$$

$$Q_{BP} = Q_{BPO} \cdot \sqrt{\frac{\Delta p_S}{\Delta p_{S0}}} \tag{19}$$

where: Δp_S [bar], Δp_{S0} [bar]- nominal and actual pressure drop in hydraulic motor

5. EXPERIMENTAL AND CALCULATION RESULTS

To verify the flow control system of the pump with staomomentum control, experimental measurements of the flow of submerged ballast pumps of the SB300 type were made. The measurements were carried out at a special measuring station in the laboratory of the pump manufacturer FRAMO / Norway. Flow measurements were made when pumping seawater with a density of $\rho_w = 1025[kg/m^3]$ and kinematic viscosity $\nu_w = 1.0[cSt]$. The diameter of the ballast pump impeller was 405 [mm]. Axial piston hydraulic motor type A2FM200 with fixed displacement/stroke $Dm = 200[cm^3/rot]$ [12] was used as hydraulic motor. HLP-46 grade hydraulic oil with kinematic viscosity was used to drive the hydraulic motor. $\nu_H = 46.0[cSt]$ in temperature measurements $T_o = 50[^oC]$. The above data were in accordance with the

manufacturer's recommendations. Pump flow measurements were made at 7 pressure drop settings on the STC-90 constant torque flow
controller: $\{\Delta p_s\} = \{250, 225, 200, 175, 150, 125, 100\}[bar]$

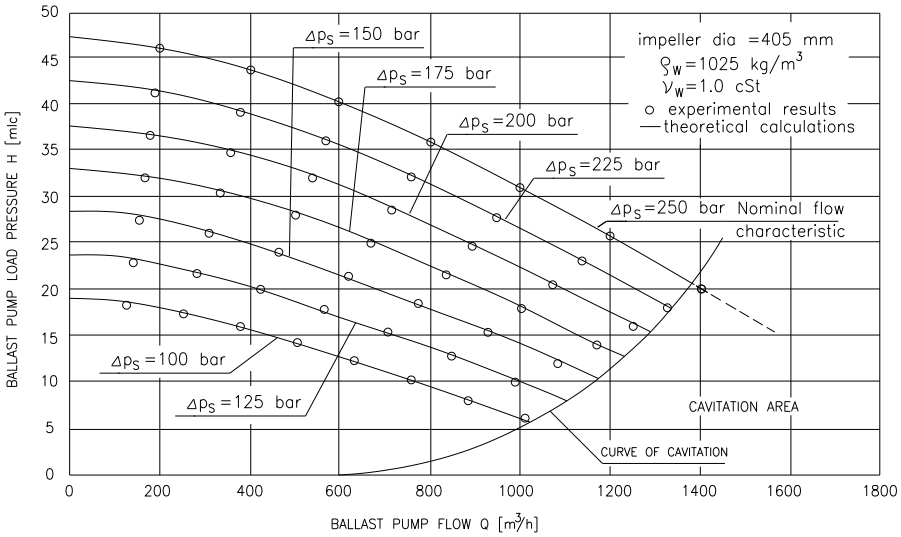


Figure 7. Experimental and calculation results of submerged ballast pump SB300 type, with the hydraulic drive and constant-torque flow control for seven (7) different drop pressure adjustments Δp_s

The results of experimental measurements of the described ballast pump are shown in Figure 7. They were compared with the results of theoretical calculations. Good agreement was obtained between the above results, confirming the usefulness of the mathematical model formulated in the paper. Using the proposed mathematical model, an example calculation of the time to carry out the ballasting process on the product tanker m/t "Simunye" was carried out for the following data : Volume of pumped water ballast: 12 000 cubic m, average flow resistance in the ballast system - 21 [mlc], setting constant torque controller STC-90 - Δp_{ss} = 200 [bar]
As a result, the time of total ballasting of the ship was obtained $t_{ballast}$= 11 hour 36 min. This result was consistent with the ship's actual ballasting time with an error of 1%, confirming the utility of the computational model.

6. CONCLUSIONS

Ballast pumps of the submersible type on board modern product and chemical tankers are the most important part of the ballast system. They are installed directly in the side ballast tanks, which are, according to the regulations of the calibration institutions, treated as a hazardous, explosive zone. Therefore, the ballast pumps described are usually equipped with a hydraulic drive, fed from a hydraulic central supply system. On the B573-I/2 class product tanker m/t "Simunye" it was decided to use the existing hydraulic central power system on board to drive the pumps. Hydraulic constant-torque controllers were used to control the flow of the ballast pumps for better stabilization of the discharge of the delivery pump with changing flow resistance in the ballast system and the water level in the individual ballast tanks. The paper presents the results of measurements of flow characteristics of ballast pumps operating at different from nominal values of hydraulic oil pressure drop in the engine, set in the constant-

torque controller. The paper presents a computational model of ballast pump performance with hydraulic drive and control by a constant-torque controller. The results of experimental measurements of the ballast pump flow were compared with the results of theoretical calculations. Good agreement between the two results was confirmed. Therefore, the article is utilitarian in nature and can be helpful to ship ballast system designers, tanker fleet management personnel and logisticians responsible for the port fuel terminal.

7. LIST OF NOTATIONS

p_G	1 the working pressure of the main power pack and hydraulic central loading system	bar
Δp_S	the value of hydraulic oil pressure drop in the hydraulic motor	bar
p_1	inlet pressure to the hydraulic motor	bar
p_2	outlet pressure from the hydraulic motor	bar
Δp_{RP}	pressure drop in hydraulic main pressure line	bar
$d_{i,j}, Q_{i,j}, \upsilon_{i,j}$	diameter in pipe, oil flow, oil velocity	$m, dm^3/\min, m/s$
Δp_{RR}	pressure drop in hydraulic main return line	bar
p_{ZP}	adjustment pressure in support relief valve in filling up system	bar
M_S	performance torque on hydraulic motor drive shaft	Nm
Dm	hydraulic motor stroke displacement	cm^3/rot
n_S	cargo pump rotation velocity	rpm
ρ	hydraulic oil density	kg/m^3
$C_{mf}, C_{m\mu}, C_{mh}$	proportional coefficients experimentally determined	-
ΔM_{mf}	torque losses caused by mechanical friction in a motor	Nm
$\Delta M_{m\mu}$	torque losses caused by viscosity friction in a motor	Nm
ΔM_{mh}	torque losses caused by hydraulic losses in a motor	Nm
ΔM_C	constant loss torque in hydraulic motor	Nm
M_{ST}	theoretical torque on motor drive shaft	Nm
ΔM_S	total loss torque in hydraulic motor	Nm
T_o	temperature of oil	^oC
ρ_w	density of sea water	kg/m^3
ν_w	kinematic viscosity of sea water	cSt
ν_H	kinematic viscosity of hydraulic oil	cSt
μ	dynamic viscosity of the hydraulic oil	cP
Q_{BPO}	nominal flow of ballast pump	m^3/h
H_{BPO}	nominal load pressure of ballast pump	M
Δp_{SO}	The nominal value of hydraulic oil pressure drop in the hydraulic motor	bar

REFERENCES

[1] Banaszek A.(2021): Methodology of flow rate assessment of submerged hydraulic ballast pumps on modern product and chemical tankers with use of neural network methods. Elsevier Procedia Computer Science 192 (2021) 1894-1903

[2] Banaszek A., Petrovic R. (2019) : Problem of Non Proportional Flow of Hydraulic Pumps Working with Constant Pressure Regulators Big Power Multipump Power Pack Unit in Open System, Technical Gazette 26, 2/2019 294-301.

[3] Banaszek A. Urbanski T**.**(2020): The flow calculation algorithm of submerged hydraulic cargo pumps working with reduced pump speed on modern product and chemical tankers, Elsevier Procedia Computer Science 192 (2020) 2868-2877

[4] Banaszek A., Petrovic R. Andelkovic M. Radosavjevic M. (2022) : Efficiency of a twin-two-pump hydraulic Power Pack with pumps equipped into constant pressure regulators with different linear performance characteristics, Energies, 21/2022, 10.3390/en1518100

[5] FRAMO Services AS (1993): *Operation of FRAMO cargo pumps*, Service Bulletin No.10, Bergen, Norway

[6] Hyunday Heavy Ind Co Ltd.(2014) : "Hi-Well Cargo pumps System Hydraulically driven" South Korea, Ulsan

[7] Shinko Ind Co Ltd.(2020) : "Cargo Oil, Tank Cleaning &Ballast Pumps" Japan, Hiroshima

[8] Bosch Rexroth Group GmbH (2018): Catalogue RD 92050, **1**-st edn,, Germany, Lohr am Main.

[9] Perepeczko A.(1995): *Exploitation installations on tankers* (in Polish), Copyright by Wydawnictwo uczelniane Wyższej Szkoły Morskiej w Gdyni, Poland

[10] Gorski Z., Perepeczko A.(1994): *Ship's pumps* (in Polish), Copyright by Wydawnictwo uczelniane Wyższej Szkoły Morskiej w Gdyni, Gdynia, Poland

[11] Troskolanski A.T., Łazarkiewicz S.(1973): *Centrifugal pumps* (in Polish), 3-th Edition, Copyright by WNT, Warsaw.

[12] Kutyrkin W.A., Postnikow W.I.(1983): *Specialnyje sitemy neftenaliwnych sudow* (in Russian), Copyright by Transport, Moskow

[13] Jedral W.(2001): *Centrifugal pumps (in Polish)*, Copyright by PWN, Warsaw, ISBN 83-01-13552-2

[14] Lloyd's Register (2022): Rules and Regulations for the Classification of Ships, July 2022, https://imorules.com/LRSHI

[15] ANSI/HI (2019)**:** Rotodynamic pumps- Guideline for Effects of liquid Viscosity on performances. ANSI/HI No 9.6.7-2019, New York, USA

[16] Szczecinska Shipyard Porta Holding SA Design Office (1996), Technical Test Raport, B573-I /2/T-542-1-1, Szczecin, 1996

[17] Zuo, Jianjun, Banjar, Hattan, Xia, Zhenyan et. al., (2016), CFD simulation and experimental study of oil viscosity effect on multistage electrical submersible pump (ESP) performance, Journal of Petroleum Science and Engineering Vol. 146 : 735-746

[18] Otuchi E.M., Stel H., Vieira T.**s**. et. al., (2017). Study of the effect of viscosity on the head and flow rate degradation in different multistage electric submerged pumps using dimensional analysis, Journal of Petroleum Science and Engineering Vol.156 : 442-450

Chapter 10

Independent Metering in Mobile Applications

COMPACT FLUID POWER CONTROL UNIT WITH INDEPENDENT METERING

peer reviewed

Mathias Niebergall[1]*, Hannes Ziegler[1,2]

[1]*Hydraulics Center of Excellence, University of Applied Science Ulm, Prittwitzstraße 10, 89075 Ulm*
[2]*Tries GmbH & Co. KG, Hydraulik-Elemente, Röntgenstr. 10, D-89584 Ehingen*

* Corresponding author: Tel.: +49 731 96537-545; E-mail address: mathias.niebergall@thu.de

ABSTRACT

safety and installation space restrictions, require new customized technological solutions on a system level. Independent meter-in and meter-out control edge arrangements with aligned electronic network and software solutions provide the freedom for the design of intelligent overall fluid control systems. The focus of the paper presented is the development of a new fluid control unit, that meets the following requirements: high power based on a maximum pressure level > 500 bar and a flow rate of up to 200 l/min, an appropriate valve stroke, valve speed, and installation space according to the specification. The fluid control unit presented has independent control edges actuated by a step motor, and a compact pilot control solution. On system level the modelling and simulation of an excavator implement drive has been carried out by means of a typical operating cycle. Firstly for a throttle system with conventional linked control edges and already with a customized load sensing pump control. Secondly for an optimized throttle system with independent metering, hydromechanical flow sharing, a customized load sensing pump control, and with an electronic control of the regeneration circuit. The comparison of the simulation results of both system variants shows already an excellent reduction of the energy consumption of 20% for the optimized implement drive system .

Keywords: Fluid power control unit, Independent metering, System simulaton

1. INTRODUCTION

Tighter guidelines on mobile machine implement drive systems in terms of power requirements, emission standards, energy efficiency, system safety, and installation space restrictions lead to new customized technological solutions on system level. To meet the increasing requirements, various drive and control technologies have been discussed since the turn of the century (e.g. [1], [2]). With focus on valve control, particularly independent metering has been the base of various developments [3], [4], [5], [6], [7], [8], [9]. Goal of the project presented is the implementation of engineering research findings in an intelligent fluid power control unit that enables advanced functionalities of high power implement drives.

The focus of the paper presented comprises the conception (section 2) and the design (section 3) of a hydraulic control unit to meet the following specification:

- significant reduction of energy consumption
- high fluid power with maximum pressure > 500 bar and a flow rate of 200 l/min
- valve spool adjustment key data: travel +/- 8 mm, speed 100 mm/s, and force 5000 N in case of direct control
- limited installation space (power density)
- intelligent control by integration of electronic hardware and software

- fail safe functionality
- acceptable investment cost (Return on Invest).

The focus in section 4 is the simulation and assessment of excavator implement drive systems with regard to their energy efficiency. The energy consumption of a conventional hydro-mechanical control system with connected control edges has been compared with the consumption of an electrohydraulic control system with independent metering.

2. CONCEPTION OF THE VALVE CONTROL SYSTEM

Within the valve control system the actuation system is of particular significance, due to its function between the Electronic Control Unit (ECU) and the hydraulic power section. Adjustment limits of various actuation systems in terms of force, travel and speed, are compared in **Figure 1** based on [10]. It becomes clear, due to restricted space requirements, that high pressure and consequently high adjustment forces in a dimension of 5000 N exclude a direct control concept. By means of a pilot control concept the adjustment limits of the actuation system as well as the space restrictions are accessible. Various actuation principles for the pilot control stage have been discussed. The preferred solution consists of a step motor with spindle, due to the sufficient travel, speed, and force for the adjustment, high position accuracy, stability, and the fail safe usability.

With respect to **Figure 1** the design of a pilot control actuator with a combination of a step motor and a spindle comply with the given adjustment limits. High power requirements in combination with limited installation space of the fluid control unit need an extremely compact pilot control valve stage. Therefore, the pilot piston is arranged concentrically inside the power stage (control chamber, control sleeve, main piston) according to **Figure 2**. By means of the actuator the axial position of the pilot piston is changed.

The central control edge of the pilot piston opens and enables a volume flow in the control chamber. Increasing pressure in the control chamber move the control sleeve together with the attached main piston as far the central control edge of the pilot piston is closed again. So the high pressure and flow rate levels of the power stage are manageable.

The forces at the actuator (spindle, step motor) for moving the pilot piston are controllable small, due to the pressure compensated pilot piston [11].

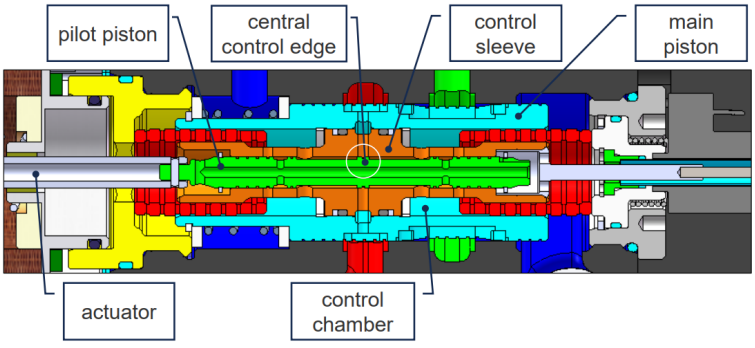


Figure 2: Compact pilot control valve concept [11]

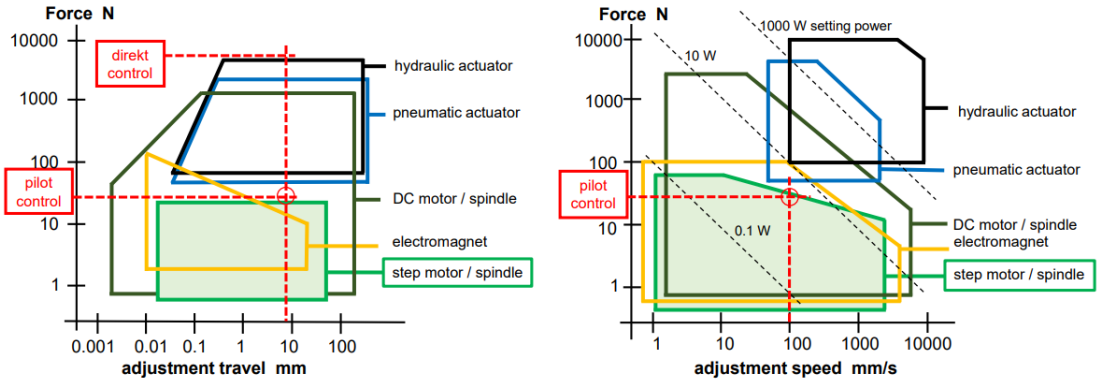


Figure 3: Comparison of relevant actuation principles (data based on [10])

To meet the specification the hydraulic control unit consist of a compact hydro-mechanical pilot valve spool arrangement (**Figure 2**), actuated by a step motor with very high power density and with a low friction spindle (compare **Figure 3**). The functional integration contains, if appropriate, additionally mechanical pressure relief valves with anti-cavitation function and a position sensor, a mechanical fail safe mechanism, an electronic control unit, and aligned software. To overcome the restrictions of conventional valve control systems with mechanically linked orifices, an independent meter-in and meter-out control edge concept with two 3/3 way proportional valves has been preferred. Separate valve control edges enable freedom for the versatile configuration of an intelligent fluid control system.

The actuation concept of pilot control has been built as a test set up (compare **Figure 4**). The actuation of the pilot valve is carried out by a step motor (1) with a smooth-running ball bearing spindle (2). Precondition for fail safe is a spindle carried out as double-threaded ball bearing screw with insignificant self-locking. In case of electric current disconnection (fail safe) the screw nut (3) (including connecting element (4) and connected pilot valve) is enabled to move back to its initial position due to the return spring mechanism (5). For the measurement of the position of the movable parts (4, 3) with reference to a fixed object (6) a laser position sensor (7) is attached.

By means of the test setup, investigations of the actuator requirements in terms of adjustment travel +/- 8 mm, adjustment speed 100 mm/s, pilot valve adjustment force (ca. 30 N), and the fail safe functionality have been carried out. Due to the pressure compensated pilot valve only the spring force is dominant and has to be overcome. In general the additional impact of flow forces results in an accelerate closing of the control edges of the pilot spool valve. This effect is useful in case of fail safe in particular.

The dynamic behaviour of the pilot control mechanism by means of the step motor with spindle is shown in Figure 4 for a current step signal of 2 A.

The measured adjustment travel of the valve spool is in the required range of ±8 mm and the moving time of ca. 0.04 s (compare **Figure 5** below) correspond to a speed level of 200 mm/s (requirement 100 mm/s).

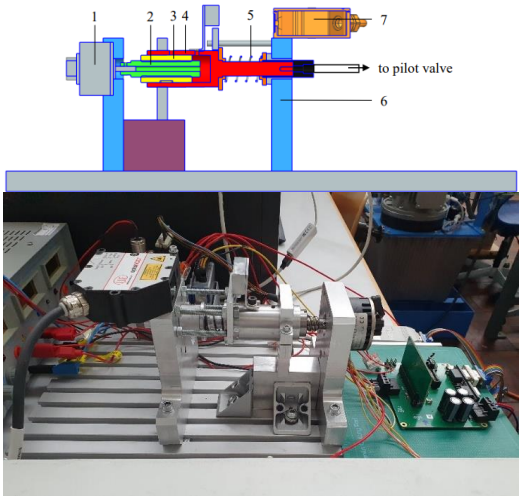


Figure 4: Schema of the actuation concept (top) and test setup (bottom)

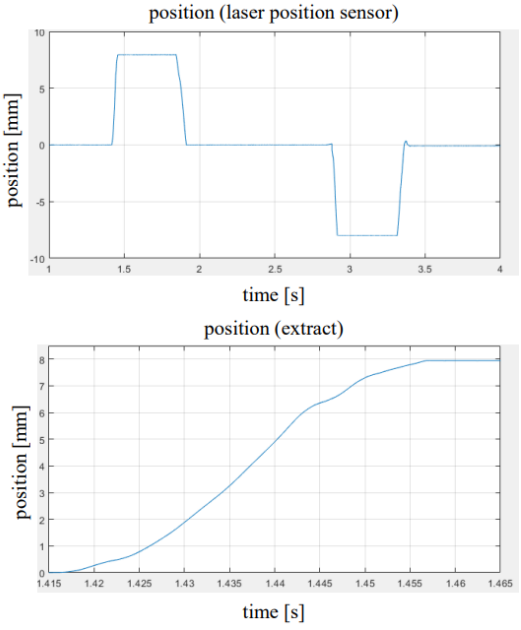


Figure 5: Measured adjustment travel of the pilot control actuator

3. FLUID POWER CONTROL UNIT

The fluid power control unit has been designed with respect to the various functional requirements and the limited installation space. The compact design of a block of several fluid power control units is illustrated in **Figure 6.** The sectional drawing of a single fluid power control unit is shown in **Figure 7**.

A single fluid power control unit contains of the following functional modules (**Figure 6**):

a) actuator module with step motor (4), electronic board (6), ball bearing screw with minimal self-locking (double thread) (2), and a fail safe mechanism (3)

b) main module with two independent stacked pilot-operated 3/3 way proportional valves (1) including centric integrated pressure compensated pilot valve spools, as well as a pressure sensor (8), and a position encoder (7)

c) outlet plates with pressure relieve and anti-cavitation valves (5)

d) connector module as electric interface (9)

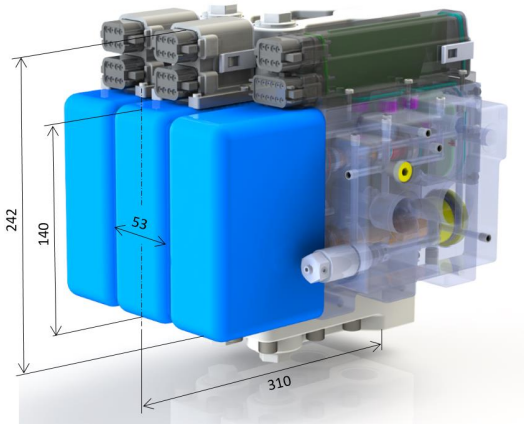


Figure 6: Valve block design of fluid power control units with independent metering (source TRIES GmbH & Co. KG)

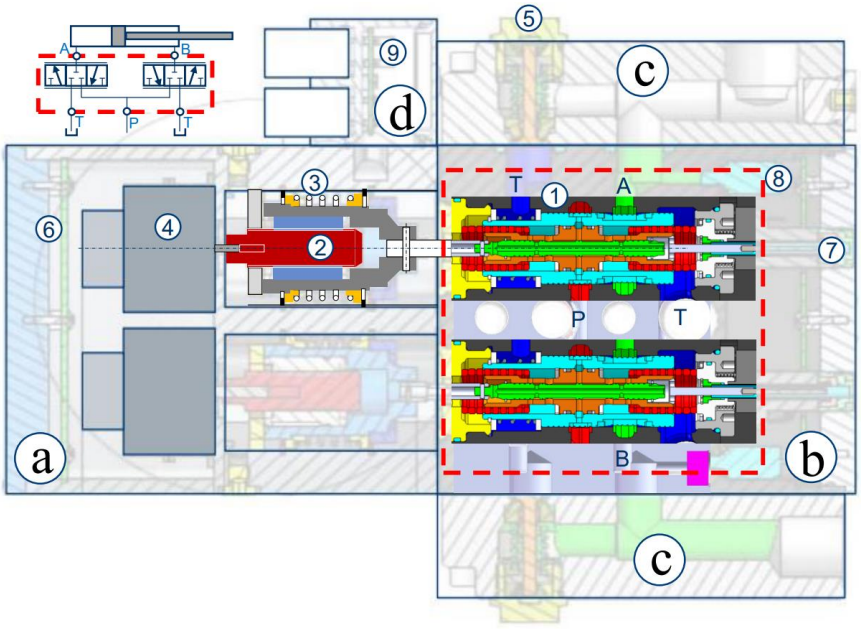


Figure 7: Overall view of the fluid power control unit with independent metering

The potential of the fluid power control unit presented in section 3 for the use in a hydraulic excavator drive have been investigated using system simulation. The fluid power control unit with separate control edges was compared in contrast to a classic hydro-mechanical throttle system with conventionally linked control edges. The fluid power control unit considered, consist of separate control edges, and electronically controlled regeneration modes. Both systems (the conventional system as well) are already equipped with demand-responsive load sensing pumps. The assessment of the system performance is carried out by means of the simulation of excavator implement drive systems. The modelling of the implement drive and the working cycle have been adapted to the practical work and usual requirements with the wheel excavator shown in **Figure 8**. The arrangement of the adjustable boom is fully extended and assumed to be rigid in the simulation. So, the simulation model contains two boom cylinders, an arm cylinder, a bucket cylinder, and a slew drive. All relevant inertia (and gravity) parameters have been determined and an additional bucket load of 800 kg has taken into account. A PI controller represents the operator of the excavator to drive a common working cycle shown in **Figure 9**.

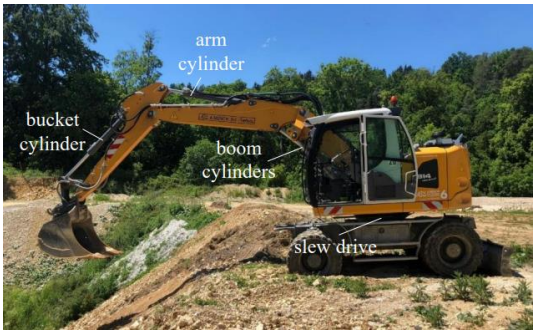


Figure 8: Excavator (Liebherr A 914 Compact Litronic) considered in the modelling

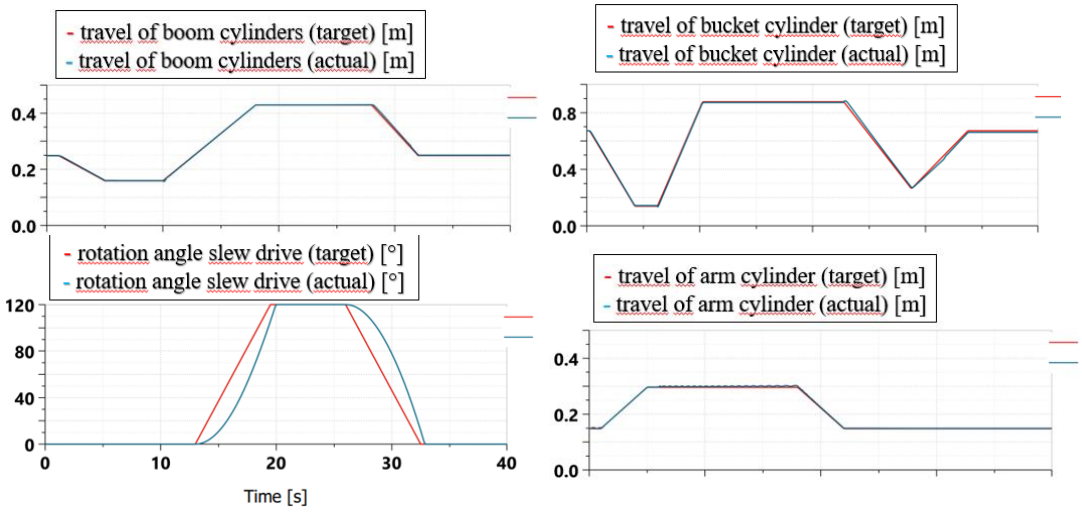


Figure 9: Working cycle of the implement drive used in the simulation

The following two system variants have been investigated by means of simulation:

1. Conventional excavator throttle system, and

2. Optimized excavator throttle system.

The conventional excavator throttle system shown in **Figure 10** consists of

- four 5/3 way proportional directional valves with linked control edges
- a customized load sensing pump with a combined pressure and flow compensator (the highest load pressure is transmitted via shuttle valves to the pump)
- a pressure relief valve to protect the hydraulic system from overload.

2. The optimized excavator throttle system shown in **Figure 11** consists of

- eight 3/3 way proportional directional valves with separate control edges
- eight flow sharing pressure compensators each with bypass and check valve
- three short-circuit valves for an energy-efficient lowering of the loads
- a customized load sensing pump with a combined pressure and flow compensator
- a pressure relief valve to protect the hydraulic system from overload
- pressure sensors for the measurement of the load pressures in the cylinder chambers as well as the pump pressure
- an Electronic Control Unit (ECU) here exemplary used for the control of the regeneration mode.

The energy consumption for the 40 s working cycle has been calculated in time steps of 0.1 s considering the simulation results of pump pressure and flow rate. The calculated energy consumption is for the conventional implement drive system 535 kWs and for the optimized implement drive system 429 kWs. **Figure 11** shows the comparison of the simulation results for both systems.

The decrease in energy consumption has dominant reasons in the general reduction of throttling losses, due to the independent control edges, and in high-pressure regeneration. Low-pressure regeneration using the short circuit valves has an additional effect.

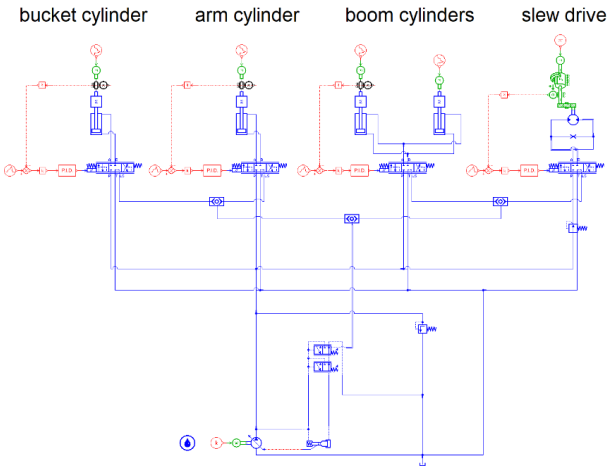


Figure 10: Conventional excavator throttle system with linked control edges (Simcenter Amesim)

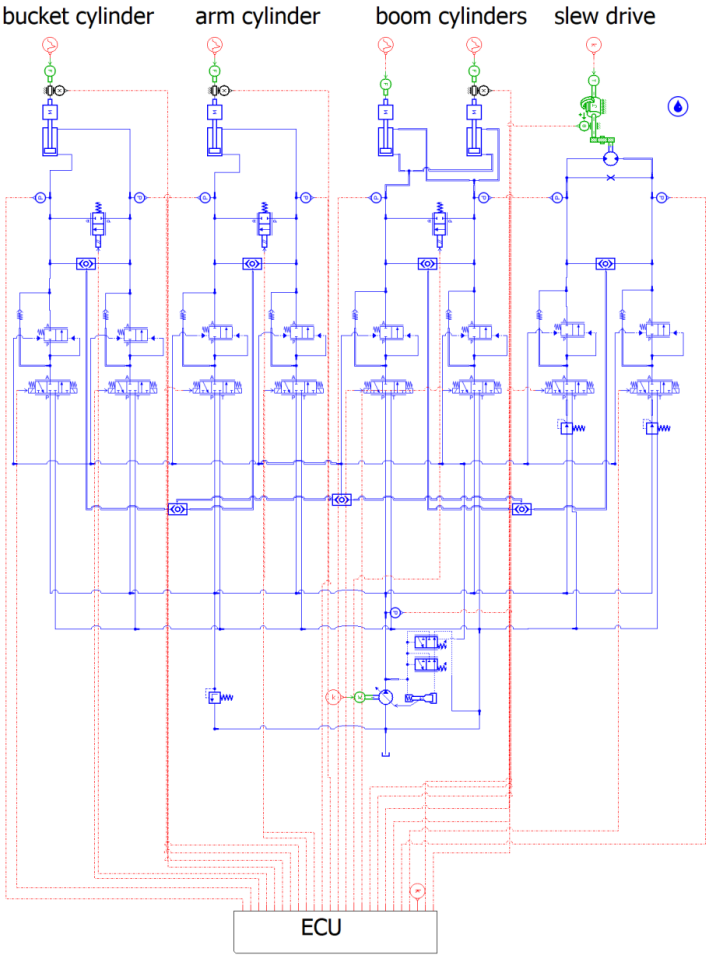


Figure 11: Optimized throttle system with independent metering combined with ECU controlled regeneration (Simcenter Amesim)

The optimized excavator implement drive system with separate control edges shows for the usual working cycle shown in **Figure 9** a good reduction of the energy consumption of 20% towards the conventional excavator hydro-mechanical implement drive system with linked control edges.

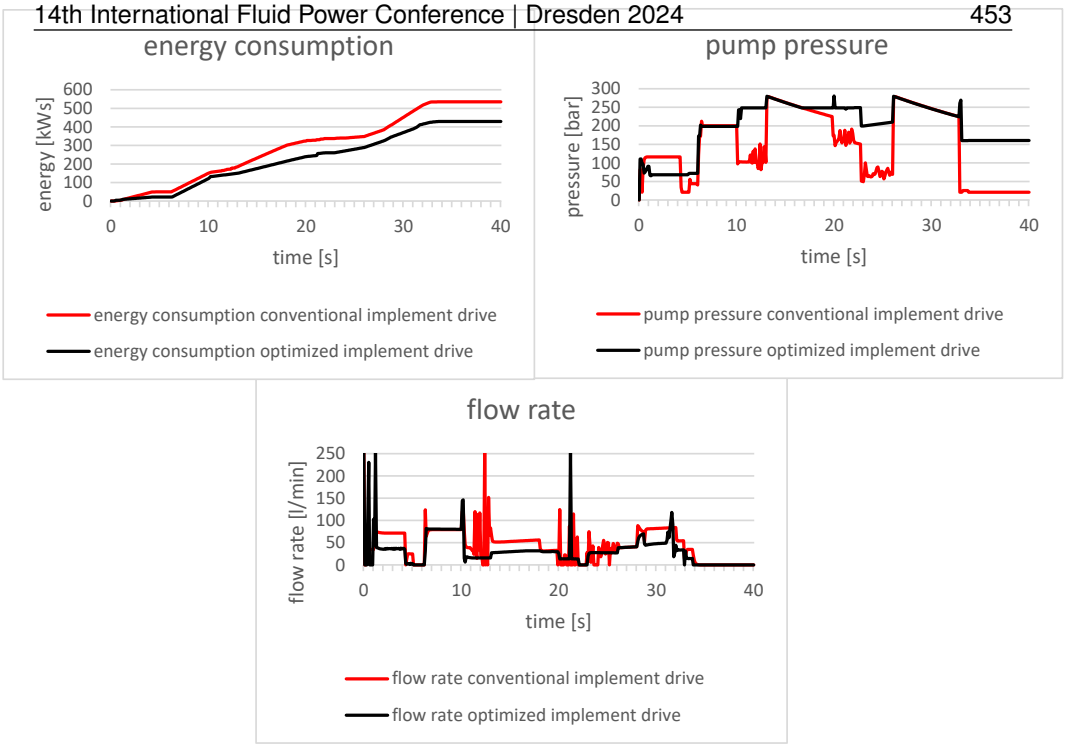


Figure 11: Simulation results of the conventional and the optimized implement drive

4. SUMMARY

The fluid control unit concept presented, has a compact pilot stage, with separate control edges, an actuator, consisting of a step motor with a ball bearing screw with minimal self-locking, pressure sensors and an ECU. The fluid power control unit is designed for high power requirements (maximum pressure >500 bar, flow rate 200 l/min), small installation space, well-defined position accuracy, stability and fail safe functionality (sections 2 and 3). By means of simulation the energy consumption of a conventional hydro-mechanic excavator implement drive system with connected control edges has been compared with an electrohydraulic excavator implement drive system with independent control edges. The comparison of the calculated results shows a good reduction of the energy consumption of 20% for the electrohydraulic solution with independent metering.

Nevertheless, higher system complexity induces higher downtime risks as well as higher investment costs. Rising fuel prices and further intensified emission standards will in future support efficient displacement control solutions. In case of resistance control solutions, separate control edges together with, sensors, electronic control and digitalization are attractive options.

On a higher system level, including implement and travel drive as well as combustion or electric engine, further digital optimization capability can be achieved, in terms of productivity, energy efficiency and emission reduction. Despite the higher investment cost and the remaining engineering challenges, further investigations are worthwhile to be considered.

Acknowledgement

This work was supported by the German Federal Ministry of Economic Affairs and Energy (BMWi) under Contract ZF4593501RP8.

REFERENCES

[1] Weber J, et al. (2016) Novel system architectures by individual drives. 10th International Fluid Power Conference, Dresden, Germany

[2] Vukovic M, Leifeld R, Murrenhoff H. (2016) STEAM – A hydraulic hybrid architecture for excavators. 10th International Fluid Power Conference, Dresden, Germany

[3] Rydberg K-E (2003) Fluid Power technology in distributed mechatronic systems. The Fourth International Symposium on Fluid Power Transmission and Control, Wuhan, China

[4] Linjama M, Vilenius M (2005) Energy-efficient motion control of a digital hydraulic joint actuator. Proceedings of the JFPS International Symposium on Fluid Power, Tsukuba, Japan

[5] Eriksson B, Larsson J, Palmberg J-O (2007) A novel valve concept including the valvistor poppet Valve. 10th Scandinavian International Conference on Fluid Power, SICFP´07, Tampere, Finland

[6] Sitte A, Beck B, Weber J (2014) Design of independent metering control systems. 9th International Fluid Power Conference, Aachen, Germany

[7] Lee J-C, Jin K-C, Kwon Y-M, et al. (2016) Development of the independent metering valve control system and analysis of its performance for an excavator. Bath/ASME 2016 Symposium on Fluid Power and motion control, Bath, UK

[8] Lübbert J, Weber J, Strauch C, Bruck P (2020) Modular independent metering system for mobile applications providing smooth mode transition. 12th International Fluid Power Conference, Dresden, Germany

[9] Fischer E, Weber J (2022) Data analysis for the evaluation and design of a model-based fault detection based on an independent metering system for mobile hydraulic drives. 13th International Fluid Power Conference, Aachen, Germany

[10] Isermann R (2007) Mechatronische Systeme. Springer, German Edition

[11] Herzog R, Zimmer T (2004) Control Valve. European Patent EP1426622A2

COMPARISON OF STRATEGIES FOR UNNOTICEABLE MODE SHIFTING IN MOBILE INDEPENDENT METERING SYSTEMS

peer reviewed

Jan Lübbert[1]*, Jürgen Weber[1], Peter Bruck[2]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*
[2] *Hydac Fluidtechnik GmbH, Industriegebiet D-66280 Sulzbach/Saar*

* Corresponding author: Tel.: +49 351 463-39932; E-mail address: jan_frederik.luebbert@tu-dresden.de

ABSTRACT

Independent Metering (IM) offers great potential to improve the energy efficiency of hydraulic systems. This is especially important in mobile applications due to the limited capacities of electric accumulators, which will probably become the primary energy source in many applications in future. One energy saving mechanism of IM are regenerative operation modes. In many applications, the load direction changes during an ongoing movement. In these cases, regenerative modes are feasible only if there is a way to shift unperceivably between the modes while moving. In this paper, four different mode-shifting techniques are described and compared on an excavator-arm test rig. These are continuous shifting with three active valves (CSA) or with two active valves and a passive path (CSP), and discrete shifting with a pressure compensator (DSMR) and without one (DS). It will be shown, that unnoticeable mode shifting is possible using serial production mobile hydraulic components – either continuously with a low-pressure regenerating valve layout comprising simple check-valves, or fast and discrete using a pressure compensator.

Keywords: mobile hydraulics, independent metering, mode shifting, operating behaviour

1. INTRODUCTION

Modern hydraulic systems for operator-controlled mobile machines have undergone a long developing history. The first fully-hydraulic excavators appeared on the market in the 1950s [1]. Developing efforts were focused on optimizing the control behaviour for a long time. The outcome of decades of system development and optimization are Open- and Closed-Centre-valve-controlled systems, with their respective specific operating behaviour. With increasing energy costs since the 1970s (first oil crisis) energy efficiency gained importance and Negative-Flow-Control- and Load-Sensing-systems emerged, being wide spread until today.

The current demand for elimination of fossil energy sources and combustion engines forces OEMs to shift to battery-electric drives in many application cases. For example in some cities it is already required by law to operate machine on municipal construction sites with electric drives, to reduce air pollution and the municipal carbon footprint [2]. Under the circumstances of very high battery costs, limited battery capacities and long charging times energy efficiency is not only a question of operating costs any more, but also of acquisition costs, operating time and productivity. This lifts the importance of energy efficiency on a higher level, which most state of the art solutions cannot meet anymore.

In respect to costs and installation space Independent Metering valve technology (IM) is a promising option in the range of medium-sized machines, while in large machines (i.e. mining) multi-pump-systems are common [3].

1.1. Energy saving mechanisms of IM

IM offers two mechanisms to save energy: Reduction of unnecessary throttling losses at the outlet throttling edges applies to consumers with altering load direction like an excavator's stick cylinder or swing drive. Conventional valves need small outlet openings to keep control of overrunning loads. These increase the required pressure for moving resistive loads. With IM the valve's opening ratio can be adapted to the current load situation thus preventing the mentioned energy loss. The second saving mechanism are regenerative operating modes, which reduce the pump volume flow. The goal is to use the mode, which requires the least pump flow. This optimization can be performed either on axis level [4] or on system level. [5] describes such a solution for a stationary application.

In applications with heavily varying load forces during ongoing movements regenerative modes can only be used, if it is possible to shift between these and normal operation while moving. This task is one of the major challenges of IM to overcome [6], because multiple valves and the pump are involved in a shifting event. Flow paths and pressure drops as well as volume flows change and all involved components must perform so well coordinated that the operator does not take notice of this complex event.

1.2. Mode shifting approaches

There are different ways to deal with the mode-shifting problem:

Prevent shifting by predefining modes for each consumer and moving direction [7]. For specific applications with foreseeable loads, like crane booms, this is sufficient, but in case of changing loads, i.e. in earthwork machinery, the method will lead to inefficient operating situations when the machine cannot react on the load change by shifting into the proper mode.

The downside of *discrete shifting* is the jerk induced into the movement. This is unavoidable, because valves and pump must react on a step signal, which exceeds their limited dynamics. In [8] and [9] suggestions for minimizing the jerk by selecting an optimal shifting moment are given. Apart from the dynamic limitations, the shifting action must not interrupt the flow path. This demand excludes some valve topologies from mode shifting.

Early extensive research on *continuous shifting* can be found in [10]. Shenouda used the Incova valve system, having a Wheatstone-Bridge layout, applied to a loader backhoe boom. The idea has been adapted to stationary applications in [5] and a different valve layout with a short circuit path in [11].

Speed-controlled regeneration is established in certain advanced conventional systems, but can also be found in some IM valves [12], [13]. In this case, a spool valve opens a regeneration path at large spool deflections, mostly from rod to piston side. The classic motivation is to obtain higher speeds with limited pump flow. This strategy is rather simple to implement but limits the usage of regeneration to high speeds, thus loosing much of the energy saving potential.

1.3. Mode shifting in serial production IM systems

State of the art IM valve systems available on the market are mostly not capable of shifting in an ongoing movement:

The CMA by Eaton [14] is able to regenerate but not to shift into regeneration while moving, because the shift would interrupt the flow path. The PVX valve by Danfoss [12] incorporates only a high-speed-regenerating edge (speed-controlles regeneration). A completely different approach has been implemented by Husco in their Incova valve [15]. This is a module comprising four 2/2-poppet-valves arranged in a Wheatstone-Bridge. This allows for normal mode, as well as high- and low-pressure-regeneration and continuous shifting between these modes. The module is intended for

decentralized mounting and connection with a P and T bus hose line. Probably this novel layout, which is incompatible to traditional machine architectures, and the anticipated high price led to its failure on the market. The latest public sources date back to 2008. To the knowledge of the authors, apart from Incova there is or has been no other IM system on the market, which allows for mode shifting at any consumer velocity.

Mode shifting enhances the application spectrum of regenerative modes and thus energy efficiency of IM systems a lot, making IM much more attractive. Apparently, there is no satisfying solution for the mode-shifting problem yet. This paper will show approaches to shift between modes without disturbing the ongoing movement in a noticeable way, using common mobile hydraulic components.

2. OPERATING MODES

The algorithms in this paper use distinct modes and continuously variable modes (CVM). In a distinct mode, the volume flows in the flow paths only depend on the work port flows. Dividing or merging of flows may occur, but with fixed ratios depending on the consumer's geometry or maximal valve openings. A CVM is characterized by a variable dividing or merging ratio at at least one junction.

The following modes will be used in the experiments in this paper:

Normal extension NM+ (Figure 1 left). In normal extension, the pump delivers the volume flow Q_S to the cylinder's piston chamber P while the rod chamber R is drained to tank (Q_R). This is a distinct mode, meaning that there is no option to divide or merge volume flow in a definable ratio.

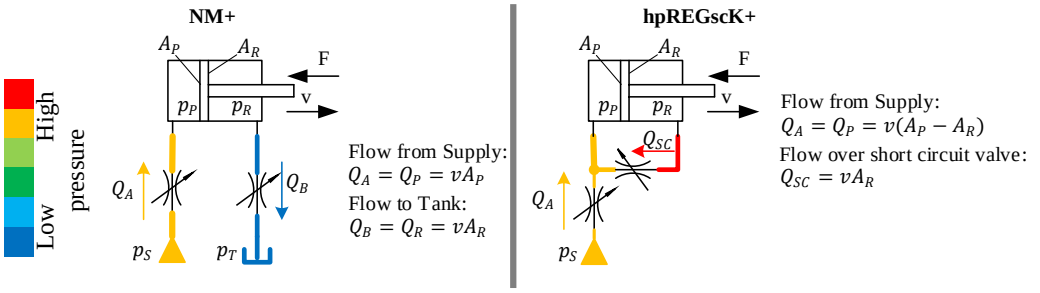


Figure 1: Normal extension NM+ (left) and high-pressure-regeneration hpREGscK+ (right). The thickness of the lines indicates the volume flow, the colour the pressure level.

High-pressure regeneration extension with junction on piston side (hpREGscK+, Figure 1 right**)**: In this distinct mode, a short-circuit-valve SC feeds the drain flow from the rod side back into the piston chamber. The regeneration flow merges with the pump flow on the piston side. The rod side pressure must be higher than the piston side pressure. The pump delivers only the differential flow. This mode is used for low loads, when the pump pressure is much higher than required for normal extension, and overrunning loads.

Continuously variable mode (CVM) NMhpREGscK+ (Figure 2 left): This is a mixture between NM+ and hpREGscK+ [11]. The return flow from the rod side divides into a share Q_B drained into the tank and a second share Q_{SC} regenerated to the piston side. A shifting variable $\alpha = 0..1$ defines the dividing ratio. The edge case $\alpha = 0$ is equivalent to distinct normal extension NM+, while $\alpha = 1$ equals distinct high-pressure regeneration hpREGscK+. The mode is intended for shifting between NM+ and hpREGscK+ over a short time span.

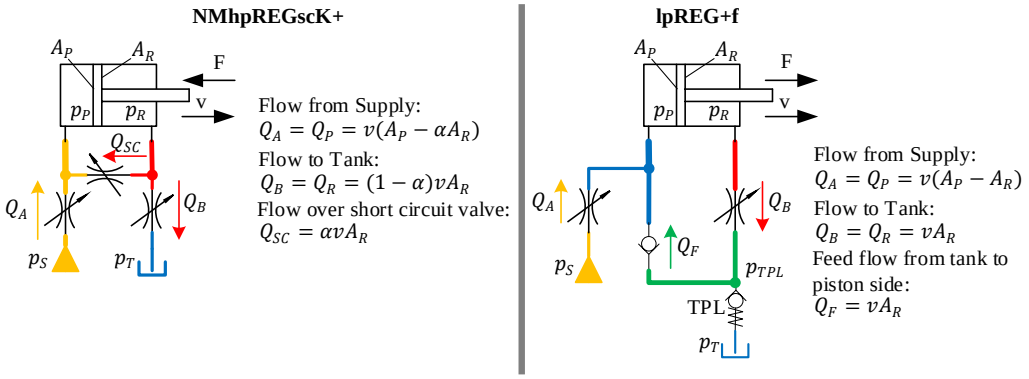


Figure 2: Extension in continuously variable mode normal extension to high-pressure short circuit regeneration NMhpREGscK+ (left) and in low-pressure regeneration lpREG+f (right), requiring active load and a pressurized tank line.

Low-pressure regeneration extension lpREG+f (Figure 2 right**)**: In this mode, the rod side flow Q_B is drained to tank while the piston side P is being fed partially from the tank line (Q_F) and partially from the pump (Q_A) [5]. The mode is feasible only for aiding loads. In this case, the load F defines the rod side pressure p_R, while the piston side pressure p_P is on a constant low level below the tank line pressure p_{TPL}. To enable regeneration from tank line to the cylinder, the tank line must be pressurized, i.e. by a spring-loaded check valve TPL (*t*ank *pre*load valve). The rod side flow regenerates to the piston side completely, making this mode a distinct mode. The rod side valve B controls the velocity while the piston side valve A delivers exactly the required differential flow.

CVM NMlpREG+f (Figure 3): This mode, being a mixture between NM+ and lpREG+f, enables continuous shifting between NM+ and lpREG+f. In contrast to lpREG+f, here the rod side flow Q_B partially drains into tank instead of completely regenerating into the piston chamber. While valve B still controls the velocity, valve A now delivers more than the differential flow $v(A_P - A_R)$ and thus controls the merging ratio α between feed flow Q_A from pump and regenerating flow Q_F from the tank line, making this mode a CVM. $\alpha = 1$ is equivalent to the distinct mode lpREG+f, whereas $\alpha = 0$ equals NM+.

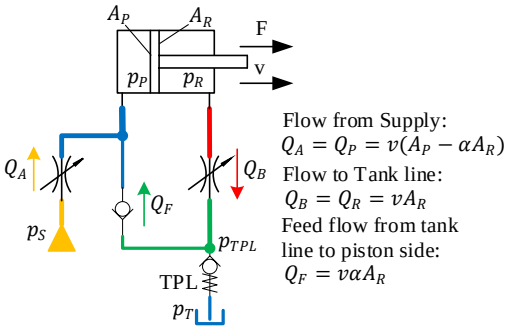


Figure 3: Extension in continuously variable mode normal mode to low pressure regeneration NMlpREG+f.

3. MODE SHIFTING METHODS

This section explains the different mode shifting methods, which will be evaluated in experiment. Independent from the mode shifting method itself, the shifting threshold defines the circumstances, under which a mode shift shall occur.

3.1. Definition of the mode shifting threshold

There are different approaches to define a mode-shifting threshold. Shenouda describes a method based on the power $P = F_L v_{Des}$, which results from the current load force F_L and desired velocity v_{Des} [10]. His aim is to use maximum acceleration at low speed and the additional speed capability of regenerative modes in low load conditions. In this paper the actuator shall never move faster than the limits of normal mode because otherwise a fast moving low loaded cylinder will shift down from regeneration to normal mode when the load force rises, causing a sudden slowdown. A manual operator would probably not accept this kind of behaviour.

Here, solely the load force F_L defines the shifting threshold. Figure 4 shows the force-controlled shifting schedule at the example of the transition between normal extension NM+ and high-pressure regeneration hpREGscK+ via the CVM NMhpREGscK+.

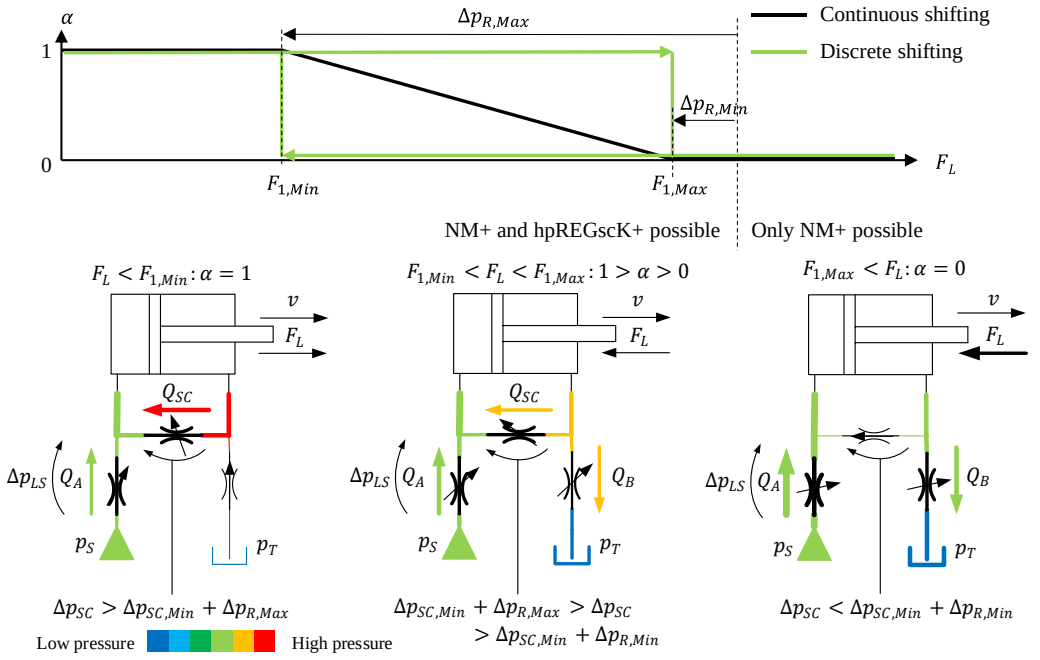


Figure 4: Definition of shifting threshold and CVM-variable α at the example of the transition between high-pressure-short-circuit-regeneration hpREGscK+ and normal extension NM+.

To make an operation mode feasible, certain pressure drops must exist over the corresponding flow paths. In the case of NMhpREGscK+ it is the pressure drop $\Delta p_{SC,Min}$ over the short-circuit-path (bottom part of the figure). If $\Delta p_{SC} > \Delta p_{SC,Min}$, NMhpREGscK+ is feasible, otherwise only NM+. An offset Δp_R added to $\Delta p_{SC,Min}$ creates a margin: This way Δp_{SC} does not fall below $\Delta p_{SC,Min}$ if Δp_{SC} drops quickly due to a sudden load change and the controller triggers the backshift into NM+

when $\Delta p_{SC} < \Delta p_{SC,Min} + \Delta p_R$. Defining a margin interval $\Delta p_{R,Min}..\Delta p_{R,Max}$ creates an interval between two load force values:

$$F_{1,Min} = (p_S - \Delta p_{LS})(A_P - A_R) - \left(\Delta p_{SC,Min} + \Delta p_{R,Max}\right)A_R \quad (1)$$

$$F_{1,Max} = (p_S - \Delta p_{LS})(A_P - A_R) - \left(\Delta p_{SC,Min} + \Delta p_{R,Min}\right)A_R \quad (2)$$

The continuous shifting methods use a linear interpolation of the shifting variable α between the two force values while the discrete methods use a hysteresis loop.

3.2. Discrete Shifting (DS)

A discrete shift is an instantaneous transition from one distinct mode to another. There are different ways to perform a discrete shift:

The volume flow is being passed from one proportional valve to another one. This causes a sudden, step-like change of the desired *flow* in the participating valves, while the pressure drops over the valves remain constant. Thus, the valves must change their opening instantly.

If the valve layout consists of a series arrangement of switch/directional and metering valves, the directional valves perform the mode shift, causing the pressure drops over the metering valves to change. In this case, the flows through the metering valves remain constant but here, the metering valves must adapt their openings to the new *pressure drops* to maintain the desired constant flow rates.

In both cases with the mode shift, the required pump volume flow changes instantaneously. The limited dynamics of valves and pump induce disturbances and jerks into the movement. Some valve systems like Eaton's CMA interrupt the flow path while shifting. The resulting pressure peaks intensify the disturbance.

3.3. Discrete Shifting with Pressure Compensator (DSMR)

If, during a discrete shifting event, only the pressure drops over the corresponding metering valves change, but not the desired volume flow (second case in the last section), it is possible to compensate for the sudden pressure drop change with an individual pressure compensator (IPC), eliminating the need for the metering valve to change its opening. An IPC reacts much faster than a mobile proportional valve with its relatively sluggish magnet.

Figure 5 shows a simplified example layout performing a compensated discrete switch from hpREG+ to NM+, induced by a changing load force. In the left part of the figure, the pump delivers flow through metering valve A into the piston chamber of the extending cylinder. Metering valve B drains the rod side, in series with the IPC, which controls the pressure drop over B. A switching valve SV compares the pressure behind B with the pump pressure. An aiding load raises the drain pressure to a high level, causing SV to close and divert the drain flow over the check valve CV into the pump line – high-pressure regeneration hpREG+. The pump delivers the required flow, so the unloading valve UV remains closed.

As the load force changes (central figure), the drain pressure decreases and the control device anticipates a shift into NM+. To prevent undersupply due to the relatively slow dynamics of the pump right after the quick shift, the controller increases the pump flow shortly in advance. The unloading valve UV opens and drains the excessive pump flow to tank, limiting the pump pressure to inlet pressure plus LS margin.

Between the central and right figure the mode shift takes place: A further drop of the rod-side chamber pressure causes SV to open. Now the pressure at the outlet of the IPC drops from pump to tank level. The IPC closes quickly and keeps pressure drop and volume flow over metering valve B constant. Since there is no volume flow regenerated any more, the pump flow matches the inlet flow through A again and the unloading valve UV closes.

Neither had one of the actively controlled components (A, B, pump) to perform a step nor have the volume flows at the cylinder ports been disturbed.

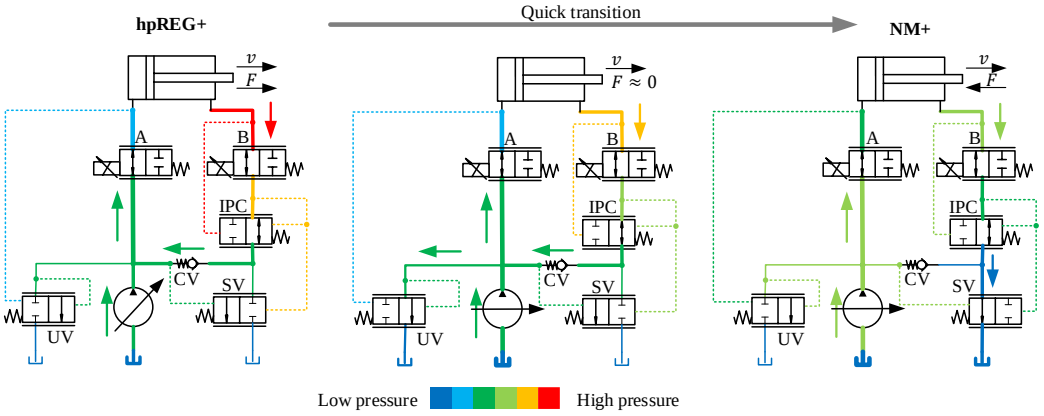


Figure 5: Discrete shifting DSMR with pressure compensator at a simplified example valve layout, in this case triggered by a load change. In this layout, also a changing pump pressure can trigger such a mode shift.

3.4. Continuous Shifting with three active Valves (CSA)

This strategy uses three active valves while shifting, similar to the approach in [10]. In at least one path, inlet or outlet, flow is divided or merged; see section 2 and Figure 4 bottom. The dividing/merging ratio is variable. In this condition the IM valve system as a multivariable system has three outputs: Velocity v, pressure level p and mode shifting variable α. Hence, always three valves are necessary to control these three variables independently.

Figure 4 bottom shows an example of a CSA shift: An extending movement of the cylinder may start in hpREGscK+ with an aiding load – bottom left. When the load force raises above the limit $F_{1,Min}$, the continuous shift begins. The shifting variable α drops below 1, and, according to the equations in Figure 2 right, a portion Q_B of the outlet flow from the rod side is drained into tank. At the same time, the regenerating flow Q_{SC} decreases and the pump flow Q_A increases.

With increasing load force, α drops further and this portion increases until the regenerating flow Q_{SC} becomes zero and the drive operates in normal operation NM+. If the shifting interval between $F_{1,Min}$ and $F_{1,Max}$ is defined large enough, the shifting process is sufficiently long for valves and pump to follow the changing set values with their limited dynamics.

During the shifting phase the outlet flow of the cylinder is the sum of both Q_B and Q_{SC}, meaning that the errors of *both* valves contribute to the velocity error of the cylinder. So it is anticipated, that some velocity deviations may occur during shifting. Since the volume flow is controlled by different valves before and after the shift, a constant velocity change due to static valve errors is expected.

3.5. Continuous Shifting with two active Valves (CSP)

Drawbacks of CSA are the need to have and control three proportional valves and the summation of flow control errors during the shifting phase. Furthermore, in CSA in general and in some DS(MR) shifting methods, depending on the valve layout, before and after the switch, different valves control the consumer's velocity. This may lead to permanent velocity deviations due to the different flow accuracy errors of the valves. These effects might be noticeable for the operator.

To prevent the aforementioned issues, a novel method shall be developed, in which no summation errors occur and that uses the same valve to control the velocity before and after the shift.

Figure 6 shows a continuous shifting between extension in low-pressure-regeneration (lpREG+f) and normal mode (NM+).

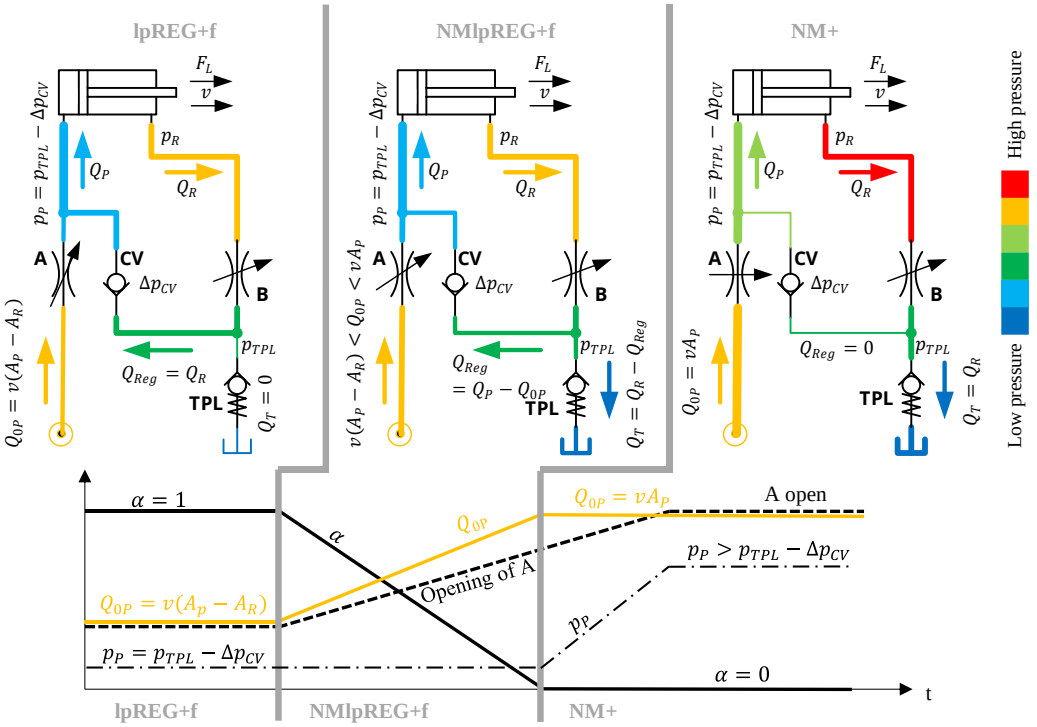


Figure 6: Shifting from low-pressure-regeneration lpREG+f to normal extension NM+ using CSP in a simplified example valve layout.

In one flow path, in this case the inlet on the piston side, there is a junction, at which and actively controlled flow is merged with a flow from a third path reacting passively on the actively controlled flow. Here the actively controlled flow is Q_{0P} through valve A, and the passive flow is the regeneration flow Q_{Reg} through the check valve CV in the centre of the setup. Together with the tank preload valve TPL in the bottom right the regeneration check valve ensures a minimal pressure $p_{TPL} - \Delta p_{CV}$ in the rod chamber.

In lpREG+f (Figure 6 left) valve B controls the velocity while valve A is in control of the operation mode. If A is partially opened so that Q_{0P} exactly matches the differential flow $v(A_P - A_R)$, the consumers operates in distinct lpREG+f. The rod side outlet flow regenerates completely into the piston path while the pump delivers exactly the required flow to fill the piston chamber (see also

Figure 2 right). Once valve A opens further, the transition to NM+ begins by gradually shifting the rod side outlet flow from the regeneration path to the tank path with the further opening of A. In this situation, still B is in control of the velocity, while A controls the mode shift. Once Q_{Reg} has become zero (Figure 6 right), normal mode NM+ is reached and the valves A and B control velocity and pressure together with the common cross-interactions between these two variables.

It should be noted, that in lpREG+f only one active valve is in control and during the shifting phase NMlpREG+f two active valves are used, meaning that in lpREG+f only the velocity is controllable and in NMlpREG+f additionally the shifting variable α. In both cases, it is not possible to control the pressure level, which results from the valve layout and load situation instead. Since valve B alone determines the velocity over the whole shifting process, valve precision errors will not lead to velocity disturbances during shifting.

4. TEST RIG SETUP

The test rig comprises multiple IM valve systems, two different excavator arms and an electronic control unit.

A TT Control TTC580, programmed in Codesys V3, serves as ECU. The control algorithms are sampled with 20 ms. A PC serves for programming, administration and measurement acquisition via Ethernet. CAN-Joysticks allow for manual control of the three available axes boom, stick and bucket. To obtain reproducible results, it is possible to perform automated movements, which is used for the measurements in this paper.

The “big implement” is a 2:1-model of a medium-sized excavator arm with two conventional supporting boom cylinders and an additional mass of 220 kg mounted at the bucket. This implement will be used for the shifting methods DS, CSA and CSP. An alternative “small implement” from a mini-excavator will be used for the method DSMR. This implement uses a top-mounted boom cylinder, which extends while lowering instead of retracting.

An electro-proportional variable-displacement-pump supplies the test rig using electronic Load-Sensing with flow feed-forward, according to [1].

Discrete shifting DS is tested with an IM valve system “MvS”, which consists of one piloted proportional 2/2-spool valve at each work port and a set of four 2/2-poppet-switch-valves to connect each proportional valve either to pressure or to tank, thus enabling high and low pressure regeneration besides normal mode. An upstream pressure compensator decouples the inlet flow from the pump pressure. For details see [16].

A second IM valve system “MvSSC” performs the CSA and DSMR shifting methods. This system comprises an electro-proportional three-way pressure control valve at each work port to connect the port either to supply or to tank; the latter in series with a 2/2-proportional poppet valve and an outlet pressure compensator. A pressure-controlled 2/2-switch valve and a check valve enable high-pressure regeneration, compare to Figure 5. Additionally the systems contains a 2/2-proportional poppet valve in the short circuit path. Details are described in [11].

A third IM valve system “MvSND” serves for the CSP shifting method test (Figure 7). Two pressure-compensated 2/2-proportional spool valves (SP and SR) meter the pump inlet flow, while the outlet flow runs through one 2/2-proportional poppet valve at each work port (PT and RT). A tank preload check valve TPL in the tank drain line and a feed check valve at each work port (FP and FR) enable low-pressure regeneration.

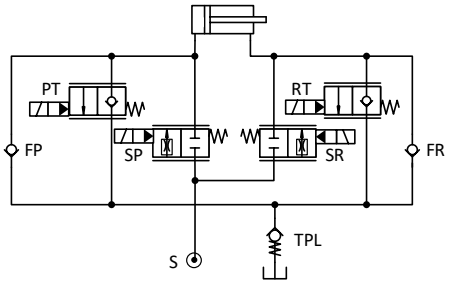


Figure 7: IM valve layout "MvSND" for test of switching strategy CSP.

5. COMPARISON IN EXPERIMENT

An excavator's stick cylinder has the highest demand for smooth mode shifting: During operation of the stick in air, without ground contact, a continuous transition from aiding to resistive load occurs, when the stick approaches its lowest position and starts lifting from that point on. The operator expects a very smooth movement. In contrast, the bucket cylinder often operates for digging, where harsh and unpredictable load forces disturb the movement, but are expected and accepted by the operator. Besides, due to the mainly passive load, the bucket drive has almost no chance to use regenerative modes. The boom drive mostly operates against/with a force in constant direction and therefore does not shift between modes while moving. Hence, the shifting methods are tested at the stick cylinder.

The experiments to evaluate the methods DS, CSA and CSP are performed at the "big implement" with 220kg additional mass at its bucket. A constant pressure supply instead of Load Sensing excludes any influence of the pump control algorithm on the mode shifting performance. The boom rests in a relatively level position while the stick cylinder is retracted completely, meaning the implement is stretched horizontally (see faint sketch in Figure 8 central bottom). From this starting position, the stick cylinder extends with a constant set velocity of 100 mm/s. The load force during the resulting movement is aiding at the beginning, slowly increasing to a resistive value. Thus, the valve system can begin the movement in regeneration but needs to shift to normal mode later on.

The central graph in Figure 8 shows the desired stick cylinder velocity and its measured velocities for the different shifting methods. As a reference, the result of a conventional LUDV valve is displayed. The upper diagram shows the measured cylinder chamber pressures (P = piston side, R = rod side), while the bottom frame displays the shifting variable α for the IM systems. At the beginning the rod side pressure is high while the piston side pressure is low, indicating the aiding load. As the load changes, the pressures approach each other, triggering the mode shift.

The LUDV system, supplied by electronic LS, delivers a smooth movement, as can be seen in velocity and pressures. It is constantly too slow due to some inaccuracy in the used valve opening look up table.

The continuous shift of the valve system "MvSSC" from hpREG+ to NM+ via CSA begins at 3.5 s, where α starts to decrease to 0 smoothly. During the shifting process, some minor pressure fluctuations and a small velocity deviation occur, but these are not subjectively perceivable.

At around 3.4 s the shifting process CSP begins. In this case, there is no low-pass-filter for the shifting variable α, since it has shown to be not necessary for smooth and oscillation-free performance of the axis. For this reason, α shows some noise in the transition phase. Velocities and pressures are almost constant, making the shift unperceivable.

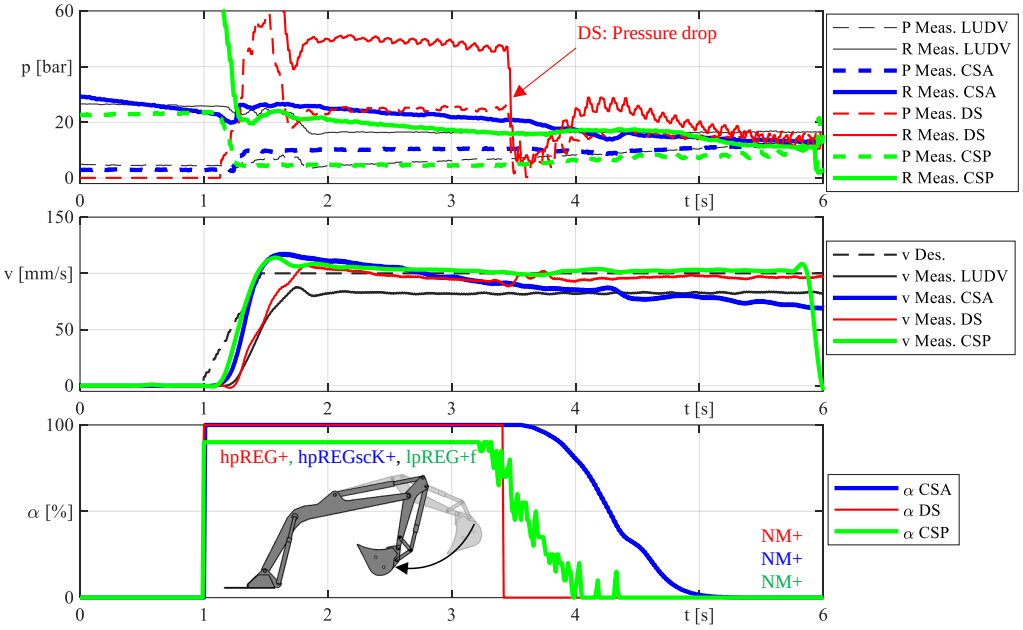


Figure 8: Comparison of shifting strategies CSA, CSP and DS.

At 3.5 s in the valve system “MvS” the discrete shift DS from hpREG+ to NM+ is triggered. This causes an abrupt pressure drop in the cylinder due to the sudden pressure drop from supply to tank level at the outlet of the rod-side metering valve. To compensate for this, the valve needs to decrease its opening in a very fast step. The limited valve dynamics cause an unintended peak in the outlet flow and velocity during the shifting phase leading to a well perceivable pressure drop and jerk.

One should notice that the shifting threshold slightly varies between the valve systems due to the different valve layouts and pressure drops across their flow paths.

The DSMR shifting method is tested with the valve system “MvSSC” with its outlet pressure compensators. The required switch valve SV and check valve CV (see Figure 5) are only available in the boom valve system. To test the DSMR shifting method between hpREG+ and NM+ with this valve setup, an aiding extension movement is required. Lowering the boom of the “small implement” with 180 kg additional mass at the bucket provides this required movement. This load results in a rod side pressure of approx. 80 bar, see the upper diagram in Figure 9. The movement is manually controlled with a constant maximal cylinder velocity of 75 mm/s. The pump pressure remains constant at 50 bar. This is achieved by delivering excess pump flow and bleeding it off over an electro-proportional pressure relief valve. With the pump pressure being lower than the rod pressure, the regeneration switch valve SV (see Figure 5) is closed and the rod-side outlet flow regenerates into the pump line, hpREG+.

At 4.8 s a manual signal triggers a step of the pump pressure to 130 bar, exceeding the rod pressure of 80 bar. This causes the valve system to shift into normal mode NM+ instantly by opening SV, causing the pressure behind the outlet pressure compensator to drop from 50 bar to tank level. The pressure compensator quickly regulates the pressure drop over the rod-side outlet valve, so that only a very short and little pressure drop is visible in the measurement, which is neither measurable in the velocity nor subjectively perceivable.

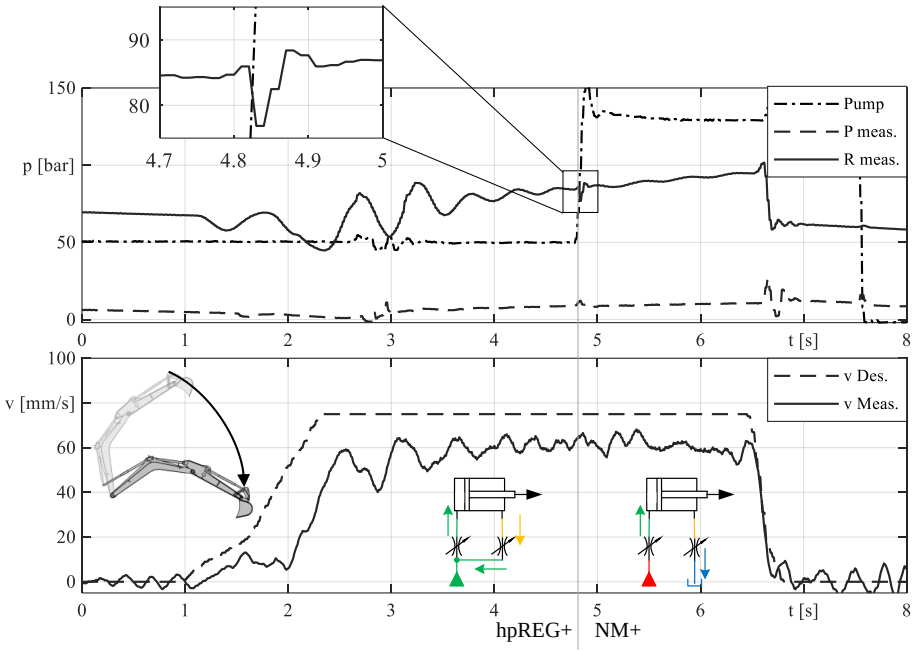


Figure 9: Result for shifting strategy DSMR.

6. SUMMARY AND OUTLOOK

Both continuous shifting with three active valves (CSA) and continuous shifting with a passive path (CSP) perform smooth, unnoticeable mode shifts. Compared to CSA, CSP is easier to implement, because one valve less needs to be controlled. Furthermore, low-pressure regeneration, as used in this paper, is very easy to implement by re-dimensioning the secondary-relief/anti-cavitation-valves, which must be integrated into most valve layouts for safety reasons anyway. A drawback of both continuous shifting methods is the relatively long shifting time, shortening the timespan in which regeneration is actually used.

Discrete shifting with pressure compensator DSMR overcomes this downside. It offers a very fast and unperceivable shift. The price is the pressure compensator as an additional component and a more complex valve layout. Apart from that, the pressure compensator reduces the drive's damping ratio compared to a valve layout without load compensation [17]. A suitable dynamic strategy for operating an electronically controlled pump together with an unloading valve may eliminate possible dynamic under- or oversupply situations, which result from the sudden flow demand changes the discrete shifts induce.

Discrete shifting without additional measures (DS) has been tested for reference purposes in this study as well, although it is evident, that this strategy is not recommendable because of the unavoidable jerks.

To the authors CSP appears as the preferred solution because of its good operating behaviour and low complexity. To improve energy efficiency, one might add a central switch valve to the control block to bypass the tank preload required for low-pressure regeneration, when none of the active consumers regenerates.

ACKNOWLEDGEMENT

The presented research activities are part of the project “LiftControl” (Ref. No. KK5023208RF1), which was funded by the German Federal Ministry for Economic Affairs and Climate Action within the ZIM programme.

Supported by:



on the basis of a decision
by the German Bundestag

NOMENCLATURE

A	Area	*F*	Feed
α	Shifting variable for continuously variable modes	*IM*	Independent Metering
Δp	Pressure drop	*IPC*	Individual pressure compensator
F	Force	*L*	Load
m	Mass	*LUDV*	“Lastunabhängige Durchflussverteilung” = conventional flow sharing valve system
Q	Volume flow	*P*	Piston side
v	Velocity	*R*	Rod side
		Reg	Regeneration
CSA	Continuous shifting with 3 active valves	*S*	Supply
CSP	Continuous shifting with a passive valve	*SC*	Short Circuit
CV	Check Valve	*T*	Tank
CVM	Continuously variable mode	*UV*	Unloading valve
DS	Discrete Shifting	*TPL*	Tank pre-load (valve)
DSMR	Discrete Shifting with mechanic regulator (pressure compensator)		

REFERENCES

1. Djurovic M (2007) Energiesparende Antriebssysteme für die Arbeitshydraulik mobiler Arbeitsmaschinen – „Elektrohydraulisches Flow Matching“. Dissertation, TU Dresden

2. Lee C (2021) “Grüne Baustelle” dank Batterie und Elektro-Bagger? https://www.dw.com/de/bauen-baustelle-nachhaltig-emissionsarm-elektro-bagger-baumaschinen-bausektor-zement-co2-klimaschutz/a-57250263. Accessed 2 Dec 2022

3. Weber J, Beck B, Fischer E, et al (2016) Novel System Architectures by Individual Drives. In: 10th International Fluid Power Conference, IFK Conference proceedings. Dresden, Germany

4. Eriksson B (2010) Mobile Fluid Power Systems Design. PhD thesis, University of Linköping, Sweden

5. Kolks G, Weber J (2016) Modiciency - Efficient industrial hydraulic drives through independent metering using optimal operating modes. In: Proceedings of the 10. International Fluid Power Conference. Dresden

6. Karem Abuowda, Ivan Okhotnikov, Siamak Noroozi, et al (2019) A review of electrohydraulic independent metering technology. ISA Transactions

7. Lee J-C, Jin K-C, Kwon Y-M, et al (2016) Development of the Independent Metering Valve Control System and Analysis of its Performance for an Excavator. In: Proceedings of the BATH/ASME 2016 Symposium on Fluid Power and Motion Control. Bath, UK

8. Ding R, Xu B, Zhang J, Cheng M (2016) Bumpless mode switch of independent metering fluid power system for mobile machinery. Automation in Construction 68:52–64

9. Xu B, Ding R, Zhang J, et al (2015) Pump/valves coordinate control of the independent metering system for mobile machinery. Automation in Construction 57:98–111

10. Shenouda A (2006) Quasi-Static Hydraulic Control Systems and Energy Savings Potential Using Independent Metering Four-Valve Assembly Configuration. Dissertation, Woodruff School of Mechanical Engineering, Georgia Institute of Technology

11. Lübbert J, Stauch C, Bruck P, Weber J (2020) Modular Independent Metering System for mobile Applications providing smooth Mode Transition. In: Proceedings of the 12th International Fluid Power Conference. Dresden

12. (2016) Proportional Valve PVX (product data sheet). https://assets.danfoss.com/documents/DOC152986483080/DOC152986483080.pdf. Accessed 23 Mar 2023

13. (2016) K220LS EcoFormance. https://www.parker.com/literature/Mobile%20Controls%20-%20Europe/Literature%20Files%20MCDE/HY17-8571-B1-UK_K220LS_Ecoformance.pdf. Accessed 7 Dec 2023

14. EATON (2016) CMA200 Advanced Independent-Metering Mobile Valve. Product data sheet by EATON Company. https://www.eaton.com/ecm/groups/public/@pub/@eaton/@hyd/documents/content/pct_1451917.pdf. Accessed 11 Feb 2019

15. Husco (2007) INCOVA® Hydraulic Control System for Excavators (Product data sheet by Husco). http://hydrodream.ru/upload/iblock/e56/e5619a3a7f4f1aed0b1be538c1b0e743.pdf. Accessed 3 May 2019

16. Lübbert J, Sitte A, Beck B, Weber J (2016) Load-Force adaptive Outlet Throttling - An easily commissionable Independent Metering Control Strategy. In: Proceedings of the 2016 Symposium on Fluid Power and Motion Control. Bath (UK)

17. Andersson R (1997) Valves Contribution to System Damping. In: 5th Scandinavian International Conference on Fluid Power. Linköping (S)

FAULT LOCALIZATION FOR INDEPENDENT METERING SYSTEMS BY MODEL-BASED FAULT DETECTION

peer reviewed

Eric Fischer[1]*, Benjamin Beck[1], Jürgen Weber[1]

[1]*Institute of Mechatronic Engineering, Chair of Fluid Mechatronic Systems (Fluidtronics), Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*

* Corresponding author: Tel.: +49 351 463-33553; E-mail address: eric.fischer1@tu-dresden.de

ABSTRACT

This paper deals with fault localization for a mobile hydraulic system with independent metering, but it is also applicable to other hydraulic systems. The basis for this contribution is a model-based fault detection, which generates multiple symptom patterns for various component failures using a set of parity equations (see previous publications from the authors). The symptom patterns are evaluated through different classification methods such as geometric, statistical, and artificial intelligence methods. The evaluation focuses on simple approaches that are applicable in practice. Taking into account the limitations of mobile systems, different operating points, and a variety of fault scenarios, a correct fault localization of up to 92% of detected faults is possible. After locating faults correctly, IM systems enable a range of reconfiguration modes to keep the machine's functionality and therefore rise the availability. Laboratory demonstrator tests confirm the simulation-based outcomes.

Keywords: fault localization, model-based fault detection, independent metering, safety, reliability

1. INTRODUCTION

Systems with independent control edges offer high energy efficiency and great structural flexibility. To fully exploit these benefits, electrohydraulic drive systems and sensor technology are required. Despite the ongoing digitalization and automation of machines, there are only a few market-ready implementations of independent metering (IM) systems such as [1] and [2]. Possible reasons for this are the increased system complexity of IM systems in combination with extensive control algorithms, software architectures and safety aspects.

Previous work shows that IM systems are capable of meeting high safety requirements if the required diagnostic coverage is provided [3]. One possibility to achieve the required diagnostic coverage is a model-based fault detection, using parity equations. The implementation of such a reliable fault detection was shown in [4]. The increased system complexity creates the opportunity to use a set of eleven different parity equations with individual boundaries. These customized balance limits lead to varied sensitivity of the parity equations under diverse fault conditions. The resulting symptom pattern provides the basis of the investigated fault localization presented here.

Starting with the introduction of the object of study, basic approaches for fault detection are derived and their practical validation is presented in this paper. While maintaining knowledge-based (model-based) approaches, different methods for fault localization are introduced and implemented. The accuracy of their localization is compared. Test bench measurements confirm the effectiveness of the localization algorithms. The well-functioning fault localization allows partial compensation of faulty components through various reconfiguration modes - structural measures or control interventions to provide defined machine movement in case of failure - thus increasing availability.

2. SUBJECT OF INVESTIGATION

Subject of investigation is the independent metering (IM) system of the excavator test rig at the research centre. The system includes two IM sections, each consisting of two pilot-operated 2/2 proportional valves (see **Figure 1**, $V_{prop,A}$, $V_{prop,B}$) and four 2/2 switching valves ($V_{sv,AP}$, $V_{sv,AT}$, $V_{sv,BP}$, $V_{sv,BT}$). These IM sections are responsible for controlling the boom and stick cylinder of a compact excavator implement. The proportional valves regulate the flow rate and cylinder pressure according to the methods described in [5], while the switching valves determine the operation mode of the IM structure. The test rig enables extensive sensory monitoring. Sensors measure the cylinder velocities, the spool positions of the proportional valves and all pressures in the cylinder chambers and between the valves.

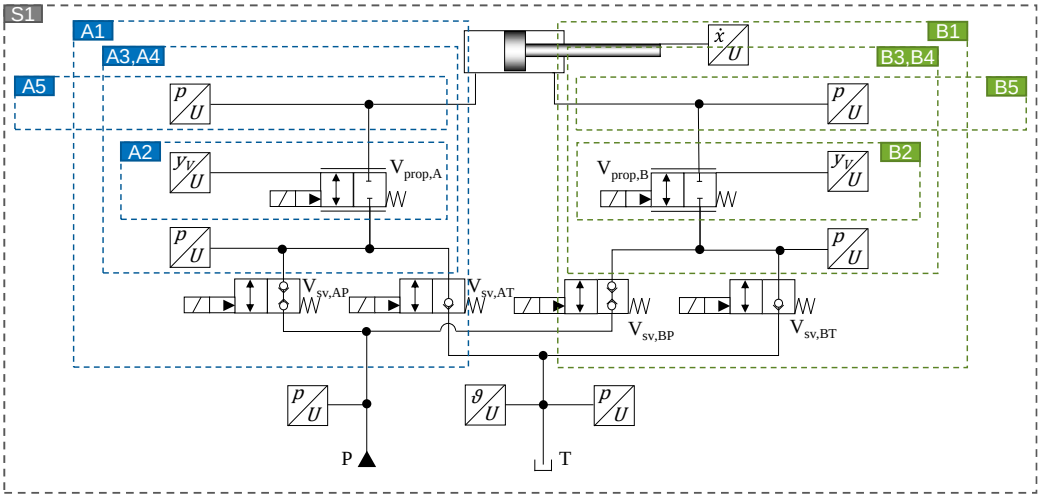


Figure 1: Design of the investigated independent metering system.

In addition to the physical test bench, a validated software-in-the-loop (SIL) environment is used for the development of the fault detection and localization algorithms. A comprehensive variant study enables the consideration of a large number of components that may cause errors, as well as different error characteristics and operating points. This takes into account the system complexity on the one hand and the operator-guided control of mobile machines on the other. The SIL environment has certain limitations that are typical in software development for mobile machinery, e.g. low computing capacities and restricted task times of the control unit. In total, 7 operating points with 542 error scenarios each form the basis of the investigation [4].

3. MODELL-BASED FAULT DETECTION

The test rig's vast sensory equipment enables the calculation of several parity equations. The aim of these equations is to generate residuals through the comparison of measured and calculated output variables. To calculate the theoretical values, practical model descriptions are used for fluid behavior, valve dynamics, pressure build-up in hydraulic capacities, and volume flow through the valves. Combined with sensor data and control values, this results in two parity equations for the valve positions (A2, B2, see Figure 1), eight for the cylinder velocities (A1 to A5, B1 to B5, except A2, B2) and one for the volume flow to be supplied by the pump (S1). The majority of parity equations have fixed, very local balance limits and are entirely defined by the associated measured and calculated values. These equations are therefore also called local parity equations. Equations A5, B5

and S1 contain set points from the machine control, which depend on the state of the overall system. Thus, these equations are termed global parity equations.

A threshold check controls the resulting residuals. Because the modeling is as simple and practical as possible, deviations between calculated and measured values are inevitable. If the residuals exceed the acceptable threshold for a prolonged duration of time, the mathematical model deviates significantly from the actual system process. In this case, a faulty state is to be assumed and an error symptom is generated. This symptom is used for error detection.

As shown in [4], model-based fault detection generates at least one symptom for up to 70% of the imprinted fault cases, dependent on the operating point and as displayed in **Figure 2**. However, a significant percentage of the errors remain undetected. To evaluate the effectiveness of fault detection, it is necessary to assess the potential danger of any detected or undetected errors. For this purpose, it is suitable to consider the position deviations of the cylinder due to the error in comparison to the fault-free state. The histograms illustrate these position deviations Δx_{cyl} 2 seconds after the fault occurs. In the case of an undetected fault (symptomless), the movement of the cylinder generally deviates only slightly from its fault-free state. However, if a fault is detected, serious and dangerous changes in movement may occur. The undetected faults therefore do not lead to a dangerous situation, while all dangerous faults are detected. This leads to Diagnostic Coverage values of up to DCavg=99 % for the safety evaluation process.

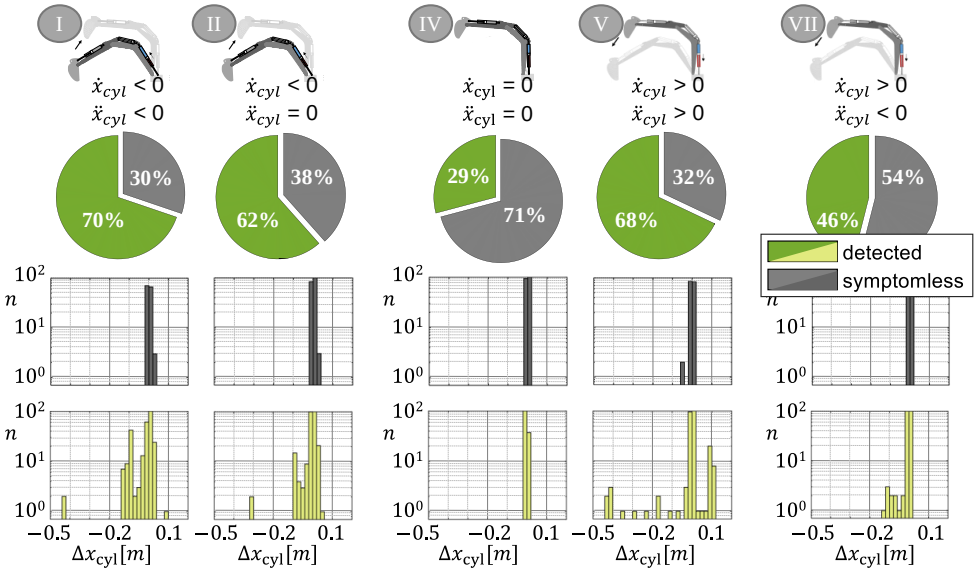


Figure 2: Amount of detected faults and their impact on the cylinder position for different operating points.

To confirm the results of the simulative variant study, various fault scenarios are reproduced on the test rig. For this purpose, the detection algorithms are transferred to the test rig controller (HY-TTC 580). To ensure consistency, the test rig demonstrator autonomously executes the reference cycle of the simulation study. A Fault Insertion Unit enables the time-controlled emulation of component faults so that the practical tests match the simulated conditions as closely as possible.

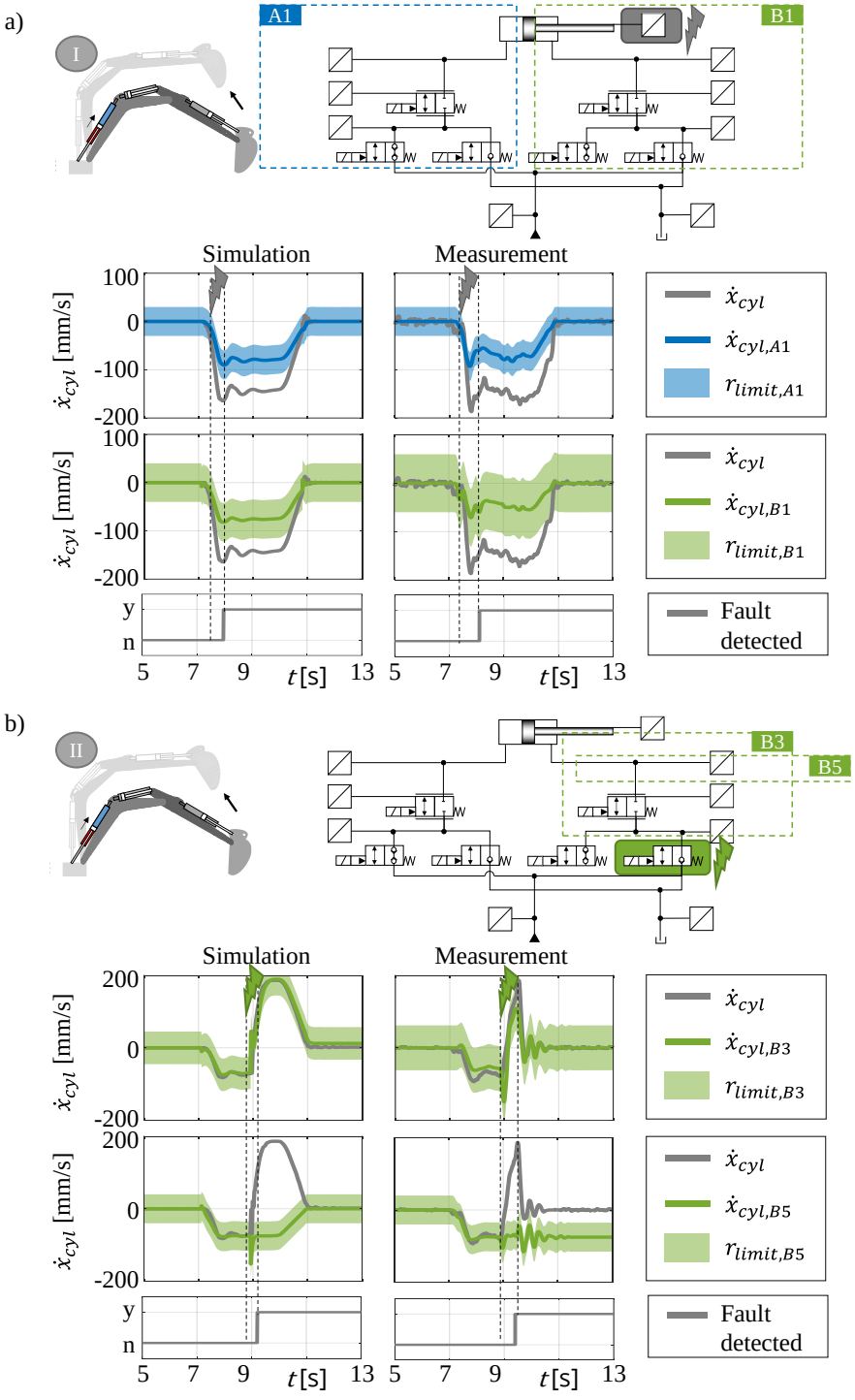


Figure 3: a) Error of the cylinder velocity sensor at operating point I, b) Error of the switching valve $V_{sv,BT}$ at operating point II.

Figure 3 shows the comparison of simulated and measured behavior for two fault scenarios. The fault detection of two parity equations for a malfunctioning cylinder velocity sensor is displayed in Figure 3 a). The absolute measured value is greater than the actual velocity. The parity equations A1 and B1 calculate a velocity the cylinder should achieve within their balance limit, the corresponding valve section and cylinder chamber. The measured velocity signal acts as a reference value for both equations. Due to inaccuracies in the model and measurement deviations, the calculated velocity is subject to an uncertainty (r_{limit}). When an error occurs, the measured value exceeds the acceptable tolerance range. If the deviation remains over for a certain period of time, the fault detection identifies the presence of a faulty condition. Since the tolerance range was exceeded by both A1 and B1, both parity equations produce an error symptom. Simulation and measurement match qualitatively and quantitatively very well.

Another example is shown in Figure 3 b). During the lifting of the boom, the valve $V_{sv,BT}$ is opened in a faulty manner. As a result, the load-holding cylinder side is relieved against the reservoir and the attachment drops to the ground. The faulty valve is outside the balance limits of the local parity equation B3, which is mainly determined by valve flow characteristic of $V_{prop,B}$. Therefore, there is no error in the scope of the parity equation. The figure clearly shows that equation B3 accurately computes the motion resulting from the error. The global parity equation B5, on the other hand, detects the fault immediately after it occurs. Only B5, which uses the set volume flow through valve $V_{prop,B}$ for calculation, produces an error symptom. Again, measurement and simulation match very well. In the absence of any error reaction measure, the boom cylinder reaches its end position in the simulation while its motion is constrained by the floor during the laboratory test.

The test rig results on the one hand validate the results of the variant study and on the other hand show the selective behavior of the generated symptoms in case of different faults. This is a prerequisite for a proper fault localization.

4. DEVELOPMENT OF FAULT LOCALIZATION

The aim of fault localization is the identification (detection) of a faulty component. This first requires the detection of a faulty state. As seen before, the fault detection method utilizing eleven parity equations and the resulting symptoms provides a solid foundation for achieving this goal. In the following, four different approaches for fault localization will be presented. These work in detail…

1. …on the basis of complete feature matching,
2. …using decision or fault trees,
3. …according to the Manhattan distance or
4. …via learning vector quantization.

As shown, the parity equations have different error sensitivities due to their different balance limits. According to these limits, there are expectations regarding the symptom patterns that occur with different faults. **Table 1** illustrates the expected symptom patterns for cylinder retraction.

Table 1: Expected symptom pattern for different fault classes during cylinder retraction

		Symptom generating parity equation							
Fault description, Error of…	Fault class	A1	A2	A3	A4	B1	B2	B3	B4
the proportional valve $V_{prop,A}$	i-p-A	x			x				
the proportional valve $V_{prop,B}$	i-p-B					x			x
the position sensor on valve $V_{prop,A}$	y-p-A-sens		x	x					
the position sensor on valve $V_{prop,B}$	y-p-B-sens						x	x	

the switching valves on the piston side or of the tank pressure sensor (no distinction possible)	pT,AT,AP	x						
the switching valves on the rod side or of the system pressure sensor (no distinction possible)	p0,BT,BP				x			
the cylinder pressure sensor on the piston side	p-A-sens	x	x	x				
the cylinder pressure sensor on the rod side	p-B-sens				x		x	x
the pressure sensor between the switching valves and the proportional valve (piston side)	p-α-sens		x	x				
similar to fault class p-α-sens, but for rod side	p-β-sens						x	x
the cylinder velocity sensor	v-cyl-sens	x	x	x	x		x	x

The first three approaches take advantage of this knowledge. However, as described above, the global parity equations do not have the required selectivity. Consequently, the localisation of these methods relies entirely on local parity equations. For this reason, only the local symptom-generating parity equations are listed in Table 1.

4.1. Fault localization based on the complete feature matching

As shown in Table 1 faults with the same symptom pattern were grouped into a fault class (e.g.: pT,AT,AP). The current set of parity equations cannot distinguish between them. To differentiate between these faults, an expanded equation base is required. For example, Richter [6] and Kramer [7] present a way to directly monitor the state of the switching valves. Each fault class has a unique symptom pattern. Therefore, if one of these individual patterns appears due to a faulty condition, it can be assumed that the component fault underlying the fault class is responsible. This constitutes a basic form of fault localization. All detected fault cases are analyzed for complete feature matching, based on the variant study. A confusion matrix presents the result of the fault localization. The matrix compares the actual error class (true class) with the class classified by the algorithm (predicted class). The matrix's numerical values indicate the frequency of each classification. On the matrix diagonal, correctly classified error cases can be identified.

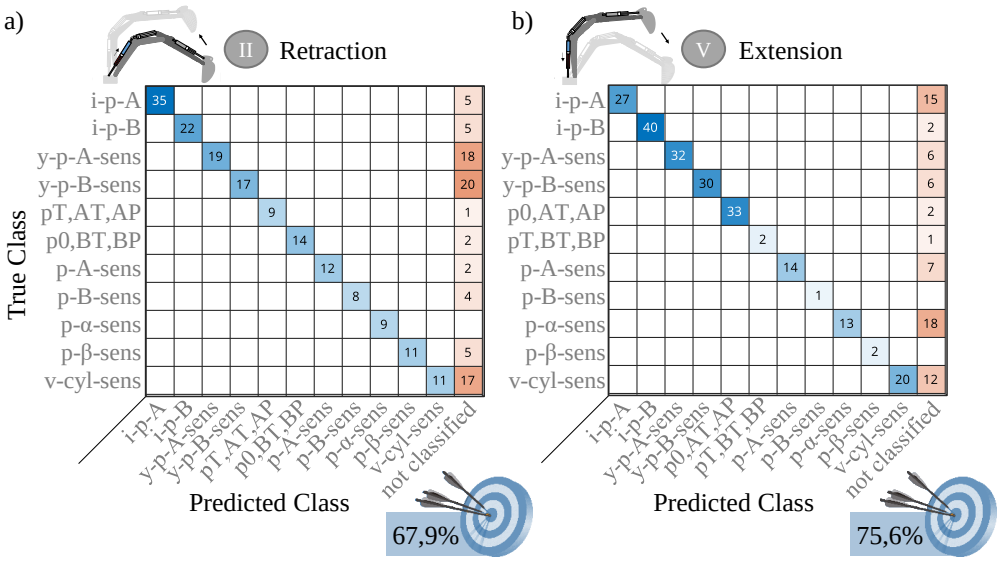


Figure 4: Confusion matrix for a) operating point II and b) operating point V with localization by means of complete feature matching.

As expected, classification based on full feature expression does not lead to any mislocalizations (refer to **Figure 4**). Nevertheless, the overall localization accuracy is only about 68% and 75%. In approximately 30% of the analyzed detected faults, not all parity equations produce an expected symptom. This leads to the absence of a complete symptom pattern, making a fault case classification impossible (not classified faults).

4.2. Fault localization using fault trees

Since a complete symptom pattern is a major limitation for fault localization, other approaches are being considered. One of these is the utilisation of fault trees. According to Isermann, error trees belong to statistical approaches [8]. This approach uses a series of hierarchical decisions to identify possible fault causes. Starting from a root decision node, the presence or absence of particular symptoms is checked. Depending on the result, the following decision node is then reached. At the end of each decision path are the individual fault classes. A complete feature expression is not required, but it improves localization quality. There are various methods to create a fault tree. One method is to train the fault tree using the existing data set (variant simulation). An other method is to use the expected symptoms (refer to Table 1) to generate an analytical fault tree. The analytical approach allows the inclusion of "expert knowledge", such as the combination of criteria as well as the prioritization of decision nodes. **Figure 5** illustrates the result of the classification based on an analytical decision tree. The confusion matrices demonstrate that improved error localization is achievable at both operating points, with an increase in overall localization accuracy of between 12 to 16%. The fault tree assigns a possible fault class to all detected faults. However, it can result in some incorrectly localized faults. For example, 16 out of 28 faults of the cylinder velocity sensor are localized correctly in operating point II.

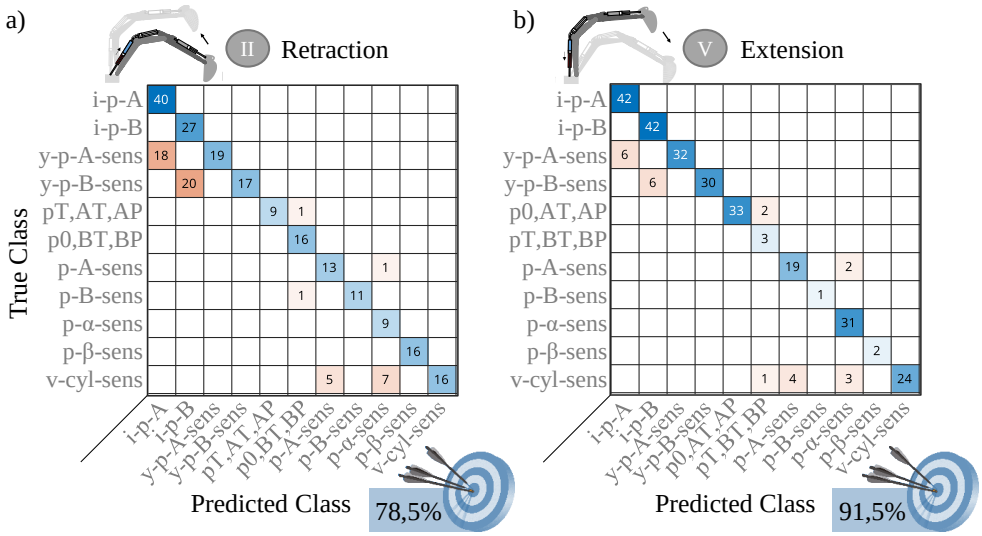


Figure 5: Confusion matrix for a) operating point II and b) operating point V with localization by means of fault tree.

The hierarchical structure of a fault tree is a drawback as it strongly prioritizes selected symptoms, making the sequence of decisions critical. Correct localization becomes challenging when a highly prioritized expected symptom is absent.

4.3. Fault localization according to the Manhattan distance (City-block-distance)

The Manhattan distance is a geometric classification method, where the assignment of individual fault classes is based on geometric distance [8]. Each class has a prototype that is specific to its class, whose symptom expression is characteristic and representative for this particular class. In this case, the prototype aligns with the expected symptom pattern. The closer the distance between a prototype and an object to be classified, the higher the likelihood of it belonging to the corresponding class [9].

The term Manhattan (or city block) distance originates from the fact that in urban areas, distance between two points can only be determined along specific coordinate directions (e.g. streets, lifts). The Euclidean distance ("as the crow flies") is not suitable for this. By interpreting the features of individual parity equations as directions (A1, A2, ..., B4) and their corresponding symptoms (0,1) as distance, this method can be used for fault localization. Neither a fully developed symptom pattern nor a hierarchical structuring of the individual symptoms is necessary. In general, distance measures offer the possibility of weighting diverse "directions" differently. However, since the equality of all symptoms is an advantage over fault trees, this possibility is not used. Nevertheless, a suitable weighting is purposeful. By choosing robust threshold values for fault detection, the occurrence of unexpected symptoms is hardly possible, while the absence of an expected symptom is probable. For this reason, the matching of existing symptoms is weighted more heavily than their absence.

The geometric classification achieves a very good overall localisation accuracy above 90% for both operating points, as shown in **Figure 6**. Due to the possibility of several prototypes having the smallest distance value, a single fault scenario can be assigned to multiple classes. The sum of all classifications can therefore exceed the sum of the investigated scenarios. If other potential causes are also assigned to a fault in addition to the correct fault class, it is still considered a correctly localized fault, since focused troubleshooting in practical use is simplified.

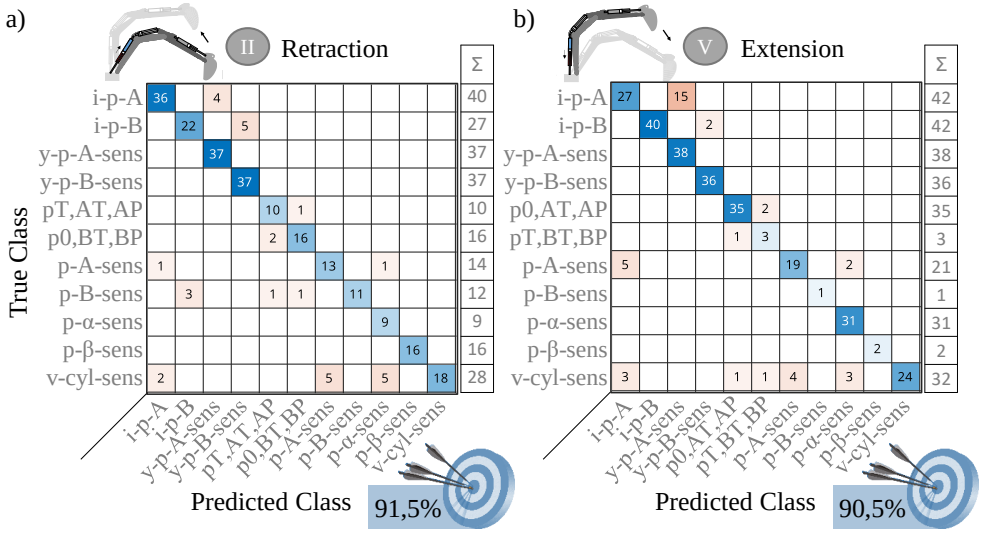


Figure 6: Confusion matrix for a) operating point II and b) operating point V with localization by means of Manhattan distance.

4.4. Fault localization via learning vector quantization

Learning vector quantization (LVQ) originates from the field of machine learning. Unlike the other three variants for fault localization, which follow a causal and analytical approach, LVQ cannot easily

transfer its results to other system structures. For this reason, this approach will only be briefly discussed. LVQ is a prototype-based artificial neural network method for data classification. By an optimization, the choice of the prototype characteristics takes place in such a way that the associated data points have a minimum, class foreign a maximum distance to the prototype. An object is then classified according to the resulting Voronoi regions using the winner-takes-it-all principle. Consequently, only one fault class is assigned to each fault scenario. Numerous LVQ algorithms are available in literature. The "Generalised Relevance LVQ" algorithm described in [10], [11] was utilised in this work. This method optimizes the features of the prototypes as well as the weighting of the symptoms.

Despite the extensive variant study, only few training data sets are available, if one applies the standards of machine learning algorithms. Furthermore, there is no uniform distribution of failure scenarios, since some components have a binary fault character (e.g. switching valve is open or closed). As a result, the promising LVQ algorithm reaches its limits and only achieves classification accuracies within the range of 80%. In addition, there are the well-known challenges in the use of machine-learning algorithms, such as transferability of results and overfitting.

5. TESTING OF FAULT LOCALIZATION ON THE TEST BENCH

For fault localization on the test bench, the geometric classification method is implemented using the Manhattan distance. This consistently continues the knowledge-based approach of model-based fault detection and achieves the best localization accuracy in the simulation study. **Figure 7** shows the detected threshold exceedances as a result of a velocity sensor error. The first symptom is generated by parity equation A3 where the status changed from false “f” to true “t”. Other symptoms appear subsequently, mainly of equations A1, B1, A4, A5 and B5. If the movement of the boom ends at 10.5s, the permitted thresholds are no longer exceeded. The parity equations A1 and A4 are the last to fall below the thresholds.

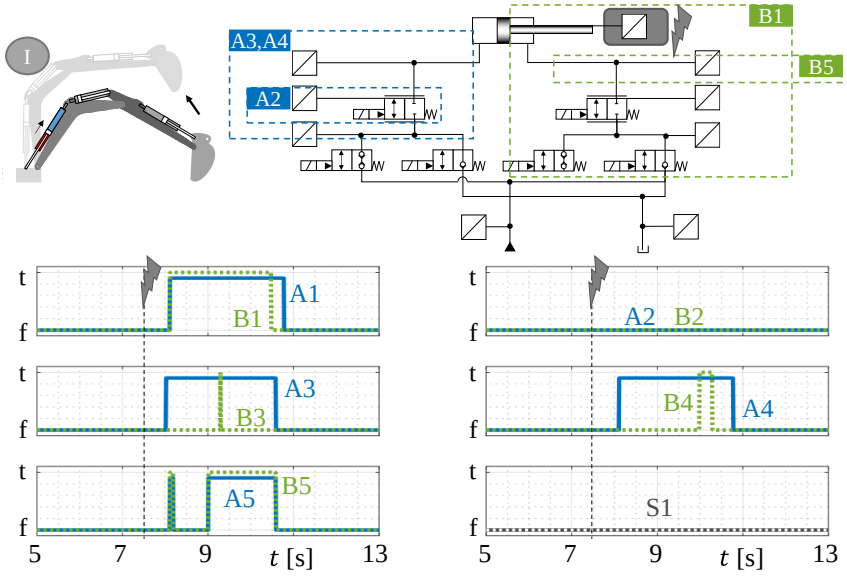


Figure 7: Currently detected threshold exceedances of all parity equations in case of a velocity sensor error at operating point I (measurement)

The fault cases are classified based on temporary threshold violations and corresponding symptoms.

Figure 8 clearly shows that the fault is mostly classified correctly during the movement phase (the classification marker "v-cyl-sens" changes from false to true), although there is no fully developed symptom pattern at any time (see Table 1). Only at the beginning and at the end of the movement the calculated Manhattan distance indicates another cause of error. Nevertheless, the measurement illustrates the effectiveness of fault localization and the validity of the simulatively obtained statements in an exemplary manner.

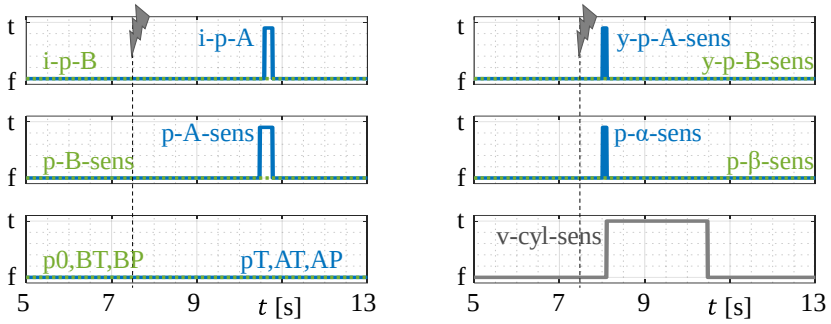


Figure 8: Resulting classification in case of a velocity sensor error at operating point I (measurement)

Another classification example is displayed in **Figure 9**, showing a malfunctioning rod-side pressure sensor p_β positioned between the proportional and switching valves. This sensor outputs the maximum value instead of the real pressure. The machine controller uses the wrong value for valve actuation, causing the proportional valve on the drain side to open too wide. The system cannot generate the necessary pressure to support the equipment, causing the boom to move too quickly towards the ground (see **Figure 10 a)**). This fault state is reliably detected by parity equations B3 and B4 and accurately localized by Manhattan distance.

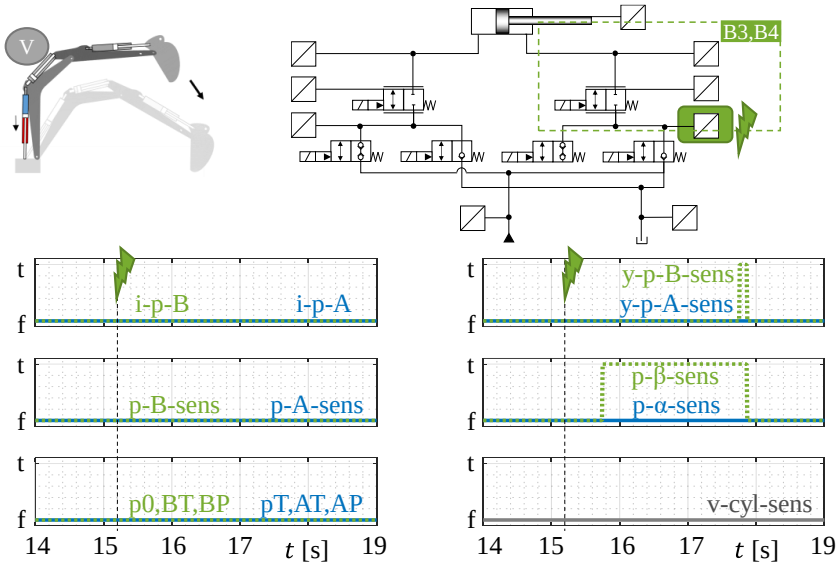


Figure 9: Classification of error “faulty value of sensor p_β (maximum value)” at operating point V.

In case of detecting a fault, appropriate responses can be taken to address the issue. One option is to close all valves and shut down the actuator completely. This is a safe option, but at the expense of

availability. The correctly localized cause of error enables another option: reconfiguration.

Reconfiguration describes the upkeep of functionality in the case of failure, through compensating the faulty component. IM systems basically offer 2 possibilities for reconfiguration:

- mode switching
- alternative control strategies

Mode switching describes the structural compensation of a fault through alternative flow paths and is a characteristic of IM systems. However, the use of alternative control strategies is applicable to all appropriate systems. A key requirement for this compensation is usually an extensive sensory set-up. When a faulty sensor is identified, a control concept for valve actuation is chosen that does not rely on the faulty sensor value. The performance of both concepts can be demonstrated on the test rig. In the case of the sensor error p_β, the use of an alternative control strategy is suitable (see Figure 10 b**)**). After the error occurs, the fault cause is accurately identified (see Figure 9) and the control strategy is adjusted accordingly. The pressure drop over the proportional valve can no longer be determined directly due to the faulty sensor. By using the system pressures in the supply and tank lines and a combined resistance model of the switching and proportional valve, adequate valve control is nevertheless possible. This ensures that the cylinder's movement remains manageable, preventing hazardous movements as well as drive shutdown.

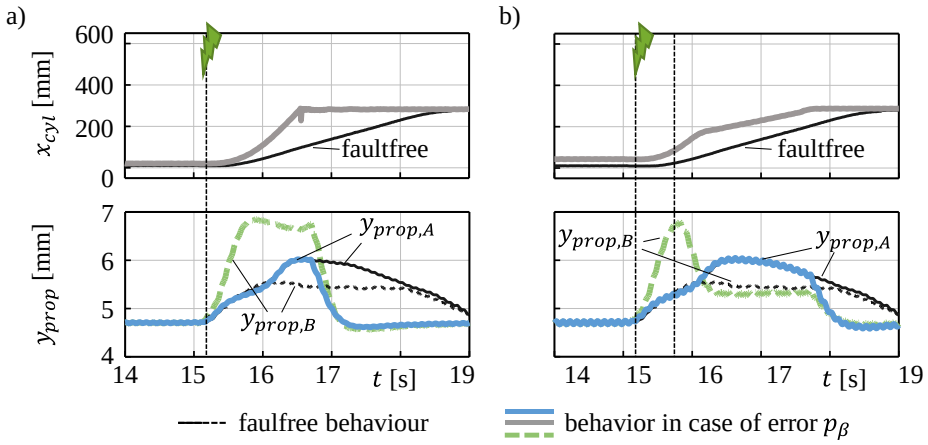


Figure 10: Measured behavior in case of error "faulty value of sensor p_β (maximum value)" a) without error detection and b) with error detection and reconfiguration at operating point V.

6. CONCLUSION AND OUTLOOK

The paper presents a model-based fault management for IM-systems. Fault detection and fault localization utilize existing system knowledge. The practical modelling approach employs a set of 11 parity equations for fault detection. The different balance limits of these equations generate selective symptoms for different fault scenarios. This is an essential prerequisite for identifying the specific fault cause.

Different approaches have been investigated for fault localization. In particular, localization based on Manhattan distance leads to promising results of up to 92% accuracy. This method can easily be transferred to other system structures. However, it turns out that an extension of the parity equation set could provide benefits. Particularly, descriptions of component-specific elements enhance the reliability of detection and increase depth of localization.

The reconfiguration of correctly localized faulty components can be successfully demonstrated. Depending on the fault and reconfiguration possibility, partial or even complete compensation is possible.

The transfer of the investigated algorithms to the real system allows the validation of the results. The exemplary verification leads to a good agreement with the results of the simulation study. The test rig results therefore confirm the theoretical analyses and conclusions.

7. ACKNOWLEDGEMENTS

The project “Adaptive modellbasierte Fehlererkennung und -diagnose mit rekonfigurierbaren elektrohydraulischen Systemen” (IGF 20246 BR/1) was financed and supervised by the Research Association Mechanical Engineering (FKM), Frankfurt am Main. In the scope of the program to promote Industrial Collective Research it was funded by the German Federation of Industrial Research Associations (AiF) with means of the Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag.

Supported by:

Federal Ministry
for Economic Affairs
and Climate Action

on the basis of a decision
by the German Bundestag

NOMENCLATURE

A1…A5	Parity equations or their symptoms on the piston side	
B1…B5	Parity equations or their symptoms on the rod side	
Δx_{cyl}	Deviation of the cylinder position due to an error	m
IM	Independent metering (system)	
r_{limit}	Threshold of the parity equation residuals	
S1	Parity equation S1 or its symptom	
t	Time	s
t/f	true/false	
$V_{prop,A/B}$	Proportional valve on the piston / rod side	
$V_{svBP/BT}$	Switching valve to connect the rod side with pump or reservoir	
$V_{svAP/AT}$	Switching valve to connect the piston side with pump or reservoir	
x_{cyl}	Cylinder position	mm
$\dot{x}_{cyl}$	Cylinder velocity	m/s
y/n	yes/no	
$y_{prop,A/B}$	Position of proportional valve $V_{prop,A/B}$	mm

REFERENCES

[1] Eaton (2016), Data Sheet, CMA200 Advanced Independent-Metering Mobile Valve. 2016

[2] Danfoss (2016), Data Sheet. Proportional Valve PVX. 2016

[3] Beck B., Weber J. (2016), Safety and Reliability of Independent Metering Systems in Mobile Machinery. In: 26th European Safety and Reliability Conference, ESREL 2016, Glasgow, Scotland, September 25-29, 2016,

[4] Fischer E., Beck B., Weber J. (2022), Data analysis for the evaluation and design of a model-based fault detection based on an independent metering system for mobile hydraulic drives. In: 13th International Fluid Power Conference, 13. IFK, Aachen, Germany, March 21-23, 2022

[5] Jansson A., Krus P., Palmberg J.-O. (1991), Decoupling of response and pressure level in a hydraulic actuator. In: Fluid power systems modelling and control. Fourth Bath International Fluid Power Workshop on Systems. Bath, UK, September 18-20, 1991

[6] Richter S. (2016), Ein Beitrag zur Integration von Maßnahmen der funktionalen Sicherheit in einen geregelten elektrohydraulischen Antrieb für stationäre Anwendungen. PhD thesis, TU Dresden, Shaker Verlag, Aachen, 2016

[7] Kramer T., Weber J. (2018) ,Intelligent Condition Monitoring of Bi-stable Process Valves. In: 11th International Fluid Power Conference, 11. IFK, Aachen, Germany, March 19-21, 2018

[8] Isermann R. (2006), Fault-Diagnosis Systems. An Introduction from Fault Detection to Fault Tolerance. Springer, Berlin Heidelberg, 2006

[9] Clauß G., Finze F.-R., Partzsch L. (1994), Statistik. Für Soziologen, Pädagogen, Psychologen und Mediziner. Band 1, Verlag Harri Deutsch, Thun, Frankfurt am Main, 1994

[10] Bunte K., Schneider P., Hammer B., et al. (2012), Limited Rank Matrix Learning - Discriminative Dimension Reduction and Visualization. In: Neural Networks, Volume 26, 2012

[11] Schneider P., Biehl M., Hammer B. (2009), Adaptive Relevance Matrices in Learning Vector Quantization. In: Neural Computation, Volume 21, 2009

Chapter 11

Sustainable Pneumatics

peer reviewed

CONTROL OF A PNEUMATIC SYSTEM FOR MATERIAL STRENGTH TESTING

Željko Šitum[1]*, Marko Miroslav Bača[2], Juraj Benić[1], Željko Alar[2]

[1]*Department of Robotics and Production System Automation, University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, Ivana Lučića 5, HR-10000 Zagreb, Croatia*
[2]*Department of Materials, University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, Ivana Lučića 5, HR-10000 Zagreb, Croatia*

* Corresponding author: Tel.: +385 91 733-4072; E-mail address: zsitum@fsb.hr

ABSTRACT

The article describes a prototype device for testing the dynamic stress of materials with the intention of analysing the fatigue behaviour of the materials. A pneumatic cylinder controlled by a proportional pressure regulator is used to achieve the required force. The experimental set-up consists of structural elements with pneumatic components, then a control system with a Controllino device and HMI interface, and the measuring system composed of a dynamometer with an amplifier characterized by high accuracy and sampling frequency characteristics. With this system, it is possible to work in both compression and tension ranges. The basic parameters of the process were identified, then the control-oriented dynamic model of the system was derived, and the synthesis of the PID controller was made for more precise following of the sinusoidal reference signal. Compared to hydraulic solutions that are used to achieve large forces, this pneumatic fatigue testing machine has a compact design, the desired reference is simply set via the HMI interface, it is portable, and it can be used to test materials with lower tensile strength.

Keywords: Pneumatic fatigue testing machine, Material strength testing, Dynamic endurance

1. INTRODUCTION

The common intersection of the main areas of production engineering such as technologies, materials science and construction overlap with the analysis of stress and deformation of materials, i.e. on the strength of materials. Material strength is a fundamental feature in the analysis and design of structural parts, as it determines the permitted capacity and load of a complex system (with the addition of a safety factor). If the permissible stress limit is exceeded, the most common approach is correction, which in practice means increasing the dimensions or changing the material. In materials science, tests of mechanical properties of different materials under conditions of long-term dynamic action of variable stress have special significance because the results of these tests determine the dimensional calculation of mechanical structures [1]. Understanding the fatigue mechanism is important for considering various technical conditions and this knowledge is essential for the analysis of strength properties of an engineering structure [2]. Experimental testing of components under static and dynamic loading tests is necessary for ensuring the reliable operation of structural elements. As a result of long-term periodic variable loading, material fatigue occurs. This causes fracture of the material at a significantly lower stress than the maximum tensile strength of the material. In mechanical laboratories, two types of testing machines are most often used to perform fatigue tests with random loading. These are servo-hydraulic devices [3], which generate force and/or torque as the input excitation and electrodynamic shaker tables [4], which apply a force to the vibrating table. For the needs of biaxial in-plane fatigue tests, the servo-hydraulic actuator type is the most common

option available in the market [5]. The analysis of mechanical properties of different materials under conditions of long-term dynamic action is often a crucial criterion when designing machine parts, and therefore there are many manufacturers of devices for material strength testing on the market. Many reputed companies around the globe manufacture fatigue testing machines, so these machines are available in different versions and in different price ranges. Three types of fatigue testing machines are the most common types of this machine: strain control, low cycle fatigue, and load control high cycle fatigue. The purpose of fatigue testing is to determine the lifespan that can be expected from a material subjected to cyclic loading. However, fatigue strength and crack resistance are also desirable information about material behaviour. The fatigue life of a material is the total number of load cycles to which the material can be subjected. Fatigue testing is also used to determine the maximum load that a specimen can withstand for a specified number of cycles. These features are extremely important for the use of various materials subject to dynamic loads. For the realization of smaller forces, the dynamic endurance of the material could be analysed on test systems (fatigue testing machines) that use pneumatic actuators. Current intensive technological development requires an increasing number of production procedures with a simultaneous reduction of the material used, so ensuring the reliable operation of structural elements requires experimental testing of components under static and dynamic straining. Modern materials testing devices require the possibility of programmed motion and/or actuator force control, as well as quick adaptation of control algorithms to new requirements in the material testing process. When designing complex systems, it is often quite difficult to predict all the sections where the critical stress occurs. Simulation models can give a good insight into the behaviour of the system during exploitation, but the need for a real experiment is necessary to validate the results. Since it is theoretically impossible to determine the exact value of the dynamic endurance of the material, in practical applications experimental tests are carried out on specialized fatigue testing machines. Such tests provide the necessary parameters for construction calculations to produce components of smaller dimensions that meet the requirements for a longer period of operation. Fatigue testing machines are most often based on a servo-hydraulic drive, they can produce very large forces, but the dimensions of such devices are very large. Nowadays, when many products are made from polymer or metal sheets, there is a need for test devices that can realize smaller forces, which makes the application of pneumatic test devices suitable for these tasks.

The aim of this work was to produce a universal pneumatically driven test device on which static and dynamic loading tests can be carried out. For static tests, the accuracy of the system is very important, and the response speed is less important, while for dynamic tests, the response speed is essential for achieving higher operating frequencies, and the accuracy is a secondary requirement [6]. The use of pneumatics makes it possible to achieve the necessary forces for testing materials of lower strength, the system will be of smaller dimensions and suitable for desktop and mobile use in laboratories for testing the dynamic properties of materials.

The experimental set-up must meet the necessary norms and recommendations for work in test laboratories and must be able to carry out a complete dynamic test of the material.

2. EXPERIMENTAL TEST DEVICE

The constructed experimental system, shown in **Figure 1**, consists of a pneumatic system, a control system with HMI interface, and a measuring device with an amplifier. The construction of the test device is made from standard aluminium profiles due to the high modularity and wide choice of shapes that allow the connection of various profiles using angle joints. The upper horizontal beam is used as a support for the pneumatic cylinder. It is movable along the vertical axis and ensures the required position of the pneumatic actuator during the testing procedure [7].

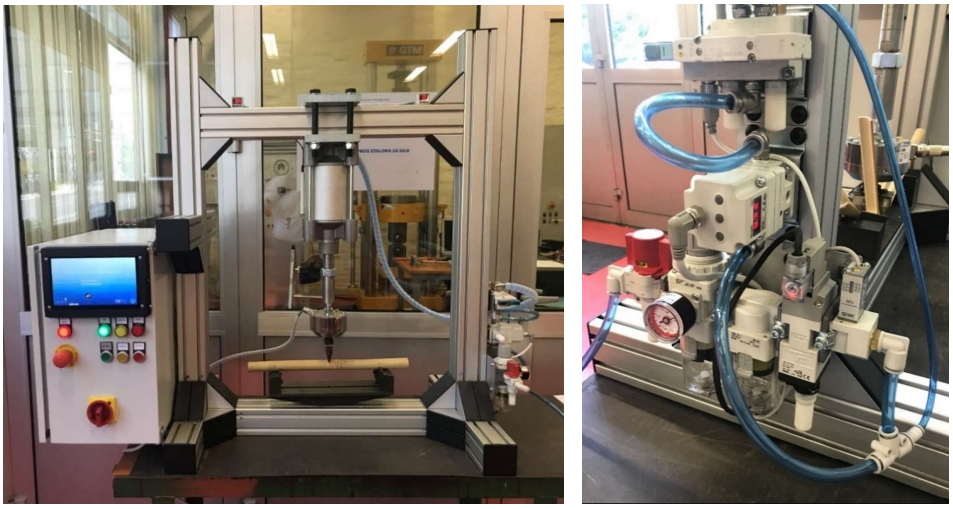


Figure 1: Fatigue testing machine

The pneumatic system of the test device is physically detached from the electrical part, which enables simple corrective maintenance with easy access to all components. The air preparation unit is made as a compact component and consists of a manual 3/2 valve, a micro air filter, an electromagnetic 3/2 valve and a pressure switch. **Figure 2** shows the pneumatic scheme of the experimental system.

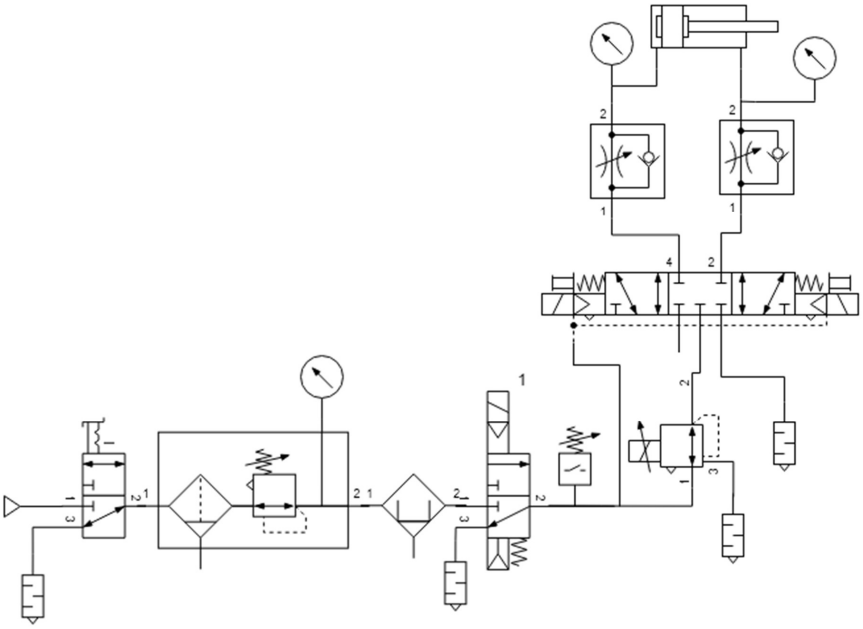


Figure 2: Pneumatic scheme of the experimental system

Dynamic loading of test samples is achieved by applying forces of different frequencies and amplitudes. The diameter of the piston and the stroke of the pneumatic cylinder are selected according

to the force requirements, considering standard industrial pressures. To prevent *stick-slip* effect, a smooth, slow-moving cylinder of the new generation, type SMC CY-96YDB125-100 was chosen, which has lower friction values at lower pressures than standard cylinders. A proportional pressure regulator SMC ITV2050-312N3 was used to control the force of the pneumatic cylinder. An external pilot solenoid 5/3 valve SMC SY5301RK-5U1 with a metal seal was used to supply compressed air to both cylinder chambers. Two flow control valves are used only in the preparatory phase, to strongly reduce the flow and decrease the speed of the actuator movement to achieve a proper effect of the actuator on the test sample in the clamping device. During the test phase, they are fully open. The maximum dimension of the test sample is 50 cm x 100 cm (w x h), and the length of the sample can be adjusted by using supports (e.g. ski testing). The maximum displacement of the cylinder is 100 mm. The amount of force depends on the pressure of the compressor, which is set to 10 bar, and multiplied by the surface of the cylinder gives a maximum force of 11 kN.

Controllino Maxi Power Automation PLC with software that is fully compatible with Arduino is used as a control device, **Figure 3**. The basis of the Controllino device is the ATmega328P microcontroller, which is an integrated chip that controls devices and processes, stores and executes the program. All input-output connectors have electrostatic discharge protection, so it is a control device that combines the advantages of open source and programming methods with the standards and safety of industrial PLCs. At the beginning of the program, Controllino initializes all necessary variables, then certain pins are assigned the property of input or output pins. The direction of movement of the cylinder is controlled by sending an electrical signal to the electromagnetic valve, and the pressure in the cylinder is determined by the voltage signal on the proportional pressure valve. The device for communication between the system and the operator is an LCD screen with a diagonal of 9", which will be suitable for entering the necessary parameters and displaying graphs of measurement data from the sensor. The working environment of the designed system are normal laboratory conditions, so industrial quality devices are not required. **Figure 4** shows the HMI system SK-90DT manufactured by 4D Systems, which enables simple communication and data exchange with the control unit. The screen is mounted on the electrical cabinet.



Figure 3: Control device



Figure 4: HMI interface of the test device

The force measuring sensor is a dynamometer axially connected to the cylinder piston. The dynamometer has a measuring amplifier which enables reading the applied force and sending a signal to the PLC. The dynamometer, HBM type U10M and ClipX amplifier, is characterized by high accuracy and sampling frequency and enables operation in the compressive and tensile range. Calibration was done on the built-in dynamometer and accuracy in class 1 was achieved. On 100,000 performed cycles, about 2% were below the required tolerance. **Figure 5 s**hows the setup for testing

the flexural strength specimen made of bamboo, where the dynamometer for force measurement is attached to the piston rod of the pneumatic cylinder. For the procedure of testing mechanical properties of bamboo sample, a device for holding the sample was also made.

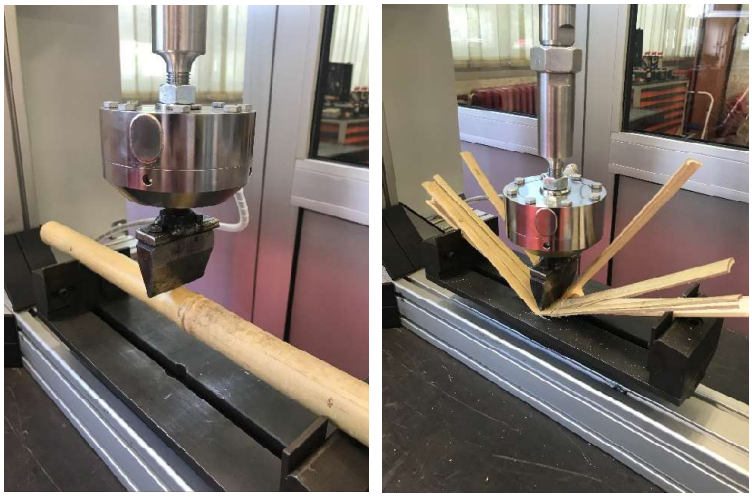


Figure 5: Testing the flexural strength

3. DYNAMIC MODEL OF THE TEST SYSTEM

The dynamic model of the test system was obtained by determining the transfer function of the proportional pressure valve and the dynamic model of the pneumatic cylinder, which includes the friction of the cylinder and the stiffness of the test object.

3.1. Dynamic model of the proportional pressure valve

The proportional pressure valve sets the output pressure proportional to the input voltage, indicating a linear characteristic as shown in **Figure 6**, with a linearity of ±1% of full scale guaranteed by the valve manufacturer [8].

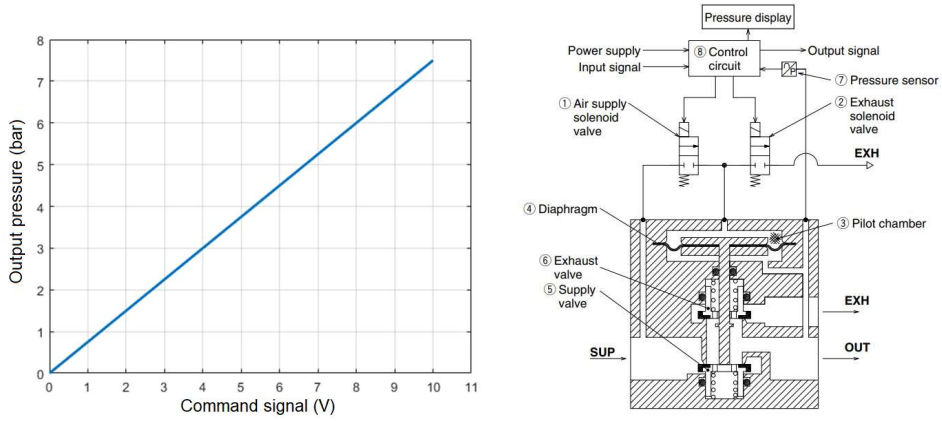


Figure 6: Static characteristic of the proportional pressure valve and working principle diagram

The forward gain K_v of the proportional pressure valve is determined from the calibration curve of the valve, which presents the relative pressure outlet versus the input voltage signal and is calculated as K_v=0.075 MPa/V. The value of the time constant T_v of the proportional pressure valve was determined experimentally based on the transient response of the output pressure to the step reference signal. The pressure is measured using a built-in sensor in the housing of the proportional valve. Three different input voltage signals of 3, 6 and 9 V are set on the pressure regulating valve. According to the static characteristic, the stationary value will be different, so the responses are scaled. **Figure 7** shows the pressure responses in the system for different voltage signals supplied to the proportional pressure valve, which are scaled to the forward gain K_v.

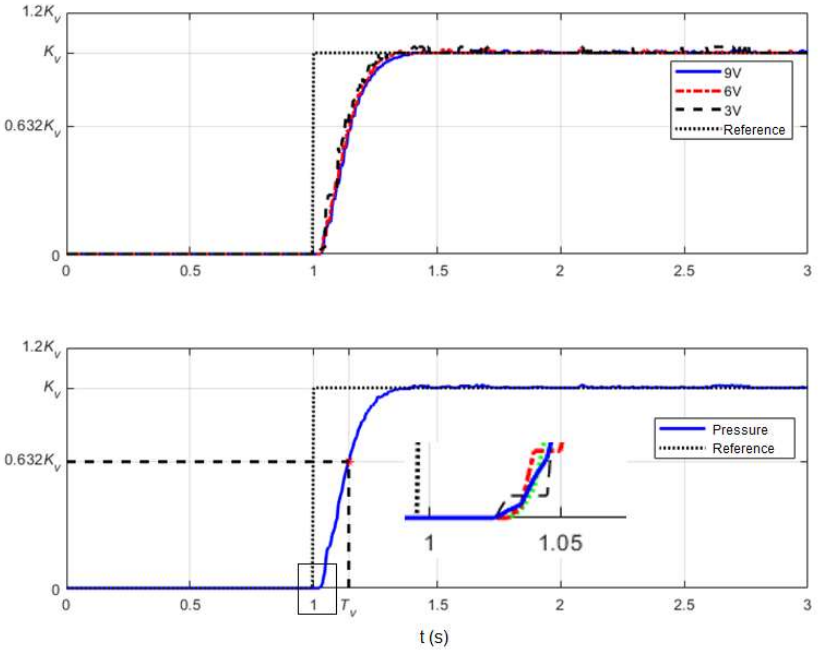


Figure 7: Pressure transient response

The second diagram in Figure 7 shows the mean value of the transient response curves for the previous three levels of voltage signals. From the transient response, the value of the time constant of the proportional pressure valve was calculated as the time required to reach the amount of 63.2% of the stationary value of the response, which gives the time constant T_v= 0.1422 s. The final expression of the PT1 term can be represented by the following transfer function:

$$G(s) = \frac{p(s)}{u(s)} = \frac{K_v}{T_v\, s + 1} = \frac{0{,}075 \cdot 10^6}{0{,}1422\, s + 1} \tag{1}$$

For a more accurate valve model, a transport delay (dead time) can be added to the proportional term in the simulation program, which according to the zoomed part in Figure 7 has a value of T_u=0.03 s. However, the transfer function with the added dead time makes the synthesis of the controller for the implementation of the experiment more complex.

3.2. Dynamic model of the pneumatic cylinder

The dynamic model of the pneumatic cylinder is obtained using the expression for the balance of forces of the mechanical system, assuming that air is an ideal gas.

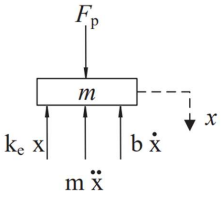


Figure 8: Forces on the cylinder piston

From [9], the reaction force of the test specimen is represented as a spring force. Given that dynamic testing of materials is primarily conducted within the elastic region of the material, such a representation of the reaction force is valid. Consequently, the first and second derivatives of the spring force are as follows:

$$x = \frac{F_c}{k_e} \quad , \quad \dot{x} = \frac{\dot{F}_c}{k_e} \quad , \quad \ddot{x} = \frac{\ddot{F}_c}{k_e} \tag{2}$$

The balance of forces on the cylinder piston can be given by the following expression:

$$m\,\ddot{x} = F_p - b\,\dot{x} - k_e\,x \tag{3}$$

where the applied force F_p is defined as the pressure differential across the piston multiplied by the cross-sectional area of the piston (A), which includes both the piston and connecting rod, m represents the piston mass, k_e is the stiffness constant, b is the viscous friction coefficient, and x is the position of the cylinder. The choice of the Newtonian friction model is based on the use of a smooth cylinder in the experimental setup, which exhibits low friction at lower speeds of cylinder movement and practically eliminates stick-slip motion.

The dynamical model of the hydraulic cylinder is obtained by inserting equations (2) into equation (3) as follows:

$$\ddot{F}_c = \frac{k_e}{m}\,(A_1\,p_1 - A_2\,p_2) - \frac{b}{m}\,\dot{F}_c - \frac{k_e}{m}\,F_c \tag{4}$$

The analysed system will work exclusively in the compression range, so the pressure in the second chamber will always be approximately atmospheric pressure.

$$p_2 = p_a \tag{5}$$

The final expression for the dynamics of the pneumatic cylinder where the force is the output value is as follows:

$$\ddot{F}_c = \frac{k_e}{m}\,A_1\,p_1 - \frac{k_e}{m}\,A_2\,p_a - \frac{b}{m}\,\dot{F}_c - \frac{k_e}{m}\,F_c \tag{6}$$

The control error e is expressed as the difference between the reference and actual force:

$$e = F_R - F_c \tag{7}$$

and this value will be the input to the controller which defines the output pressure on the proportional valve. A dynamometer is additionally connected to the piston rod of the cylinder, as well as other equipment for achieving contact with the test specimen and for load measurement. Equation (6) contains parameters whose values need to be determined, such as the mass m of the cylinder piston

and the attached measuring equipment, the coefficient of viscous friction b and the stiffness constant k_e. The mass of the cylinder piston and attached measuring equipment was determined using an external dynamometer and amplifier connected to the piston rod of the cylinder. The average value of three force measurements is converted into mass to be consistent with the unit of measurement in equation (6), and thus we obtain a precise amount of mass m=6.4269 kg.



Figure 9: Estimation of process parameters using a rubber bumper

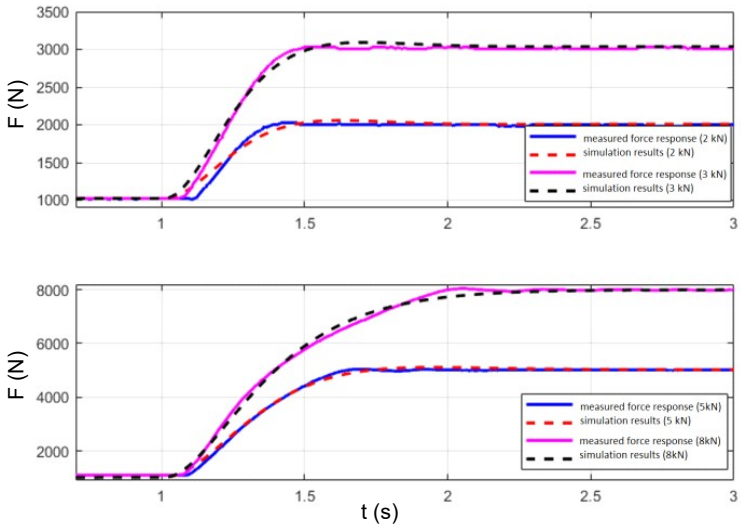


Figure 10: Comparison of measured force response and simulation model with estimated parameters

The values of the parameters b and k_e are difficult to determine directly as in the case of mass, especially in the case of the spring stiffness constant k_e which approximates the reaction force of the test specimen, since the specimens are always different. By measuring the experimental response of the system, it is possible to obtain the actual behaviour of individual system components, so this approach was used to determine unknown process parameters. The parameters b and k_e were determined from the experimental response of the force applied to the rubber bumper and the

estimation of the parameters of the simulation model using the normalized root mean square error - NRMSE method. This dimensionless method is used to measure the difference between prediction models and recorded data and is often used in statistical data analysis. A rubber bumper, shown in **Figure 9**, was used to simulate the test specimen, which can achieve larger deformations for the applied force. In the initial position, the cylinder piston touches the surface of the rubber bumper. The solenoid valve is open during the entire experiment, so its dynamics does not affect the pressure building process. The pressure increase in the cylinder is controlled by a proportional pressure valve, and its dynamics will be considered. A comparison of the simulation and experimental results for several cylinder pressure (force) values is shown in **Figure 10**, while the process parameters used in the simulation are given in **Table 1**.

Table 1: Parameters of the dynamic system model

Cylinder parameter	Value	Unit
Mass m	6.4269	kg
Piston area A_1	12.272×10^3	mm^2
Rod end area A_2	11.468×10^3	mm^2
Friction coefficient b	74	Ns/mm
Stiffness constant k_e	266	N/mm

By estimating the unknown parameters from the measured responses of the process and comparing them with the simulation results, a good matching between the experimental and simulation results can be observed for several pressure values (and thus the cylinder force).

4. EXPERIMENTAL SYSTEM CONTROL

For system control, the parallel structure PID controller is often applied, which has the derivation of the error signal as an input to the D-term. Although this approach is often used in industry, it should be avoided due to the appearance of large forcing of the control signal during step reference changes. Also, during each measurement of the controlled value, the noise of the measured signal appears, and the derivation significantly increases the noise level, so such a signal is almost unusable. To eliminate this problem, a PID controller with a speed estimation filter was used, which is shown in **Figure 11**.

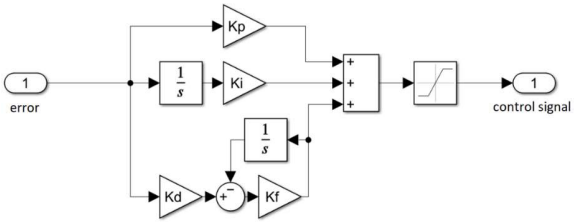


Figure 11: PID controller with speed estimation filter

The parameters of the PID controller used in the simulation and experiment were: K_p=0.0337, K_i=0.0961, K_d=0.0038, while the speed estimation filter was set to the value K_f=250. The initial parameters of the controller were determined from the experimentally recorded transient response of the process in the open loop, and then minor adjustments were made during the control process. The designed controller was checked on the experimental system and the response comparison is shown in **Figure 12**. It can be observed that there is quite a good matching between the simulation results and the experimental response in the control process for lower frequencies of the sinusoidal reference signal and a certain phase delay for the case of higher frequencies.

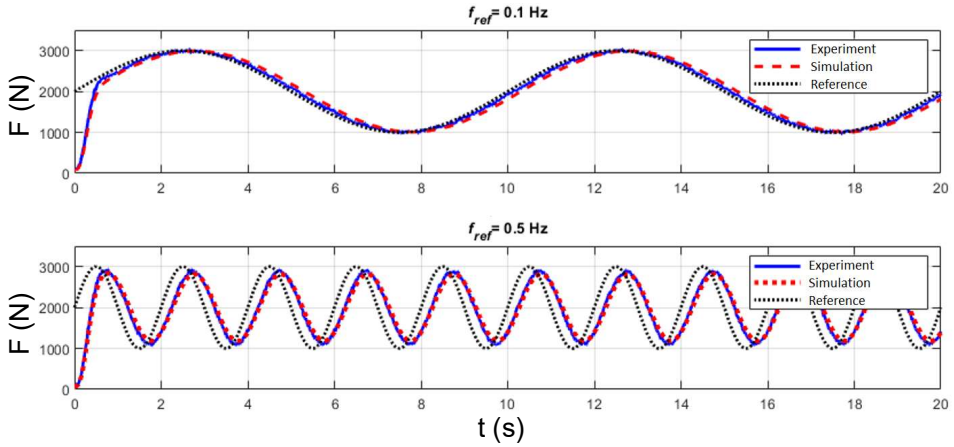


Figure 12: Comparison of experimental and simulation results using PID controller

During the testing, it is important to keep the amplitude deviation below the required value (e.g. below 10%), so the regulator is set to achieve the force output value slightly below the reference value. If the achieved force is significantly lower than the required value, in the control program the number of test cycles will be increased by the number of such inappropriate amplitudes. During dynamic testing of materials, the phase delay is less relevant and does not significantly affect the implementation of the experiment. Further improvements towards more accurate tracking of higher frequency reference signals are expectedly possible using feedforward control methods.

5. CONCLUSION AND OUTLOOK

Practical realization of own laboratory educational test device, which can be used for experimental demonstration of the fracture mechanics of materials caused by dynamic stress, analysis of the resistance of materials to cyclic stress and the development of cracks due to material fatigue is a valuable achievement in deepening knowledge of materials. An important feature of this test device is the use of pneumatic components to achieve the required dynamic forces and a user interface with a touch screen to set the parameters of the material testing process, while reference and measured force values are displayed on the screen in real time. A simplified control-oriented dynamic model of a pneumatic cylinder controlled by a proportional pressure valve has been derived. The process parameters were analysed and experimentally identified. Based on the dynamic model of the process, the parameters of the PID controller were determined and the experimental results for the task of tracking the sinusoidal reference force were obtained. Satisfactory results of tracking the reference force were obtained for sinusoidal signals of lower frequencies, but due to the slow dynamics of the control components when tracking sinusoidal references of higher frequencies, there were significant delays in tracking the reference force. However, for dynamic testing of materials, the realization of a pulsating response is required, and the accuracy of tracking the periodic signal is a secondary criterion. In further research, some more advanced control methods will be applied to explore the possibilities of improving the dynamic response and accuracy of the system. To avoid readjustment of the controller for each test specimen (depending on the stiffness of the test specimen), a state controller with parameter estimation will be applied. The developed system is completely open for upgrading new user applications, which is quite difficult with available commercial systems. Also, the price of used components of approximately 5,000 euros is many times lower than similar available systems offered on the market, especially those with hydraulic actuators.

NOMENCLATURE

A_1 , A_2	Cylinder piston area and rod end area	m^2
e	Force control error	N
F_c	Spring force of the test specimen	N
F_p	Applied force on the cylinder piston	N
F_R	Reference force	N
$G(s)$	Transfer function	-
K_f	Speed estimation filter	-
K_p , K_i , K_d	PID controller gains	-
K_v	Forward gain of the proportional valve	Pa/V
k_e	Stiffness constant	N/m
b	Viscous friction coefficient	Ns/m
m	Mass of the piston and measuring equipment	kg
p_1 , p_2	Pressures in chamber 1 and chamber 2	Pa
p_a	Atmospheric pressure	Pa
s	Laplace operator	-
T_v	Time constant	s
u	Control signal	V
x	Cylinder position	m

REFERENCES

[1] Gela T (2006) Mechanical Design Handbook, Chapter 6: Properties of Engineering Materials. McGraw-Hill Book Company

[2] Murugan S (2020) Mechanical Properties of Materials: Definition, Testing and Application. Int. J. of Modern Studies in Mechanical Engineering (IJMSME) 6, 2, 28-38

[3] Costa P, Nwawe R, Soares H, Reis L, Freitas M, Chen Y, Montalvão D (2020) Review of Multiaxial Testing for Very High Cycle Fatigue: From 'Conventional' To Ultrasonic Machines. Machines, 8, 25; doi:10.3390/machines8020025

[4] Ghielmetti C, Ghelichi R, Guagliano M, Ripamonti F, Vezzù S (2011) Development of a fatigue test machine for high frequency applications. Procedia Engineering 10, 2892–2897. https://doi.org/10.1016/j.proeng.2011.04.480

[5] McAlorum J, Rubert T, Fusiek G, Niewczas P, Zorzi G (2018) Design and Demonstration of a Low-Cost Small-Scale Fatigue Testing Machine for Multi-Purpose Testing of Materials, Sensors and Structures. Machines, 6(3), 30, https://doi.org/10.3390/machines6030030

[6] Marques J M E, Benasciutti D, Niesłony A, Slavič J (2021) An Overview of Fatigue Testing Systems for Metals under Uniaxial and Multiaxial Random Loadings. Metals, 11, 447. https://doi.org/10.3390/met11030447

[7] Šitum Ž, Benić J, Pejić K, Bača M M, Radić I, Semren D (2021) Design and control of mechatronic systems with pneumatic and hydraulic drive, Int Conf Fluid Power 2021. Maribor, September 16-17, 2021, Maribor, Slovenia, 179-193.

[8] Technical documentation Electro-Pneumatic Regulator/Electronic Vacuum Regulator ITV Series, SMC ITV 2050-312N3, https://static.smc.eu/pdf/ITV-G_EU.pdf

[9] Benić J, Rajčić N, Šitum Ž (2017) Precise force control for hydraulic and pneumatic press system, Int Conf Fluid Power 2017. Maribor, September 14-15, 2017, Maribor, Slovenia, 57-71.

PRODUCT CARBON FOOTPRINT OF HYDRAULIC AND PNEUMATIC COMPONENTS – CHALLENGES IN ACCOUNTING AND COMPARABILITY

peer reviewed

Johannes Sprink[1]*, Katharina Schmitz[1]

[1] *RWTH Aachen University – Institute for Fluid Power Drives and Systems (ifas), Campus-Boulevard 30, 52074 Aachen*

* Corresponding author: Tel.: +49 241 80 47709; E-mail address: Johannes.Sprink@ifas.rwth-aachen.de

ABSTRACT

To achieve global climate protection goals, a reduction of greenhouse gas emissions is necessary. Accordingly, emissions from technical products such as fluid power components must be known over the entire Product Life Cycle. The Product Carbon Footprint (PCF) is used to systematically record the greenhouse gas emissions of a product.

For the calculation of a PCF, for each step in the life cycle, the inputs of material and energy have to be identified and analyzed. Within the context of a cradle-to-gate analysis, the focus is on the production phase and upstream processes. In a previous study carried out by the authors, a variety of fluid powered components were evaluated in regards of their greenhouse gas emissions during the production. Challenges that can arise in such accountings are presented in this paper. For example, in many cases the use of database factors is unavoidable due to the lack of primary data. However, determining of exact factors poses a problem that cannot be neglected. It is shown that even with similar preconditions extreme deviations in the results can occur. Examples for such results will be presented and explained.

Conclusions are drawn that may also be helpful when accounting products. Furthermore, recommendations are presented on how to deal with the calculation results of third parties.

Keywords: Product Carbon Footprint, Sustainability, Greenhouse Gas Emissions, Cradle-to-Gate

1. INTRODUCTION

The effects of climate change are now ubiquitous. It is reflected in the increase in temperatures and extreme weather conditions. These will occur even more frequently in the future [1]. Climate change is thus becoming one of the most significant risk factors for humankind and is determined mainly by the concentration of greenhouse gases (GHG) in the atmosphere. To counteract this, it is necessary to emit fewer greenhouse gases. In the Paris Climate Agreement, 190 countries agreed to the goal of limiting global warming to 1.5°C compared to pre-industrial times [2]. Other international agreements lay down further obligations in this regard. At the European level, for example, there is the EU Green Deal, in which the European Union commits itself to be ultimately climate neutral by 2050 [3]. In addition, the individual member states have, in some cases, made even more far-reaching commitments. The Federal Republic of Germany, for example, sets even stricter targets for annual emissions with its Climate Protection Act 2021 [4].

Industrial processes have a significant share in the emission of greenhouse gases [5]. The various regulatory measures mentioned are therefore aimed in particular at the manufacturing sector. In order to achieve savings, a systematic and uniform measurement of emissions is first necessary.

Various methods are available for accounting these emissions, which are standardized in different norms. These are generally based on the life cycle analysis (LCA) approach. A number of ISO standards should be mentioned, with ISO 14067 providing details on the balancing of greenhouse gases. The GHG Protocol standards are also widely used in the economy.

The fluid power industry, as a supplier of products for use in energy-intensive machinery, is of particular importance, as also shown in [6] and [7]. For exemplary products from the fluid power industry, cradle-to-gate studies regarding the GHG emissions were conducted. The usage phase of the products and the end-of-life treatment were not inspected for different reasons. For example the utilization cycles of fluid powered components vary greatly from one application to the next and are not known to the manufacturer.

In this paper it will first be described what kind of difficulties can arise during the accounting, because uniform accounting is not easy to achieve. This concerns, for example, the choice of uniform system boundaries, the level of detail of data collection and the choice of applicable factors from databases. Furthermore, the results of different accountings will be compared and discussed. Finally suggestions on the application of the findings of this study will be made.

2. ACCOUNTING OF GHG-EMISSIONS

GHG emissions during the life cycle of a product can be divided into three different scopes. Scope 1 includes a company's direct emissions. For example, these can result from the use of company vehicles or the use of fossil fuels in manufacturing processes or building operations. Scope 2 includes indirect emissions from the provision of purchased energy. All other emissions that occur indirectly during the life cycle of a product are included in scope 3. Depending on the stage of the life cycle, Scope 3 emissions are referred to as upstream or downstream. The Upstream includes all processes by third parties before the reporting company becomes involved. Downstream emissions are only considered in a cradle-to-cradle analysis and include, for example, energy consumption during the use of sold products and end-of-life treatment of products. An illustration of these scopes has been created by the GHG Protocol as shown in **Figure 1**.

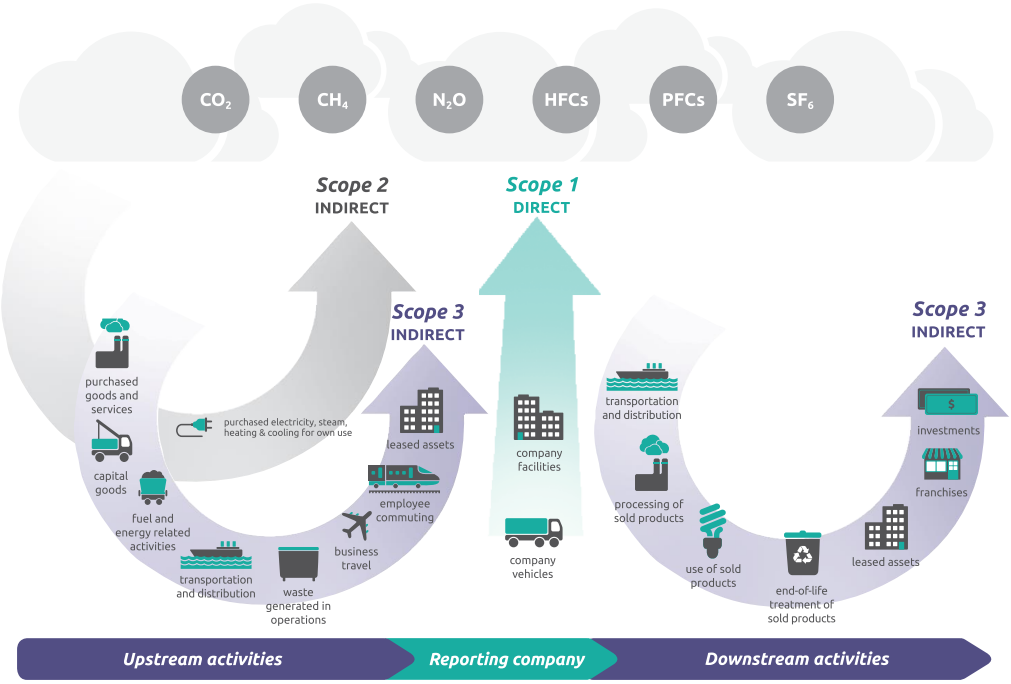


Figure 1: Overview of GHG Protocol scopes and emissions across the value chain [8]

In the context of carbon emissions accounting, a major challenge lies in the collection of data. This task involves a series of crucial steps. First, a detailed flowchart should be created to illustrate the entire operational process. Subsequently, it is imperative to identify and document all process steps that could potentially contribute to emissions. This not only involves direct emissions, but also the indirect emissions. Therefore, any inputs to the process must be considered. This may be the use of energy but also the application of components and raw materials.

Concurrently, a comprehensive inventory of the materials and energy inputs utilized within the process is essential to gain a comprehensive view of emissions sources. The practical implementation often involves the retrieval of data from Enterprise Resource Planning (ERP) systems. But overall, the process of data collection demands considerable effort of the involved parties. Not only does it involve a small number of people in a specific department, but it also requires a lot of communication with each unit involved in the production subsequent treatment of a product.

GHG accounting over the lifecycle of components has already been studied for many fluid-powered components, as shown in [9] and [10]. At ifas a further study has been conducted to compare the accountings of a variety of fluid powered components in regards of their greenhouse gas emissions during the production. This includes different hydraulic and pneumatic valves and actors as well as hydraulic pumps. Each of these were produced by different manufacturers. The goal was to achieve a uniform accounting process of all involved companies. The challenges of a consistent approach are described in the following. It is shown that the results are often poorly comparable despite the fact that they are conform to the GHG Protocol.

3. CHALLENGES IN ACCOUNTING

To get a comparable result, it is crucial to have uniform procedures for the data collection. In practical terms, however, various difficulties arise in the application of the norms and standards described. On the one hand, these standards leave a lot of leeway in their application. For example, the system boundaries and thus the scope of data collection in a large area can be chosen individually by each company. On the other hand, this is necessary to ensure general validity so that applicability in different subject areas is possible. The intention is that, if necessary, even services can be compared with the production of equivalent products. However, without further guidelines that are specific to particular product categories, this simultaneously makes the comparability of results more difficult. It also follows that a numerical value as the result of an accounting alone is not sufficient to assess the result of an accounting process. In addition, it is necessary to prepare a report that contains additional information. The preparation of such a report is required by all standards. However, only the contents of the report are specified, not a uniform format.

Within the context of the research carried out, it showed that the largest share of greenhouse gas emissions within the cradle-to-gate analysis of the inspected components is accounted for by Scope 3 emissions. This includes all emissions by upstream suppliers. Especially the production of raw materials is responsible for a large share of the total emissions. Depending on the scope of the individual analysis, transport-related emissions typically fall within the range of 1 % to 10 %. Nevertheless, it should be noted that emissions in other categories are also of significant importance. This is primarily due to the fact that changes are occurring in all areas of the life cycle. For example, in the case of "Fossil-free steel" production, the replacement of coal and coke with hydrogen was observed, signifying a critical shift towards sustainability [11]. These findings underscore the multifaceted nature of emissions accounting, where changes and innovations across diverse domains collectively contribute to the overall trajectory of carbon emissions.

Databases

At present, no exact numerical values are available for many processes in the life cycle. The reasons for this are manifold; especially in the case of multi-layered supply chains, it is still difficult to obtain reliable information on greenhouse gas emissions from each party involved. Therefore, in practice, so-called emission factors have to be taken from databases, although the data obtained from these can at best be a rough estimate of real processes. The data sets stored there originate from different sources and are regularly adapted to changed boundary conditions. For example, a change in the electricity mix in the country of manufacture of a material affects the balance. Therefore, in today's global supply chains, large fluctuations can occur over time despite unchanged products. The greater the proportion of database values used, the greater the impact on the final result.

A temporal influence also arises separately from the use of database factors. In the case of regular deliveries of raw materials, it happens that suppliers obtain the same material from different sources. Since greenhouse gas emissions are strongly dependent on the country of origin, this also leads to a temporal change in the environmental impact.

In the course of the research, many numerical values were calculated using emission factors taken from databases. For example, no participant had reliable data on the raw material used. Freely accessible data sources have cost advantages, but in some cases only provide older or inaccurate data sets, in extreme cases even obviously wrong data sets. For specific material, researching specific studies has also proven helpful. However, this is time-consuming and the data obtained in this way often still have to be transferred into a comparable context.

As has been described, databases provide poor accuracy. Provided that the databases used are of sufficient extend, more accurate information about one's materials is helpful. For example, many databases break down the emission factor of raw materials by percentage of primary material and country of origin, allowing a better approximation to be made. However, primary data is preferable at all times in order to realistically represent the actual processes and thus obtain an accurate calculation result.

An example on database values is presented in **Table 1**, where values for milling of metal are displayed. These values are represented per kilogram of material to be removed. It is important to note that specific definitions distinguishing between large and small parts were not provided within the dataset. Furthermore, it was observed that there is a substantial variation in these values, which is contingent upon the size of the respective parts. Also, there may exist uncertainty regarding the comprehensiveness of the information encompassed within the database values, including aspects such as the selection of specific transport routes and material treatment methods.

Table 1: Example: Milling of metals – Database ecoInvent 3.8 [1]

	Average	Large parts	Small parts
Cast iron	0,148 kWh	0,0659 kWh	0,706 kWh
Steel	0,474 kWh	0,211 kWh	2,26 kWh

4. COMPARIBILITY OF CALCULATION RESULTS

In order to objectively compare the greenhouse gas emissions of similar products across manufacturers, it is necessary to have identical prerequisites. The aim of the research, which was carried out in cooperation with various manufacturers of pneumatic and hydraulic components, was to archive such a uniform accounting across different manufacturers. Exemplary components such as hydraulic and pneumatic valves and actuators, as well as hydraulic pumps, were examined and an accounting according to the GHG Protocol was performed. For various reasons, which are explained below, it was not possible to achieve reliable comparability, and a comparison has not yet been possible within the context of the research carried out.

The overall results show that the numerical values of the results reflect a different level of detail between the calculation results of the individual participants. In particular, scopes 1 and 2 are mapped in much greater detail than scope 3 upstream for most companies. This is due to the fact that the data on the companies' own processes is available to a higher degree of accuracy for all companies than the data from the suppliers, insofar as they were able to provide data at all. Scope 3 downstream was not investigated during the study, but might be even harder to account, because no standardized utilisation-cycles apply for all the specific applications of different customers.

It was not possible to determine final results for the emissions during the manufacturing phase of some products. For example, specific numerical values were not available yet. Therefore, assumptions were used for these in many places. In some cases, for example, no data is available on material processing in the supply chain. In these cases, the GHG Protocol specifies that conservative estimates be used for the missing data. This prevents a realistic comparison of the products in the individual categories.

For some products, certain necessary data is not known in detail. This concerns, for example, the power consumption of individual processing machines, for which assumptions and estimates have been made. In the absence of measuring instruments for individual processing machines, only the general consumption of a larger unit can be taken and allocated to all the products produced in that unit. As a result, only a less precise result can be achieved. Similarly, for many other products,

some small components whose weight is in the range of a few grams are not taken into account. Provided that an appropriate estimate of the impact on the final result predicts only an insignificant impact on the overall result, this disregard is consistent with the requirements of the GHG Protocol.

Examples on specific comparisons

Two internal gear pumps from different manufacturers were analysed as part of the research. A comparison of the two pumps is shown in **Figure 2**. Normalization to 100% of the value of pump A was performed. Pump B weighs approximately 4 times more than pump A. Nevertheless, the calculation result of pump B in comparison with pump A indicates a significantly lower greenhouse gas emission. However, this is most likely due to the different level of detail in the analysis. Irrespective of the assumptions made in database values, for example, processing efforts at the suppliers were explicitly recorded for pump A. The value of the raw material in relation to the product weight also differs by a factor of about 4 for both pumps. However, this is probably also due to the greater depth of observation. For example, experience shows that emission factors taken from databases for a general structural steel are significantly lower than factors for a precisely specified alloy of the material actually used.

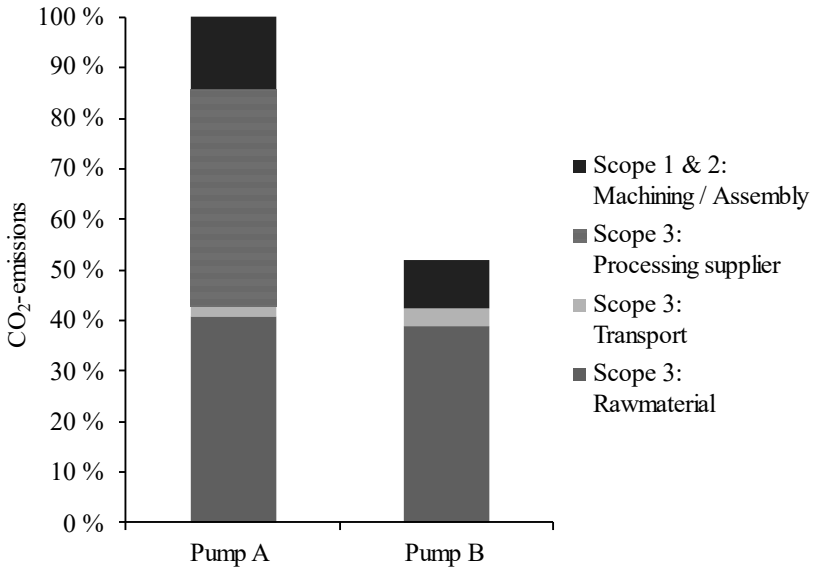


Figure 2: Comparison of internal gear pumps

A similar result is also found for pneumatic valves. **Figure 3** illustrates a comparison of the balancing results for two 5/2-way valves. It can be seen that there is also a large discrepancy between different manufacturers in this case Valve B is approximately 50% heavier as Valve A, but seems to emit 500% of the carbon emissions. The reasons for this again might be a more detailed analysis. In the case of valve B, for example, scrap treatment and surcharges for uncertainties are taken into account. The apportionment of general energy consumption of (electricity, heating, cooling) is also comprehensively taken into account, and even the heating of corridors and other production-related areas, for example, is included in the result of valve B. Corresponding values may optionally be added in accordance with the GHG Protocol, which represents a realistic balancing, but also leads to a higher result. In addition, the different prerequisites prevent a uniform comparison between different manufacturers.

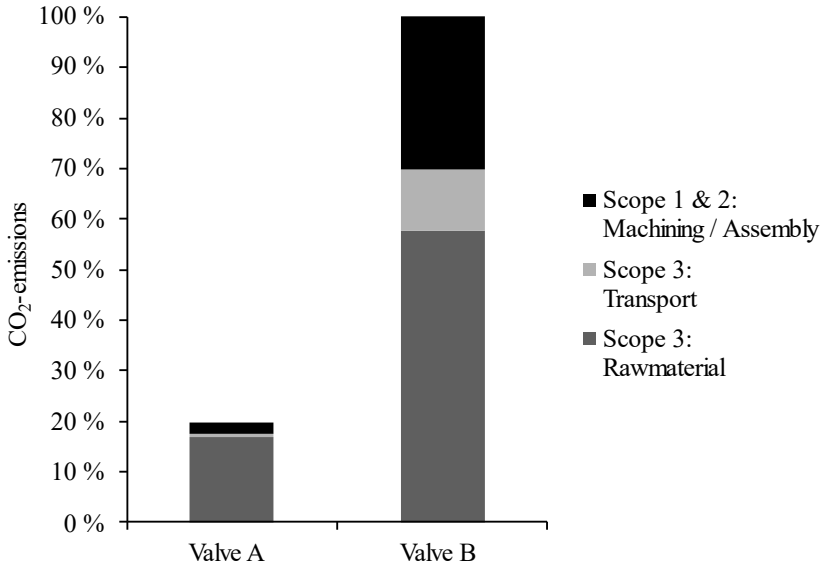


Figure 3: Comparison of pneumatic valves

5. CONCLUSION AND RECOMMENDATIONS

Based on the results shown, it can be seen that even with a fully standard-compliant procedure, a less intensive analysis can result in a lower numerical value for greenhouse gas emissions in the calculation. In this context, there is a risk that companies preparing their balance sheets will deliberately consider certain process steps less precisely in order to convey a positive impression of their own greenhouse gas balance. The intended estimation of unknown correlations can also easily be chosen in such a way that the results are presented more positively than they are. Particularly in competition with other companies, there is therefore a risk that the published figures do not reflect the real environmental impact.

An important step in communicating with suppliers is therefore to always request the corresponding report in addition to the balancing result. This is because the system boundaries selected and assumptions made must be documented in it, so that an insight into the calculation methodology is gained.

A further step towards uniform balancing can be taken by more specific regulations and standards. For example, a VDMA standard sheet has been drawn up in which more detailed approaches are prescribed for many points, which can also be applied in a practical manner in mechanical and plant engineering in particular. In addition, various characteristics are defined on the basis of which different quality criteria can be derived. By specifying them, further transparency can be created and a faster assessment of the result can be made. The creation and application of specific Product Category Rules (PCRs) in accordance with DIN EN ISO 14025 can also help to achieve standardization.

However, trust in business partners will remain an important criterion for assessing published emission values in the foreseeable future. Certification of a company's accounting methodology by neutral parties can have a positive influence here. However, neither more detailed standardization nor certification can completely prevent differences in the results of different accounting methods.

The necessary reduction in greenhouse gas emissions can therefore only be achieved by working together to make all industrial processes as climate-neutral as possible.

ACKNOWLEDGEMENT

The authors thank the Research Association for Fluid Power of the German Engineering Federation VDMA for its financial support. Special gratitude is expressed to the participating companies and their representatives in the accompanying industrial committee for their advisory and technical support.

NOMENCLATURE

ERP Enterprise Resource Planning
GHG Greenhouse Gas
LCA Life Cycle Analysis
PCF Product Carbon Footprint
PCR Product Category Rule

REFERENCES

[1] Masson-Delmotte, V et al. „Climate change 2021: the physical science basis. “ Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change 2 (2021).

[2] „Paris Agreement“, United Nations Treaty Collection, Chapter XXVII 7. d, 2015

[3] Europäische Kommission, „Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen – Der europäische Grüne Deal“, Brüssel, 11.12.2019

[4] Bundes-Klimaschutzgesetz vom 12. Dezember 2019 (BGBl. I S. 2513), geändert durch Artikel 1 des Gesetzes vom 18. August 2021 (BGBl. I S. 3905)

[5] Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (BMUV), „Novelle des Klimaschutzgesetzes beschreibt verbindlichen Pfad zur Klimaneutralität 2045“, 12.05.2021

[6] Love, Lanke, Alles, „Estimating the impact (energy, emissions and economics) of the US fluid power industry“. Oak Ridge National Laboratory, Oak Ridge, TN, 2012

[7] Radermacher, Merx, Sitte, Boyko, Unger „Potenzialstudie Energie- / Kosteneinsparung in der Fluidtechnik“ 2021

[8] Greenhouse Gas Protocol “Corporate Value Chain (Scope 3) Standard”, World Resources Institute, 2011

[9] Merkelbach, S., „Analysis of the economic and ecological properties of pneumatic actuator systems with pneumatic transformers.” Lehrstuhl und Institut für fluidtechnische Antriebe und Steuerungen, 2020

[10] Zhang, Y. Zhang, and M. Y. Cai, “Overall life cycle comprehensive assessment of pneumatic and electric actuator,” (eng), Chinese Journal of Mechanical Engineering, vol. 27, no. 3, pp. 584–594, 2014

[11] AB Volvo “Volvo Group and fossil-free steel” https://www.vol-vogroup.com/en/news-and-

media/news/2022/jun/volvo-group-and-fossil-free-steel.html (letzter Zugriff: 01.12.2022)

[12] Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., and Weidema, B., 2016. The ecoinvent database version 3 (part I): overview and methodology. The International Journal of Life Cycle Assessment, 21(9), pp.1218–1230.

EXERGY ANALYSIS FOR THE INTERMITTENT AIR SUPPLY IN PNEUMATIC MACHINES

Dominik Gryboś[1]*, Jacek Leszczyński[1]

[1] *AGH University of Krakow, Faculty of Energy and Fuels, Department of Department of Thermal and Fluid Flow Machines, Av. Mickiewicza 30, 30-059 Krakow, Poland*

* Corresponding author: Tel.: +48 12 617 33 89; E-mail address: dgrybos@agh.edu.pl

ABSTRACT

Pneumatic systems are widely used to automate production lines in manufactured plants. Their big disadvantage is their low energy efficiency, 10-20%. It is mainly due to the overconsumption of compressed air by the oversized pneumatic actuator and other components. Reducing air consumption at the utilisation stage can result in significant exergy savings. This can be done by lowering the supply pressure, introducing back pressure, or using expansion energy in the actuator. The most promising is the last method that can be implemented using the intermittent air supply in the pneumatic actuator. The literature lacks an exergy analysis of the utilisation stage of pneumatic and intermittent air supply systems. There is also no optimisation of the control algorithm of intermittent air supply control algorithm in terms of minimising exergy consumption. In this paper, we demonstrate a mathematical model and exergy analysis as a tool for assessing the efficiency of the utilisation of pneumatic system and conduct computer simulation. Exergy analysis showed that for intermittent air supply the reduction in exergy consumption decreased by more than 60% compared to the classic oversized system. The results of the computer simulation give the opportunity to optimise the operation of the utilisation stage in pneumatic systems. Furthermore, exergy analysis can be a useful tool for energy analysis and assessment of pneumatic systems, as well as providing information on the desired direction of changes in the installation.

Keywords: Compressed air system, pneumatics, air use, intermittent air supply, Exergy analysis, sensitivity analysis

1. INTRODUTION

Compressed air systems (CAS) (pneumatic systems) are widely used in numerous industrial sectors such as light, food, paper, plastic and automotive. It is estimated that pneumatic systems consume 10% of the electricity consumed in the industrial sector in the United States, the EU and China. The use of compressed air to power machines and production lines is characterised by high power density, low investment costs, simple expansion, and high reliability [1]. However, pneumatic systems have very low energy efficiency, approximately 5%-10%, cite [2]. Furthermore, the compressed air system is characterised by poor controllability [3] and high noise [4]. Energy losses in CAS occur during compression, transmission due to leaks [5] and the utilisation stage through excessive air consumption [6]. Methods to improve efficiency in compressors and air leaks are well described in the literature and technology [7–9]. However, methods for reducing compressed air consumption are less well presented and can achieve an increase in energy efficiency by up to 40%.

The oversizing of pneumatic actuators is the main reason for the overconsumption of compressed air by CAS due to the design and selection process of pneumatic actuators and their standardisation [10, 11]. The pneumatic actuator is selected for at least 33% greater or even 50% greater force

generated from the load and a piston speed condition of 0.3 to 0.5 m/s [10, 12]. For operational reliability and safety, it is customary to select a one-size larger pneumatic actuator. The degree of actuator oversizing can be determined by the pneumatic frequency factor presented by [13]. Additionally, dead volumes in system components also result in excessive energy consumption of compressed air [14].

In the literature, there is a preliminary classification of methods for reducing air overconsumption: reducing the supply pressure [11, 15], increasing the back pressure [16, 17] or using the expansion power in the actuator [18]. This results directly from the reduction of the compressed thermodynamic cycle field [19]. Reducing the supply pressure is the most commonly used method by using pressure reducers or, rarely, a meter in systems. The potential energy savings of compressed air range from 20% to even 40% [11, 16]. Increasing back pressure involves introducing a certain pressure greater than the ambient pressure at the outlet ports of the pneumatic actuator. As a result, the pneumatic actuator operates within the limit of the bottom pressure range. Back pressure air is stored in a buffer tank and its reuse allows for energy savings of 40-75%. In the commonly used algorithm for supplying compressed air to actuators, only air transmission power is used, while air expansion power is lost. The method of using expansion power is implemented through intermittent air supply and a directional control valve bridge system [18, 20]. Air expansion power is used through an appropriate algorithm controlling the suppling and cutting off of compressed air and monitoring the position of the piston and the pressure in the chambers during stroke. This method can achieve energy savings of up to 85% [18]. However, these systems require complex control algorithms, additional sensors, and a fast switching direction control valve [21]. Although this is the method that gives the greatest savings, the proper optimisation of compressed air supply is complex.

Assessment of the energy efficiency of a compressed air system is complex, and therefore it is advisable to use a system analysis approach. The first law of thermodynamics for a compressed air system (energy analysis) has limited application due to near-isothermal transformations and slight changes in air temperature and energy [22]. Much more adequate is the second law of thermodynamics and exergy analysis, which takes into account not only changes in temperature but also pressure and the irreversibility of changes. In the literature there is some static exergy analysis of the air compression and preparation stage [23–27]. These analyses lack an estimate of the actual consumption of compressed air, which is defined by the final devices in the utilisation stage. Rakova et al. [14] present a calculation of the accumulated exergy of compressed air in separate pneumatic components. Merkelbach and Murrenhoff [28] conducted experimental and computer comparisons of exergy consumption for several air savings configurations of the compressed air system and obtained a reduction in exergy consumption of 20-50%. The temporary consumption of compressed air results from the non-stationary transition of compressed air from state A to B, which is then repeated cyclically. Therefore, the tools presented are not suitable for a detailed analysis of processes at the utilisation stage.

In the paper, we show the exergy analysis for compressed air system and intermittent air supply. For this purpose, an original mathematical model of compressed air systems was used and computer simulations were performed. The level of exergy consumption is shown for various parameters that shape the air supply signal to the actuator. This tool has the ability to use an exergy criterion to optimise the operation of pneumatic systems seems to be the key novelty. This work focused on the exergy analysis of only the utilisation stage, and a mathematical model of system dynamics was implemented only for this stage.

2. METHODS

Industrial compressed air systems (see **Figure 1**) consist of three stages: compression and preparation, storage, and utilisation. Each CAS stage consists of characteristic elements: a compressor, a heat exchanger, a

dryer, and a filter in the air compression and preparation stage. Air tank with equipment in the storage stage and pneumatic lines, directional control valves, pneumatic actuator, and muffler. According to the theory of air power [29], it can be divided into transmission and expansion power, where only the former is used in classic control.

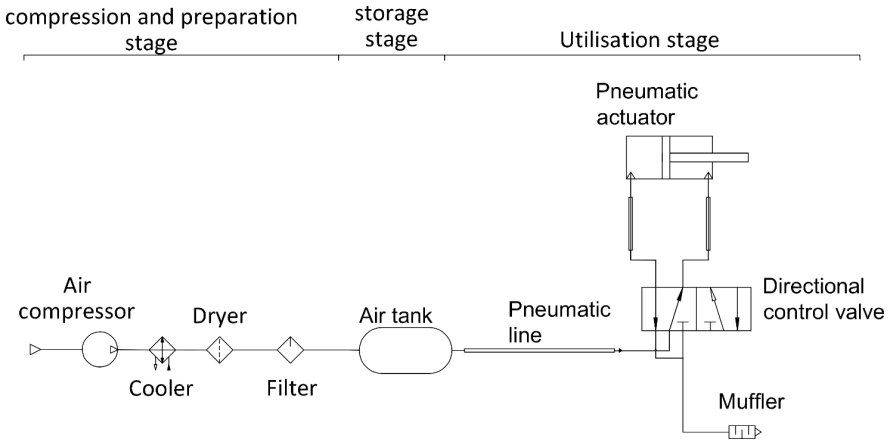


Figure 1: Compressed air system.

2.1. Mathematical model

A detailed description of the the mathematical model of individual components of the compressed air system in the utilisation stage is presented in [30]. The compression and utilisation stages in the compressed air system operate indirectly, independently of each other, with the buffer as a tank. Therefore, the work of individual stages can be considered separately with the connecting element, which is the tank. The mathematical model of the dynamic of the the utilisation stage of compressed air system is based on the flow of compressed air between certain volumes (constant – tanks, valves, mufflers) and variables (actuator) through the transmission line. For each volume, a system of three differential equations should be written describing the mass balance (1), the energy balance (2) and, for variable volumes, an additional equation for the dynamics of motion (3). An example system of equations for the i-th volume and where the index 1 is the inlet and 2 is the outlet:

$$\frac{\mathrm{d}m_i}{dt} = \dot{m}_{i,1} - \dot{m}_{i,2} \tag{1}$$

$$\frac{dT_i}{dt} = \frac{1}{m_i\, C_{v_i}} \dot{m}_{i,1}\left(h_{i,1} - h_i\right) - \dot{m}_{i,2}\left(h_i - h_{i,2}\right) + \dot{Q}_i - \dot{W}_i \tag{2}$$

$$\frac{d^2x_i}{dt^2} = A_1\left(p_{i,1} - p_{i,2}\right) - A_2 p_0 + F_c + F_{fr} \tag{3}$$

where: m is air mass in the volume, $\dot{m}$ is air mass flow, T is air temperature, C_{v_i} is specific heat capacity at constant volume, h is air enthalpy, $\dot{Q}$ is heat flux, $\dot{W}$ is work flux, x is piston position, A_1 is piston cross section, A_2 is piston bore cross section, p is air pressure, F_c is contact force, F_{fr} is Columbus-viscous friction force with Stribeck effect. In addition, the pressure equation (4) and the air mass flow equation should be written for each volume. More details about the developed model and its validation can be found in [30].

$$p_i = \frac{RT_i\, m_i}{V_i} \tag{4}$$

where: R is individual gas constant for air and V is volume. The air mass flow model is calculated by

formula presented by Leszczyński i Gryboś [31] where ζ_1 is the reduction coefficient related to flow contraction phenomenon and $\zeta_2$is geometry coefficient of outlet/inlet port.

$$\begin{cases} \zeta_1 A_i \sqrt{\frac{2\gamma}{\gamma-1}\rho p_1 \left(\left(\frac{p_2}{Zp_1}\right)^{\frac{2}{\gamma}} - \left(\frac{p_2}{Zp_1}\right)^{\frac{\gamma+1}{\gamma}}\right)} & \text{for } \frac{p_2}{p_1} > \beta \\ \zeta_1 A_i \sqrt{\gamma\rho p_1 \left(\frac{2}{\gamma+1}\right)^{\frac{\gamma+1}{\gamma-1}}} & \text{otherwise} \end{cases} \tag{5}$$

where: A_i is inlet/outlet port cross section, γ is polytropic exponent, ρ is air density and Z is scaling function. The ζ_2=1 means ideal smooth orifice edges and results in the critical pressure ratio β movement to lower values.

$$\beta = \zeta_2 \left(\frac{2}{\kappa+1}\right)^{\frac{\kappa}{\kappa-1}} \tag{6}$$

$$Z = \frac{1-\zeta_2}{1-\beta^{2\zeta_2}}\left(\frac{p_1}{p_2}\right)^{2\zeta_2} + \frac{\zeta_2-\beta^{2\zeta_2}}{1-\beta^{2\zeta_2}} \tag{7}$$

For a pneumatic muffler, the air mass flow is calculated using the Ergun formula [32] as a phenomenon of flow through a porous surface. The outlet mass flow $\dot{m}_{\mathrm{i},2}$ from a pneumatic can be determined as the positive square root of the quadratic function:

$$c_3 \cdot \left(\dot{m}_{i,2}\right)^2 + c_4 \cdot \left(\dot{m}_{i,2}\right) - \frac{p_{i,2}-p_i}{\delta_i} = 0 \tag{8}$$

Where: where δ_{i} is thickness of the porous layer, c_3 is internal losses in porous material and c_4 viscous losses in porous material. The mathematical model of dynamic flow through the pneumatic line based on one-dimensional ordinary differential equations for pressure and mass flow for each segment [11, 33]:

$$\frac{dp_j}{dt} = -\frac{p_j}{\rho_j A_j}\frac{\dot{m}_{i+1}-\dot{m}_i}{\Delta x} \tag{9}$$

$$\frac{d\dot{m}_\imath}{dt} = -A_i\frac{p_{j-1}-p_j}{\Delta x} - \frac{\lambda_i}{(\rho_{j-1}+\rho_j)A_i d_i}\dot{m}_\imath|\dot{m}_\imath| \tag{10}$$

2.2. Exergy analysis

The complete exergy balance is presented in [30]. The change in exergy over time of an air in a given control volume of a compressed air system is the sum of the exergy flow through this volume transported with mass $(\dot{X}_\imath^k(t) - \dot{X}_{\imath-1}^k(t))$, work $\dot{X}_{w_i}^k(t)$, heat $\dot{X}_{Q_i}^k(t)$ and exergy destruction $\dot{X}_{d_i}^k(t)$. The sign before the exergy term indicates the direction of exergy flow ("+" to the system, "-" from the system) and k is component pointer. Compressed air systems are open thermodynamic systems, therefore exergy equation should be:

$$\frac{dX_i^k(t)}{dt} = \dot{X}_\imath^k(t) - \dot{X}_{\imath-1}^k(t) + \dot{X}_{w_i}^k(t) - \dot{X}_{Q_i}^k(t) - \dot{X}_{d_i}^k(t) \tag{11}$$

In the exergy analysis of the complete operating cycle of the pneumatic system, understood as the extend and retract movement of the actuator. It is assumed that the final air parameters at each point of the system return to their initial values in the ideal case. Therefore, the actuator movement is for extend movement from t_b to t_e, (from t_b to t_m) and retract movement (from t_m to t_e). Furthermore, constant ambient conditions (reference conditions) were assumed throughout the entire analysis period and the kinetic and potential energy of the air in the system were neglected. Taking into

account the destruction exergy, the total change in exergy in a given volume over the complete actuator cycle from t_b to t_e is equal to zero $\frac{dX_i^k}{dt} = 0$. Therefore, the exergy of destruction in a given volume of the system can be defined as the balance of the remaining exergy streams in the system.

$$X_d^k = \int_{t_b}^{t_m} \left(\sum_{i=1}^{n} \left(\dot{X}_i^k(t) - X_{i-1}^{\dot{k}}(t) - \dot{X}_{Q_i}^k(t) + \dot{X}_{w_i}^k(t) \right) \right) dt + \int_{t_m}^{t_e} \left(\sum_{i=1}^{n} \left(\dot{X}_i^k(t) - X_{i-1}^{\dot{k}}(t) - \dot{X}_{Q_i}^k(t) + \dot{X}_{w_i}^k(t) \right) \right) dt \quad (12)$$

2.3. Computational setup

Computer simulations were carried out for two stages of utilisation of compressed air systems: classic and intermittent air supply, shown in **Figure 2**.

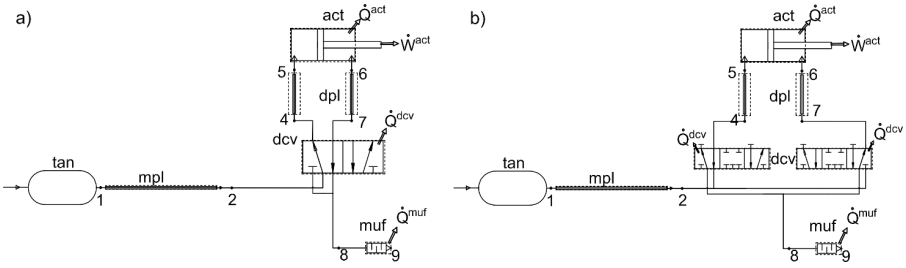


Figure 2: Stage of the compressed air system: a) Classical; b) Intermittent air supply. Where: tan – tank, mpl – main pneumatic line, dcv – directional control valve, dpl – distributed pneumatic line, act – actuator, muf – muffler.

The difference in the structure of the systems is the use of a directional control valve bridge in the case of intermittent air supply. The parameters of the system components and the process parameters are presented in **Table 1**.

Table 1: Parameters of the compressed air system

Parameter	Value
Supply gauge pressure	6.3 bar
Piston diameter	63 mm
Stroke	300 mm
Extend mass load	110 kg
Port size	G3/8
Main/distributed line diameter	12/8 mm
Main/distributed line length	5/2 m
Pneumatic frequency ratio	1.89

For intermittent air supply, a rigid algorithm was used to air supply to the actuator to examine the parameters of compressed air pulses and the sensitivity of the movement dynamics. The valve control signal (VOS) is based on the pressure signal (optional, can be omitted) in the chambers and the piston position and is defined as:

- For extend stroke

$$VOS == \begin{cases} 1 \text{ for } (x \le D) \cup (x \le a_1 \cap x \le (a_1 s + a_2 D)) \cup (x \le b_1 s \cap x \le (b_1 s + b_2 D)) \\ 0 \; for \; otherwise \end{cases} \quad (13)$$

- For retract stroke

$$VOS = \begin{cases} 1 \text{for} (x \leq D) \cup (p_2 > p_1) \\ 0 \; for \; otherwise \end{cases} \tag{14}$$

where x is piston position, p_2 is pressure in retract chamber, p_1 is pressure in extend chamber and D is duty factor:

$$D = \frac{\Delta l}{s} \tag{15}$$

where: s is actuator stroke and Δl is the distance at which the actuator is supplied by compressed air. Figure 3 shows the shape of the compressed air supply signal for the extend and retract movement.

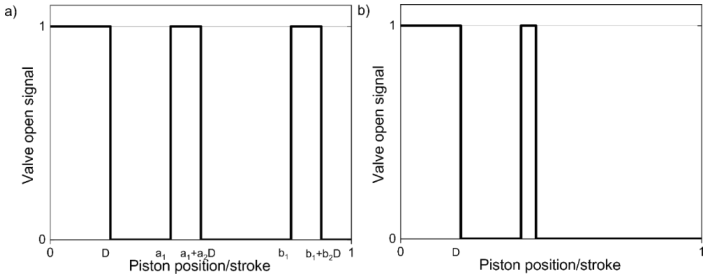


Figure 3: Valve open signal for: a) Extend stroke; b) Retract stroke.

Individual parameters defining the valve open signal are defined in **Table 2**.

Table 2: Sensitivity analysis of valve open signal

Parameter	Value
D	0.01/0.03/0.06
a_1	0.4
a_2	10/30/60
b_1	1
b_2	5

A sensitivity analysis was performed for 3 values of the D parameter and 3 values of the c parameter. A detailed model of the mathematical utilisation stage of the compressed air system was presented, implemented, and validated in [30]. The computer simulation was conducted in MATLAB~R2021b software with Runge-Kutty-Fehlberg solver. Furthermore, the thermodynamic library of moist air ASHRAE RP-1485, Thermodynamic Properties of Real Moist Air, Fog, Ice Fog, Dry Air, Steam, Water, and Ice, MATLAB version (FluidLAB) was used.

3. RESULTS

Computer simulations were carried out for a classic compressed air system and for an intermittent air supply with D = 0.01 and a_2 = 10. For such a case, the dynamics of the actuator operation was compared, and the changes in pressure, position, speed, and exergy flux were presented. The results for classical and intermittent air supply are presented in **Figure 4** and **Figure 5**. The use of an intermittent air supply instead of a classical system did not cause any disruptions in the actuator's operation. This means that the jump time remained practically unchanged and did not exceed 0.6 s, which means the average speed is more than 0.5 m/s. Additionally, during the extend movement, the piston speed increases faster, and the stroke remains constant for a longer period of time.

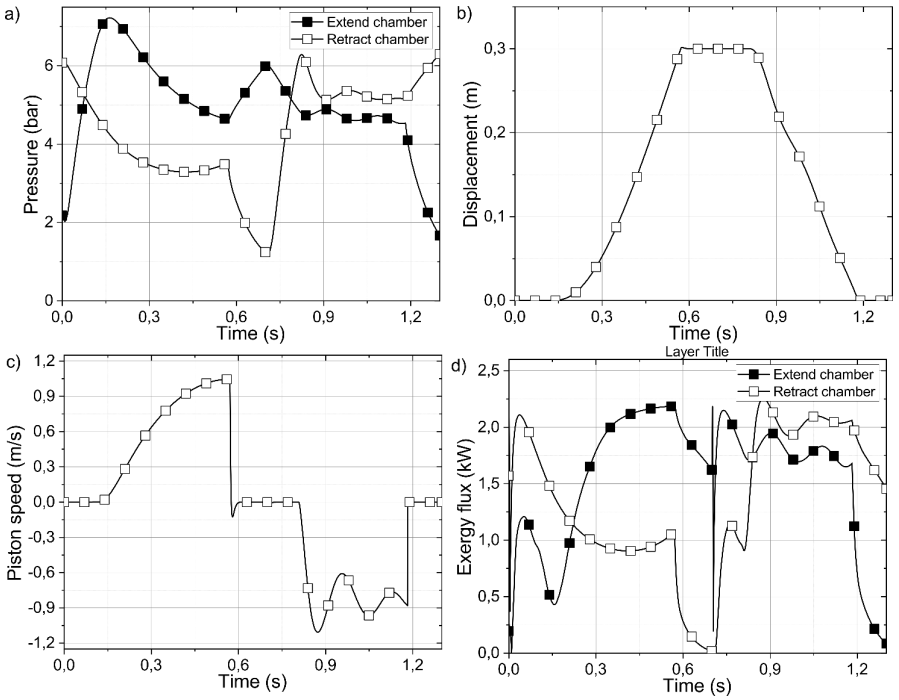


Figure 4: Pneumatic actuator dynamics for classical compressed air system: a) Pressure; b) Displacement; c) Speed; d) Exergy flux.

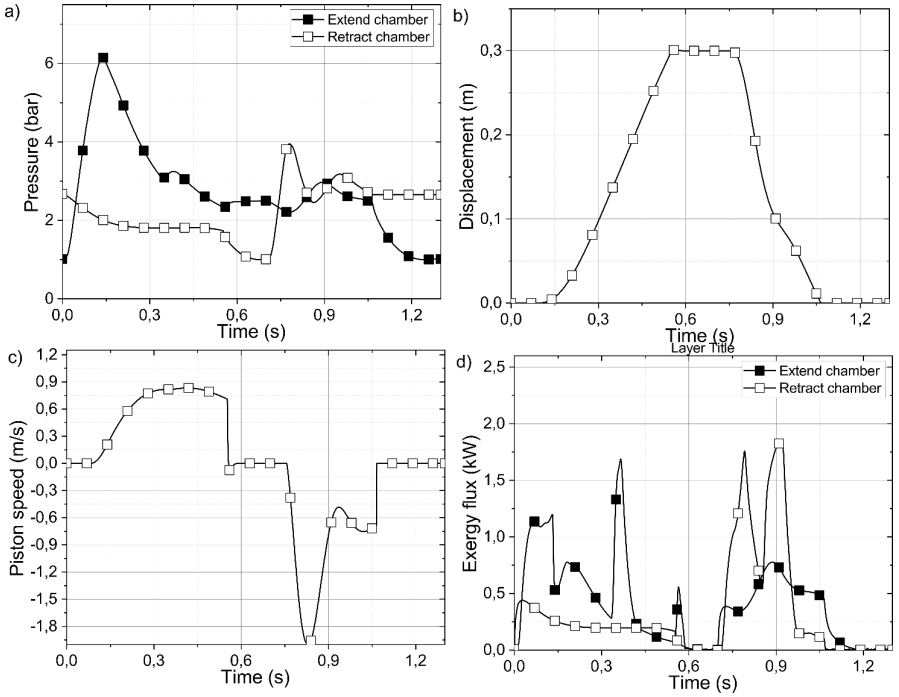


Figure 5: Pneumatic actuator dynamics for intermittent air supply: a) Pressure; b) Displacement; c) Speed; d) Exergy flux.

This is due, among other things, to the pressures that prevail in both chambers. In the case of intermittent air supply, the pressure during the retract movement is lower than the pressure in the network, and therefore the actuator needs less time to reload the harmful volumes. It also has a positive effect on the exergy flux, which is even 5 times smaller than in the classical case. The retract movement, although it is without load in both cases, is performed faster for intermittent air supply. **Figure 6** shows the Sankey diagram for the energy consumption of the classic system and intermittent air supply.

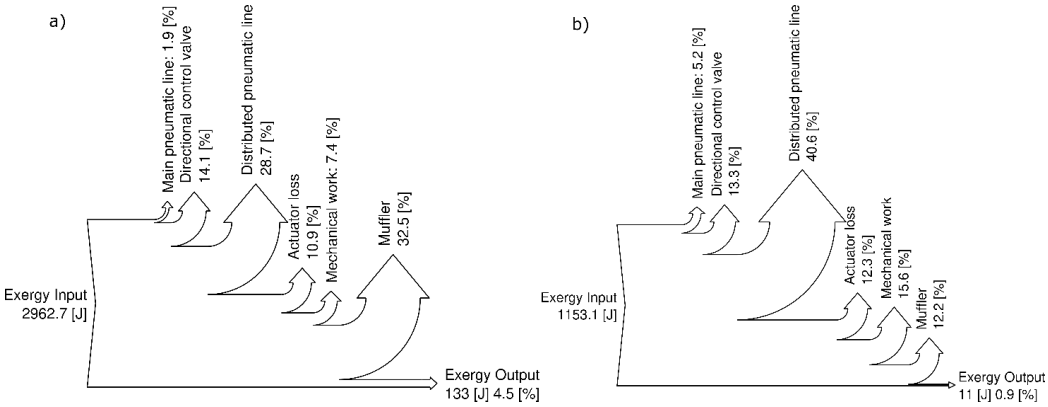


Figure 6: Sankey diagram: exergy flow in a) Classical; b) Intermittent air supply.

Exergy consumption decreased from 2962.7 J for the classical system to 1153.1 for intermittent air supply. These are savings in exergy consumption of 61%, which is caused by the compensation of dead volumes and the use of expansion air power during actuator operation. The burden of exergy losses has been transferred from the muffler and output to the distributed pneumatic line, and the exergy used in the actuator has been increased relative to the input exergy.

Then, a sensitivity analysis was performed for the shape of the valve open signal in terms of parameters D and a_2 (see Table 2). **Figure 7** shows the impact of changing these parameters on actuator dynamics.

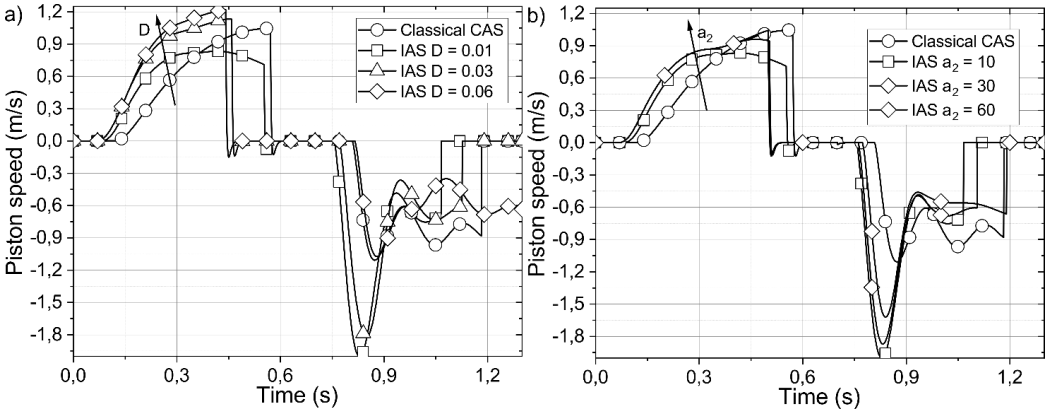


Figure 7: Sensitivity analysis of the piston speed in intermittent air supply.

As the D parameter increases, the actuator speed reaches higher values and becomes steeper. Additionally, the average speed of extend movement increases. However, the a_2 parameter affects the final speed of movement. With the larger parameter a_2, that is, with the extension of the second part of air supply, the final speed of the actuator is higher. This allows you to achieve the shock speed

parameters for the process. However, as shown in **Figure 8**, with increasing parameters *D* and *a2*, exergy consumption increases. For the parameters a_2 the increases are much smaller than for *D*.

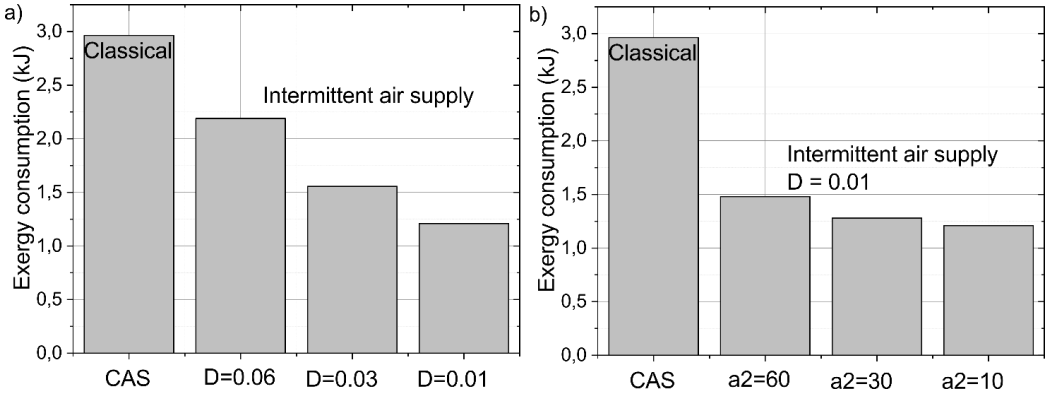


Figure 8: Sensitivity analysis of exergy consumption in intermittent air supply.

The results presented show that intermittent air supply provides large savings in exergy consumption exceeding 60-70%. However, exergy analysis can be very useful when programming the movement of an actuator piston via a valve open signal. By defining the limit conditions for the actuator movement, such as movement time, required average speed,, required final speed or return movement time, it is possible to programme intermittent air supply in such a way as to minimise exergy consumption in the entire system. Furthermore, exergy analysis can be a useful tool for energy analysis and assessment of pneumatic systems, as well as providing information on the desired direction of changes in the installation.

4. CONCLUSION

The paper presents an exergy analysis and a comparison of exergy consumption for the utilisation stage of an oversized classic CAES system and an intermittent air supply system. The intermittent air supply system consists in providing a discontinuous signal to open the air supply to the actuator. A brief description of the mathematical model and the exergy balance of the system is presented. Then, a sensitivity analysis was performed showing the influence of the valve open signal shape parameters on the piston speed profile and exergy consumption by the utilisation stage. Significant reduction in exergy consumption by the system during an actuator operation cycle by more than 60%. This is caused by the use of air expansion power and dead volume compensation in the system. Additionally, the piston speed profile has not changed significantly, and its course is even more favourable. In percentage terms, the highest exergy losses in intermittent air supply are on the distributed pneumatic line as dead volumes. By modifying the parameters of the open valve signal, we influence the speed profile and the energy consumption of the system. Exergy analysis can be used to optimise the engine in the intermittent air supply to minimise the exergy consumption of the system under given process constraints. The use of exergy analysis to optimise actuator operation and obtain air savings is not common. Therefore, preliminary results of the analysis of the sensitivity of the valve open signal to exergy consumption and dynamics will be further investigated. They will include optimisation of the movement of the actuator dynamics with the function of minimising exergy consumption.

ACKNOWLEDGEMENT

Research project partly supported by program ‘‘Excellence Initiative – Research University’’ for the AGH University of Krakow and research subvention supported by the Polish Ministry of Education

and Science, Poland grant number 16.16.130.942.

Supported by the Foundation for Polish Science (FNP)

NOMENCLATURE

A	Cross section	m^2
D	Duty factor	-
F	Force	N
V	Volume	m3
$\dot{Q}$	Heat flux	W
T	Temperature	K
$\dot{W}$	Work flux	W
$\dot{X}$	Exergy flux	W
X	Exergy	J
Z	Scaling function	-
a,b	Signal parameters	-
c	constant	-
h	Specific enthalpy	kJ/kg
$\dot{m}$	Mass flow	kg/s
m	Mass	kg
p	Pressure	Pa
s	Stroke	m
t	Time	s
x	Position	m
Δ	Interval	-
β	Critical pressure ratio	-
γ	Adiabatic exponent	-
δ	Thickness	m
ρ	Density	kg/m^3
ζ	Scaling factor	-

CAS	Compressed Air System
IAS	Intermittent Air Supply
VOS	Valve Open Signal

REFERENCES

[1] Cai M, Kawashima K, Kagawa T (2006) Power Assessment of Flowing Compressed Air. J Fluids Eng 128:402. https://doi.org/10.1115/1.2170129

[2] Saidur R, Rahim NA, Hasanuzzaman M (2010) A review on compressed-air energy use and energy savings. Renewable and Sustainable Energy Reviews 14:1135–1153. https://doi.org/10.1016/j.rser.2009.11.013

[3] Saravanakumar D, Mohan B, Muthuramalingam T (2017) A review on recent research trends in servo pneumatic positioning systems. Precis Eng 49:481–492. https://doi.org/10.1016/j.precisioneng.2017.01.014

[4] Czopek D, Gryboś D, Leszczyński J, Wiciak J (2022) Identification of energy wastes through sound analysis in compressed air systems. Energy 239:1–19. https://doi.org/10.1016/j.energy.2021.122122
[5] Dindorf R, Wos P (2018) Automatic measurement system for determination of leakage flow rate in compressed air pipeline system. Metrology and Measurement Systems 25:159–170. https://doi.org/10.24425/118170
[6] Nehler T (2018) Linking energy efficiency measures in industrial compressed air systems with non-energy benefits – A review. Renewable and Sustainable Energy Reviews 89:72–87. https://doi.org/10.1016/j.rser.2018.02.018
[7] Nourin FN, Espindola J, Selim OM, Amano RS (2022) Energy, Exergy, and Emission Analysis on Industrial Air Compressors. J Energy Resour Technol 144:. https://doi.org/10.1115/1.4051682
[8] Vittorini D, Cipollone R (2016) Energy saving potential in existing industrial compressors. Energy 102:502–515. https://doi.org/10.1016/j.energy.2016.02.115
[9] Doner N, Ciddi K (2022) Regression analysis of the operational parameters and energy-saving potential of industrial compressed air systems. Energy 252:124030. https://doi.org/10.1016/j.energy.2022.124030
[10] Fleischer H (1995) Manual of pneumatic system operation, 1st ed. McGraw-Hill, New York San Francisco Washington
[11] Beater P (2007) Pneumatic Drives. Springer Berlin Heidelberg, Berlin, Heidelberg
[12] Leszczyński JS, Gryboś D (2019) Compensation for the complexity and over-scaling in industrial pneumatic systems by the accumulation and reuse of exhaust air. Appl Energy 239:1130–1141. https://doi.org/10.1016/j.apenergy.2019.02.024
[13] Doll M, Neumann R, Sawodny O (2015) Dimensioning of pneumatic cylinders for motion tasks. River Publishers 16:11–24. https://doi.org/10.1080/14399776.2015.1012437
[14] Rakova E, Weber J (2015) Process simulation of energy behaviour of pneumatic drives. Procedia Eng 106:149–157. https://doi.org/10.1016/j.proeng.2015.06.018
[15] Harris P, O'Donnell GE, Whelan T (2012) Energy Efficiency in Pneumatic Production Systems: State of the Art and Future Directions. In: Proceedings of 19th CIRP International Conference on Life Cycle Engineering. pp 363–368
[16] Hepke J, Weber J (2013) Energy saving measures on pneumatic drive systems energy consumption of pneumatic drives. In: The 13th Scandinavian International Conference on Fluid Power, SICFP2013. Linkoping University Electronic Press, pp 475–483
[17] Hepke J, Weber J (2012) Improving Energy Efficiency of Pneumatic Handling System. In: Proceedings of The 7th International Fluid Power Conference. pp 1–5
[18] Doll M, Neumann R, Sawodny O (2011) Energy efficient use of compressed air in pneumatic drive systems for motion tasks. In: Proceedings of 2011 International Conference on Fluid Power and Mechatronics. IEEE, pp 340–345
[19] Cai M, Kawashima K, Kagawa T (2006) Power assessment of Flowing Compressed Air. Transaction of ASME 128:402–405. https://doi.org/doi:10.1115/1.2170129
[20] Raisch A, Hulsmann S, Sawodny O (2018) Saving Energy by Predictive Supply Air Shutoff for Pneumatic Drives. In: 2018 European Control Conference (ECC). IEEE, pp 965–970
[21] Pfeffer A, Glück T, Schausberger F, Kugi A (2018) Control and estimation strategies for pneumatic drives with partial position information. Mechatronics 50:259–270. https://doi.org/10.1016/j.mechatronics.2017.09.012
[22] Brol S, Czok R, Mróz P (2020) Control of energy conversion and flow in hydraulic-pneumatic system. Energy 194:116849. https://doi.org/10.1016/J.ENERGY.2019.116849

[23] Krichel S, Sawodny O, Hülsmann S, et al (2012) Exergy flow diagrams as novel approach to discuss the efficiency of compressed air systems
[24] Bader, W. T.; Kissock JK (2000) Exergy Analysis of Industrial Air Compression. IETC - Industrial Energy Technology Conference 89–98
[25] Harris P, Nolan S, O'Donnell GE, Meskell C (2013) Optimising Compressed Air System Energy Efficiency - The Role of Flow Metering and Exergy Analysis. In: Re-engineering Manufacturing for Sustainability. Springer Singapore, Singapore, pp 129–134
[26] Taheri K, Gadow R (2017) Industrial compressed air system analysis: Exergy and thermoeconomic analysis. CIRP J Manuf Sci Technol 18:10–17. https://doi.org/10.1016/j.cirpj.2017.04.004
[27] Tong Z, Wang H, Xiong W, et al (2021) Accumulated and transient exergy analyses of pneumatic systems with isochoric and isobaric compressed air storage tanks. Energy Storage 1–10. https://doi.org/10.1002/est2.269
[28] Merkelbach S, Murrenhoff H (2015) Exergy based analysis of pneumatic air saving measures. ASME/BATH 2015 Symposium on Fluid Power and Motion Control, FPMC 2015. https://doi.org/10.1115/FPMC2015-9513
[29] Yu Q, Cai M (2011) Research on expansion ratio of air power engine system. Proceedings of 2011 International Conference on Fluid Power and Mechatronics, FPM 2011 987–990. https://doi.org/10.1109/FPM.2011.6045906
[30] Gryboś D, Leszczyński J (2023) Exergy analysis of pressure reduction, back pressure and intermittent air supply configuration of utilization/expansion stage in compressed air systems. Energy 285:129419. https://doi.org/10.1016/j.energy.2023.129419
[31] Leszczyński JS, Gryboś D (2019) Compensation for the complexity and over-scaling in industrial pneumatic systems by the accumulation and reuse of exhaust air. Appl Energy 239:1130–1141. https://doi.org/10.1016/j.apenergy.2019.02.024
[32] Zhong W, Xu K, Li X, et al (2016) Determination of pressure drop for air flow through sintered metal porous media using a modified Ergun equation. Advanced Powder Technology 27:1134–1140. https://doi.org/10.1016/j.apt.2016.03.024
[33] Krichel S V., Sawodny O (2014) Non-linear friction modelling and simulation of long pneumatic transmission lines. Math Comput Model Dyn Syst 20:23–44. https://doi.org/10.1080/13873954.2013.811268

Chapter 12

New and Special Applications-II

DIGITAL REDUNDANCE FOR COMPACT SUBSEA ELECTRO-HYDROSTATIC ACTUATORS USING SENSOR FUSION

M Eng. Joao Pedro Duarte da Silva[1,2]*, D. Eng. Ali Emad[1], D. Eng. Alexandre Orth[1], Prof. Victor Juliano De Negri[2], Guilherme Prudente da Silva[1,2].

[1] *Bosch Rexroth AG*
[2] *Federal University of Santa Catarina*

* Corresponding author: Tel.: +49 9352 18-5170; E-mail address: JoaoPedro.DuartedaSilva2@boschrexroth.de

ABSTRACT

Compact electro-hydrostatic actuators (EHAs) offer a promising solution for subsea production with their cost-effective and energy-efficient design, combining the benefits of electromechanical and electro-hydraulic systems. However, adapting these compact EHAs to fit within the limited space of traditional subsea systems poses a challenge, particularly in maintaining system reliability. This study introduces a Digital Twin (DT), composed of a physical EHA model and multiple Kalman Filters for parameter estimation, aimed at creating digital redundancies for critical sensors. The effectiveness of this approach was validated using co-simulation with Dymola software, where a simulation model emulated both the Plant (Real Twin), as a Modelica model, And a mathematical model as a software object (Digital Twin). The results demonstrate reliable digital redundancies for position and load measurements, with minor deviations that are within acceptable limits.

Keywords: Compact Electric-Hydrostatic Actuator, Fault tolerance, Digital Twin, Subsea Valve Actuator.

1 INTRODUCTION

The transition from conventional hydraulic and electro-hydraulic to electro-mechanical apparatus in subsea oil & gas exploitation fields is a growing trend, referred to as the "All-Electric Subsea" approach. This shift offers multiple advantages including reduced installation (CAPEX) and operation costs (OPEX), quicker system response, increased energy efficiency, reduced umbilical cable diameter (due to the elimination of hydraulic lines), enhanced operational flexibility, and environmentally sustainable design [1][2]. Electro-hydrostatic actuators (EHAs) are integral to this approach, amalgamating the benefits listed above with those of traditional electro-hydraulic systems, such as compactness, robustness, high power density, high load capacity, and effective overload protection with fail-safe functions performed by springs [3][4].

A significant challenge in implementing EHAs – self-contained control systems with numerous components – is the requirement to accommodate these within the space that is traditionally occupied by subsea hydraulic actuators. Figure 1 illustrates the various modules comprising a rotary electro-hydrostatic subsea valve actuator, designed to fit within the dimensions of a conventional hydraulic cylinder [4].

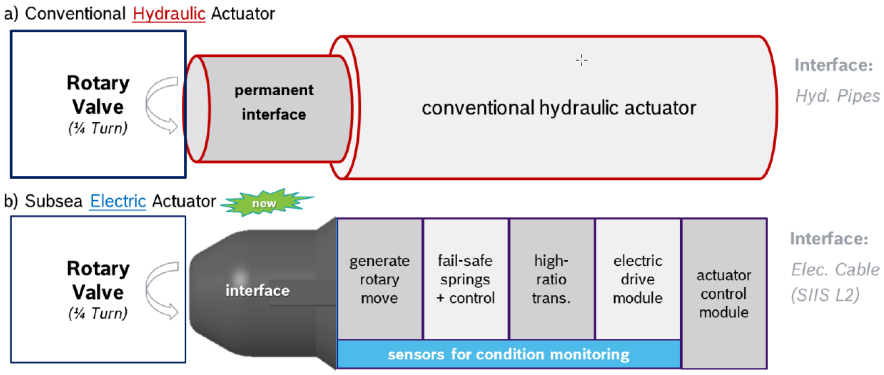


Figure 1 - Comparison of components integrated in a new Subsea rotary EHA [7].

The limited space in these intricate systems restricts the inclusion of redundant components, calling for alternative approaches to improve system reliability. Orth *et al*. [7] and Placido Neto *et al.* [8] highlight a significant challenge in applying EHA technology to replace hydraulic Subsea Valve Actuators (SVAs) where the new designs must fit into the space occupied by conventional units. This spatial limitation can restrict system redundancies, necessitating alternative approaches to bolster reliability.

To tackle this issue, a Digital Twin (DT) – a high-fidelity simulation model integrated into the actuator's controller [6] – presents itself as an effective solution. Capable of accurately reflecting the actual system's behavior, it acts as a digital backup for essential sensors. To achieve this, the DT must continuously update its parameters using sensor signals and control actions from the system, stepping in to substitute the feedback from a faulty sensor in case of failure.

2 SUBSEA VALVE ACTUATOR

The focal physical system modeled in this study is an electro-hydrostatic Subsea Valve Actuator (SVA) depicted in Figure 1. The SVA is engineered for the operation of rotary small-bore valves at water depths reaching up to 4,000 meters, designed concisely to match the size and weight of traditional subsea hydraulic actuators [7]. It not only accommodates a standard electric interface but also seamlessly integrates essential components, including electric drives and controls. Moreover, the actuator is equipped with an embedded system to facilitate motion control and offers a communication interface compliant with the Subsea Instrumentation Interface Standardization (SIIS) - specifically, a fault-tolerant CANOpen (SIIS Level 2) - ensuring reliable communication for operators.

According to Kritzinger *et al.*[6], a Digital Twin must ensure that any alteration in the physical object's state is mirrored in the digital object's state. This necessitates the establishment of a well-defined set of system elements that constitute the Digital Twin, thereby creating a control volume. This control volume defines the scope of the system covered by the Digital Twin, essentially forming a control system. A critical aspect of this process is the mapping of the system's inputs and outputs, which forms the basis for continually updating the information within the Digital Twin.

The Subsea Valve Actuator (SVA) illustrated in Figure 1 and detailed in Figure 2 is designed to operate rotary small-bore valves at depths of up to 4,000 meters [7]. The application software, housing the Digital Twin, is tailored to meet various normative, proprietary, and customer requirements, enabling the actuator to perform multiple functions. These functions, partially outlined in Placido Neto *et al.* [8], include (1) Preload fail-safe mechanism, (2) Command open of the process valve, (3) Command close of the process valve, (4) Emergency shutdown of the process valve (Fail-safe closing), and (5) Hold position.

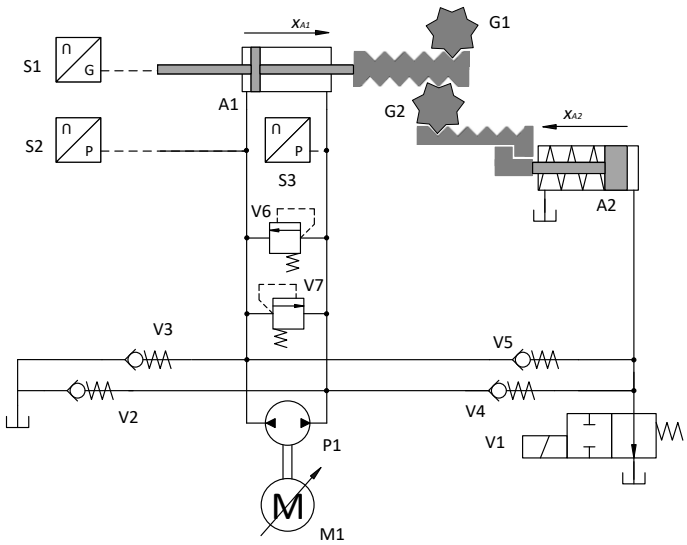


Figure 2 - Exemplary hydraulic circuit diagram

In the operation of the SVA R2, its Digital Twin must accurately replicate the behavior of its physical components. This requires the development of a robust mathematical model capable of adapting to the unique characteristics of the physical system, including manufacturing variances, environmental uncertainties, and component degradation. The details of this mathematical model and the methods for its continual updating are discussed in the following sections.

3 DIGITAL TWINS

Introduced in 2003 at the University of Michigan's Executive Course on Product Lifecycle Management (PLM) [5], the concept of a Digital Twin (DT) has evolved, yielding various definitions encompassing diverse systems and applications [9][10]. Rosen *et al*.[11] describe a DT as a highly realistic model reflecting the real-time state and behavior of a system or process interacting with the real world. According to Grieves (2014), a DT comprises three elements: physical products in real space, virtual products in virtual space, and the data connections binding the two.

Borangiu *et al*. [14] enumerate the primary benefits of DTs, including visibility (enabling operation visualization of individual equipment to larger systems), prediction (facilitating future state forecasting through modeling techniques applied to DT models), interaction (allowing condition simulation for "what if" scenario analyses impractical with physical prototypes), documentation (providing insight into the behavior of components), and integration (implementing DT models for connection with backend business applications in large systems).

However, the proliferation of definitions, along with related concepts like digital simulation, Cyber-Physical Systems (CPSs), and the Internet of Things (IoT), often results in confusion [12]. Some works between 2016 and 2019 referred to models as DTs even though they lacked essential DT attributes [13]. To clarify, Kritzinger *et al.* [6] propose classifying DTs into three subcategories: Digital Models, Digital Shadows, and Digital Twins, each representing different degrees of data integration and interaction between the physical and digital entities:

- Digital model: a digital representation of an existing or intended physical object that does not use any form of automatic data exchange between physical and digital objects. The communication between both objects is done manually. The digital representation may be more or less accurate in relation to the physical object;
- Digital shadow: In the digital shadow, there is already an automatic one-way flow of data between the state of the existing physical object and the digital one. A change in the state of the physical object leads to a change of state in the digital object, but not in a opposite way;
- Digital twin: data flow between the existing physical object and the digital object is fully integrated into both directions. The digital object can also act as a control instance for the physical object.

Kritzinger *et al*. [6] underscore that a Digital Twin must reflect changes in the physical object state with corresponding changes in its digital counterpart. Consequently, it is imperative to delineate a precise set of system elements comprising the Digital Twin. This process effectively constructs a control volume, delineating the scope of the system encompassed by the Digital Twin and forming a control system framework. Essential to this endeavor is the accurate mapping of the system's inputs and outputs, serving as the foundation for information updates within the Digital Twin.

4 SENSOR FUSION

According to Liu and Ma [15], sensor fusion is about combining information from different sensors to get a complete picture of an area or object. For example, in autonomous vehicles (AVs), various

sensors are placed in different spots to help the vehicle understand its surroundings from all angles. Sensor fusion becomes really important when there is a lot of information coming from different directions. It helps identify the same object seen by different sensors and puts together all the information from these sensors to create a big, unified picture of the area being sensed.

In the Industrial Internet of Things (IIoT) context, sensor fusion is an approach also applied to predict faults in industrial equipment where the data from multiple sensors is combined in order to aid the continuous operation without interruptions [16]. The sensor fusion approach is also applied in hydraulic power systems in the detection of faults, as presented by Zhong *et al.* [17], Chen *et al.* [18], and Shi *et al.* [19]. Kalman filter algorithms are also strongly used in sensor fusion approaches as shown in Kheirandish *et al.* [20] and Demirci *et al.* [21].

For hydraulic systems of medium and/or high complexity, the application of a sensor fusion technique becomes more challenging, since the correlation between sensor signals can or not happen depending on the states of individual components. For example, the current in a HPU electrical motor is proportional to the system pressure just while the motor is activated. Therefore, the causality between components, depending on the system's state must be known, for a proper sensor fusion algorithm to be implemented.

In addition, to create digital redundancies for critical sensors, which is to replace the feedback signal of the fault sensor with an estimated value based on a model behavior and other sensors, it is imperative that the Kalman filter algorithm is applied to update parameters of the system, so the model can be able to work similarly to the real system even when a parameter is not updated anymore. Therefore, a finite state model of the SVA R2, as shown in Figure 3 was developed. For each state the correlation between signals is obtained through a causality analysis. Fundamentals about finite state modeling can be seen in Dathan and Ramnath [22].

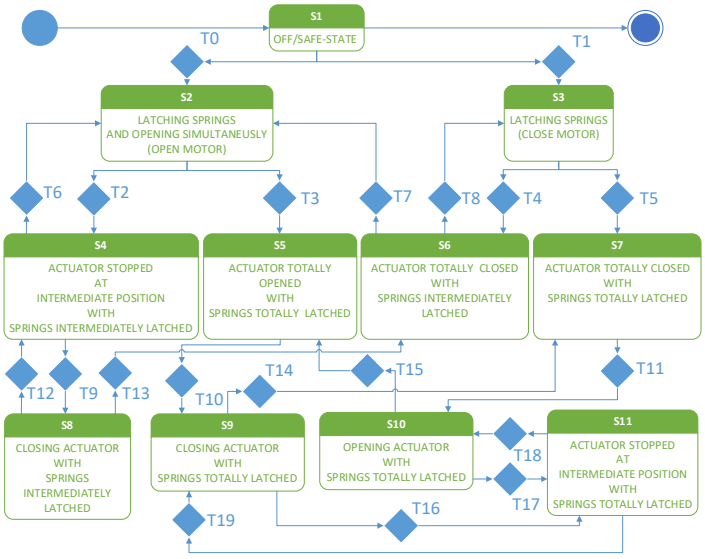


Figure 3 - State transition diagram for the SVA R2

In Figure 3 , S1 to S11 are the possible defined states of the actuator and T0 to T19 are the transitions between these states. Then, by knowing the states of the actuator, the parameter estimation algorithm uses the causality between elements to update the Digital Twin parameters.

5 DIGITAL TWIN MODELLING

There are several ways of building the structure of dynamic systems models, such as Physical Models, where is system is divided into subsystems whose behaviors are known, Identification Models, where observations from the real system are used to fit in a chosen model structure, or even neural networks [23] [24] [25] [26]. Dynamic models are also divided into deterministic and stochastic models, being deterministic if they present an exact relationship between measurable and derived variables and work without uncertainty or stochastic if they present uncertainty or probability concepts [24].

The objective of this work was to develop a physics-based Digital Twin, with a parameter estimation system to update the DT main parameters based on sensors' readings and control actions. Hence, state-space modeling, or dynamic linear modeling, presented itself as a fit solution for its development, once it is a general modeling structure that allows the combination of deterministic equations with uncertainty components [27].

As given by Shumway & Stoffer [27], the basic form of the state-space model employs an order one, n-dimensional vector autoregression as in

$$x_k = \Phi_k x_{k-1} + w_k \quad (1)$$

$$y_k = H_k x_{k-1} + v_k \quad (2)$$

where y_k is a vector of observable variables *n x 1*, x_k is a vector of unobservable variables *m x 1* called state variables, H_k is a matrix of known coefficients and v_t is a white noise *n x 1* vector, called measurement error, with covariance $\mathrm{E}(e_k e_k') = R$, Φ_k is an *m x m* matrix of autoregressive coefficients and w_t is a white noise *m x 1* vector with covariance $\mathrm{E}(\varepsilon_k \varepsilon_k{}') = Q$, called state error.

Inputs variables may enter into the states or into the observations, where a *r x 1* inputs vector is added to equations 1 and 2, as

$$x_k = \Phi_k x_{k-1} + \Upsilon u_k + w_k \quad (3)$$

$$y_k = H_k x_{k-1} + \Gamma u_k + v_k \quad (4)$$

where Υ is a $n\ x\ r$ input matrix and Γ is a $q\ x\ r$ feedforward matrix.

In order to obtain a modular design of the Digital Twin program (see Dathan *et al.* [22]), composed of individual modules, which are the mathematical models of individual components or component assemblies of the system, a modular state-space modeling based as described by Wang *et al.* (2018) and Yang and Wang [28] is chosen to be implemented.

6 PARAMETERS ESTIMATION USING KALMAN FILTER

The Kalman filter has been applied in several areas to describe the evolution of dynamic systems, where its main objective is to update the knowledge of a system each time a new observation is brought in [30] [31]. The algorithm objective is to find estimates of unobservable variables based on related observable variables through a set of equations called a state space model, which were already depicted in Equations 1 and 2.
In this work, the Kalman filter was implemented to estimate the parameters of the Digital Twin, instead of states. As shown in Liu [32] and Grewal and Andrews [33], the Kalman filter can be implemented to simultaneously estimate states and parameters by the addition of the observed parameters in the state vector. As an example, assuming two parameters a and b, included in the equation and as follows

$$x_{k+1} = f(x_k, u_k, a_k) + w_k \tag{5}$$

$$y_k = g(x_k, b_k) + v_k \tag{6}$$

$$a_k = a_{k-1} \tag{7}$$

$$b_k = b_{k-1} \tag{8}$$

The implemented algorithm was written as described in Costa & Alpuim [30]. Below, the iterative algorithm for updating the DT pump volumetric efficiency is depicted. During states 9 and 10, see Figure 3, the algorithm produces an estimator of the parameter η_{v_k} at each time k, based on the information up to time $k-1$, as given by

$$\hat{\eta}_{v_{k|k-1}} = \hat{\eta}_{v_{k-1|k-1}} \tag{9}$$

when at time k, η_{v_k} is available, the prediction error or innovation v_k is given by

$$v_k = \eta_{v_k} - \hat{\eta}_{v_{k|k-1}} \tag{10}$$

η_{v_k}, in steady-state, is calculated through

$$\eta_{v_k} = \frac{A_{A1}\frac{dx_{A1}}{dt}}{D_{P1}\omega_{P1}} \tag{11}$$

where A_{A1}[m^2] and x_{A1}[m] are respectively the cylinder A1 area and position and D_{P1}[m^3/rad] and ω_{P1}[rad/s] the pump P1 volumetric displacement and angular velocity. The error v_k is used then to estimate $\hat{\eta}_{v_k}$ with the equation

$$\hat{\eta}_{v_{k|k}} = \hat{\eta}_{v_{k|k-1}} + K_k v_k \tag{12}$$

where K_k is called the Kalman gain and is given by

$$K_k = P_{k|k-1}\left(P_{k|k-1} + R\right)^{-1} \tag{13}$$

where R is the measurement error covariance. Additionally, the mean squared error MSE matrix of the updated estimator $\hat{x}_{k|k}$ represented by $P_{k|k}$ verifies is updated by

$$P_{k|k} = P_{k|k-1} + Q - K_k P_{k|k-1} \quad (14)$$

where Q is the so called state error covariance.

7 COO-SIMULATION SETUP

The Digital Twin model and parameter estimation algorithms were developed in C code using the aforementioned equations. The C program was then compiled as an external object using the software Dymola. In the Dymola program, an already developed model simulates the Real Twin where both the DT and Modelica model received the same inputs coming from the Controller. Figure 4 shows shows the Dymola model that interconnects the three different programs.

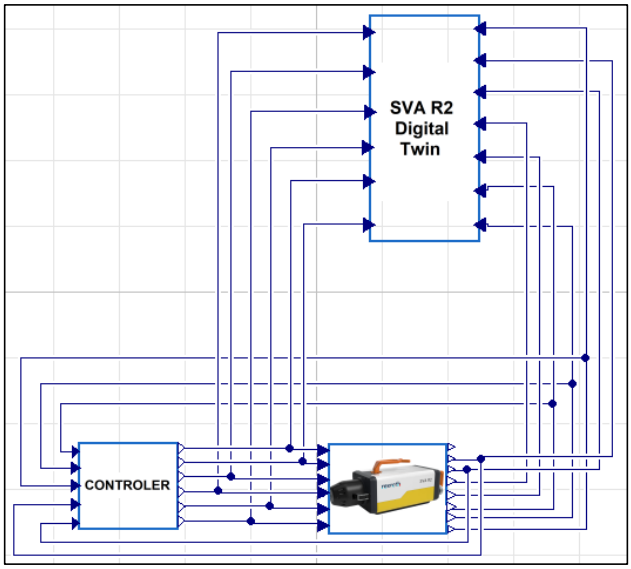


Figure 4 - Coo-simulation setup

8 RESULTS

With the coo-simulation setup developed, two different scenarios were simulated with the objective of observing the capacity of the Digital Twin parameter estimation program to adapt chosen parameters, e.g. volumetric efficiency and coulomb force, to the Real Twin and serve as a digital redundance.

8.1 Scenario 1: Regular actuation cycle with constant resistive load – No fault sensor

In the initial scenario, the effectiveness of the Digital Twin's parameter estimation algorithm is evaluated. For this test, the hydraulic pump P1 (Plant) in the Dymola model is set with a volumetric

efficiency of 87.7%, whereas the initial parameters of the Digital Twin are configured at 72%. Figure 5(a) displays the position sensor signal of the SVA (Plant) [°], the digital twin position [°], the SVA (Plant) position [°], and the difference between the Plant and Digital Twin positions [°], and Figure 5(b) presents the plant and tDT pump model volumetric efficiencies. The real value of the Plant position and its sensor are shown because in scenario 2 a fault in the position sensor will be simulated and the difference between the DT position and Plant position will be of interest.

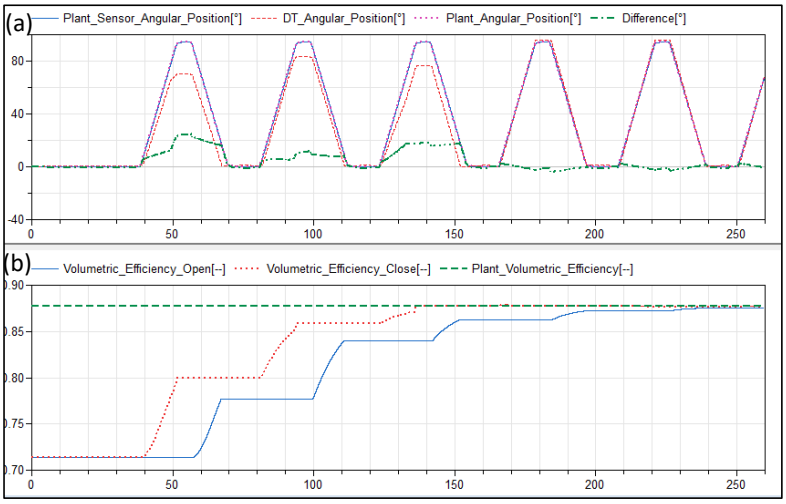


Figure 5 - *Scenario 1: Plant and Digital Twin positions as the Kalman filter parameter estimation algorithm updates the volumetric efficiency.*

It can be observed from Figure 5 (a) and (b) that the values of the open and closed volumetric efficiencies take 4 cycles to converge. At 250 seconds the difference on angular position between the plant and Digital Twin is 0.2% for open function and 0.04% for closing. Another aspect to be observed is the difference between Plant and Digital Twin positions, that has a maximum value of 24.8° for the 1st cycle and 3.36° for the 4th cycle.

It can also be observed that the parameters are just updated during the expected operational state, that means, the open and close volumetric efficiencies are updated, respectively, during the open and close operations, which are the states 10 and 9 in Figure 3. Another important variable to be monitored in a subsea valve actuator is the load torque required to operate the production valve. In this first scenario, a constant resistive load torque was applied. Figure 6 shows the plant and digital twin.

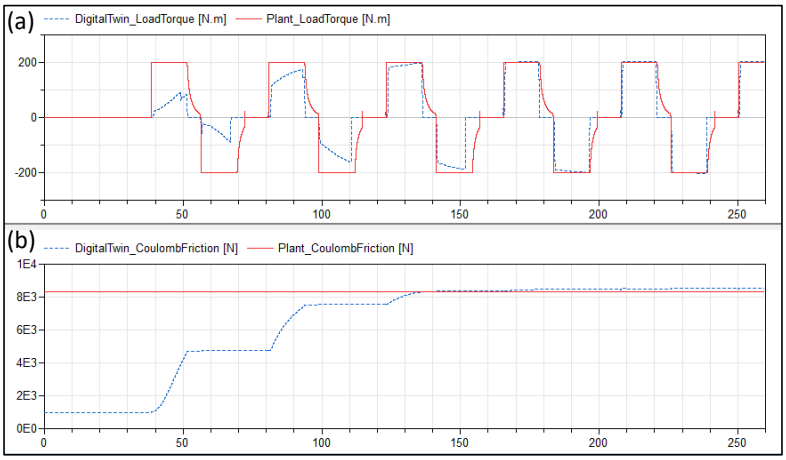


Figure 6 - Scenario 1: Plant and Digital Twin Load Torques.

The updated parameter presented in Figure 6 is the coulomb friction from the SVA load, which also updates as the cycles go forward, resulting in a torque difference of 3.5 Nm between Plant and Digital Twin model in the 4th cycle.

8.2 Scenario 2: Regular actuation cycle with constant resistive load – Fault sensor at 200 seconds

In scenario 2, a fault in the position sensor is simulated. In 200 seconds, the value of the position sensor is nullified. A *Stuck-at* detection algorithm (see Liu *et al.* [34]) is implemented to detect the faulty sensor.

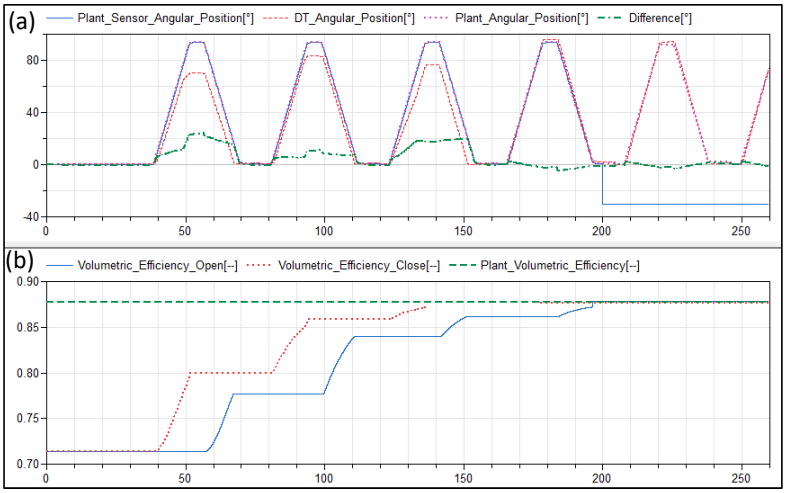


Figure 7 - Scenario 2: Plant and Digital Twin positions as the Kalman filter parameter estimation

algorithm updates the volumetric efficiency.

In Figure 7 it can be noticed that after 200 seconds, the value of the angular position sensor goes to -30°, which simulates that the sensor voltage output is zero. In this point the volumetric efficiency and load torque parameter estimation algorithm stop working and the Digital Twin position signal starts being used as feedback signal for the system controller. It can be observed by the comparison between the Plant angular position value and the Digital Twin Angular position, as well as the difference value, that the Controller is still able to control the plant while using the Digital Twin feedback signal.

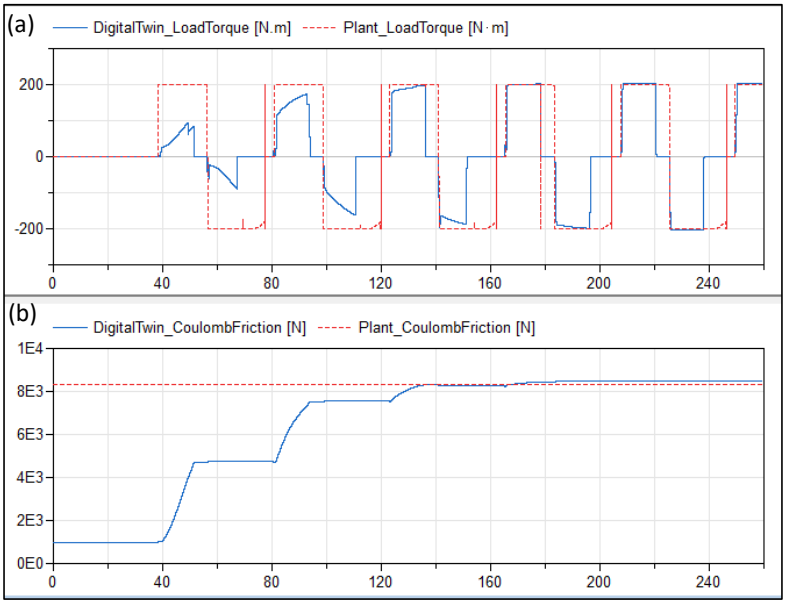


Figure 8 - Scenario 2: Plant and Digital Twin Load Torques.

Figure 8 demonstrates that the Digital Twin continues presenting a fit value regarding the real system, even after the parameter estimation algorithm is deactivated at 200 seconds.

9 FINAL REMARKS

The outcomes from the simulated scenarios show that the developed Digital Twin can successfully mimic its real-world counterpart's behavior offering a reliable level of confidence, given the required time for it to adapt its parameters as a function of the real system sensors readings and control actions. This enables tracking unmonitored variables and more importantly in the studied system, replacing malfunctioning sensors, e.g. position sensor, with digital versions that are less precise, but yet allow the system to continue its key functions with only minor dips in accuracy.

Additionally, this method can be applied to all actuated valves in a Christmas Tree subsea setup, significantly aiding in developing a comprehensive Christmas Tree Digital Twin for various subsea applications. Such widespread use enhances the detailed monitoring of equipment parameters, improving the overall condition assessment of the Christmas Tree subsea systems. The ability to enhance a fault tolerant automation system through software, without adding to its hardware complexity or incorporating extra physical components, is extremely beneficial for oil and gas subsea production or processing systems. It's also vital for cutting costs and boosting reliability and safety in emerging energy sectors like Carbon Capture Utilization and Storage systems, or offshore green hydrogen (H2) production.

REFERENCES

[1] Abicht D, Halvorsen G (2017) Ramberg R M. Subsea All-Electric. Offshore Technology Conference (OTC), Houston, Texas, USA.

[2] Mackenzie R, Halvorsen G, Henrik V (2020) Subsea All Electric – A Game Changing Technology Going Forward. Paper presented at the Offshore Technology Conference, Houston, Texas, USA. https://doi.org/10.4043/30515-MS.

[3] Weber J *et al*. (2016) Novel System Architectures by Individual Drives. 10th International Fluid Power Conference (10th IFK). Dresden, Germany. 2: 29-62 p.

[4] Orth A, Placido Neto A, Gottfried H (2022) Enabling All-Electric Subsea Control Systems without Compromising Safety - A Case Study Comparing Functional Safety Systems Using Springs or Batteries. Paper presented at the Offshore Technology Conference, Houston, Texas, USA. doi: https://doi.org/10.4043/32129-MS.

[5] Grieves M (2014) Digital twin: Manufacturing excellence through virtual factory replication Melbourne, FL: Florida Institute of Technology: White paper.

[6] Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital Twin in manufacturing: A categorical literature review and classification, IFAC-PapersOnLine, Volume 51, Issue 11, 2018, Pages 1016-1022, ISSN 2405-8963, https://doi.org/10.1016/j.ifacol.2018.08.474.

[7] Orth A, Placido Neto A, Gottfried H (2022) Enabling All-Electric Subsea Control Systems without Compromising Safety - A Case Study Comparing Functional Safety Systems Using Springs or Batteries. Paper presented at the Offshore Technology Conference, Houston, Texas, USA. doi: https://doi.org/10.4043/32129-MS.

[8] Placido Neto A. Duarte da Silva J P, Orth A, De Negri V J (2022) Development and Qualification of a New Rotary Subsea Valve Actuator for Small Bore Process Valves using Continuous Integration Approach. Paper presented at the 13th International Fluid Power Conference, 13. IFK, Aachen, Germany. Pages 1199 – 1209. Available in < http://www.hp-aachen.de/pdf/13th_IFK_Proceedings_Open-Access.pdf >.

[9] Glaessgen E H, Stargel D S (2012) The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. 53rd Structures, Structural Dynamics, and Materials Conference. Honolulu, Hawaii, USA, 23 - 26.

[10] Tao F, Xiao B, QI Q, Cheng J, Ji J (2022) Digital twin modeling, Journal of Manufacturing Systems, Volume 64, Pages 372-389,ISSN 0278-6125, https://doi.org/10.1016/j.jmsy.2022.06.015.

[11] Rosen R, Von Wichert G, Lo G, Bettenhausen K D (2015) About the importance of autonomy and digital twins for the future of manufacturing. IFAC-PapersOnLine28 (3), 567–572. https://doi.org/10.1016/j.ifacol.2015.06.141.

[12] Errandonea I, Beltrán S, Arrizabalaga S (2020) Digital Twin for maintenance: A literature review. *Computers in Industry*, Volume 123, 103316. ISSN 0166-3615. https://doi.org/10.1016/j.compind.2020.103316.

[13] Fuller A, Fan Z, Day C, Barlow C (2020) "Digital Twin: Enabling Technologies, Challenges and Open Research," in IEEE Access, vol. 8, pp. 108952-108971, doi: 10.1109/ACCESS.2020.2998358.

[14] Borangiu T, Oltean E, Răileanu S, Anton F, Anton S, Iacob I (2020) Embedded Digital Twin for ARTI-Type Control of Semi-continuous Production Processes. In: Borangiu, T., Trentesaux, D., Leitão, P., Giret Boggino, A., Botti, V. (eds) Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future. SOHOMA 2019. Studies in Computational Intelligence, vol 853. Springer, Cham. https://doi.org/10.1007/978-3-030-27477-1_9.

[15] Xiaohui Liu, Wei Ma (2022) Chapter 21 - Ubiquitous sensing for smart cities with autonomous vehicles, Editor(s): Amir H. Alavi, Maria Q. Feng, Pengcheng Jiao, Zahra Sharif-Khodaei, The Rise of Smart Cities, Butterworth-Heinemann, Pages 523-549, ISBN 9780128177846, https://doi.org/10.1016/B978-0-12-817784-6.00006-0.

[16] Deepak S, Anuj K, Nitin T, Sunil S C, Syam M P G (2023) Towards intelligent industrial systems: A comprehensive survey of sensor fusion techniques in IIoT, Measurement: Sensors, 100944, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2023.100944.

[17] Qi Z, Enguang X, Yan S, Tiwei J, Yan R, Huayong Y, Yanbiao L (2023) Fault diagnosis of the hydraulic valve using a novel semi-supervised learning method based on multi-sensor information fusion, Mechanical Systems and Signal Processing, Volume 189, 110093, ISSN 0888-3270, https://doi.org/10.1016/j.ymssp.2022.110093.

[18] Hongyue C, Hongyan C, Yajun X, Desheng Z, Ying M, Jun M (2022) Research on attitude monitoring method of advanced hydraulic support based on multi-sensor fusion, Measurement, Volume 187, 110341, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2021.110341.

[19] Jinchuan S, Jiyan Y, Yan R, Yong L, Qi Z, Hesheng T, Leiqing C (2021) Fault diagnosis in a hydraulic directional valve using a two-stage multi-sensor information fusion, Measurement, Volume 179, 109460, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2021.109460.

[20] Kheirandish M, Azadi Yazdi E, Mohammadi H, Mohammadi M (2023) A fault-tolerant sensor fusion in mobile robots using multiple model Kalman filters, Robotics and Autonomous Systems, Volume 161, 104343, ISSN 0921-8890, https://doi.org/10.1016/j.robot.2022.104343.

[21] Merve D, Haluk G, M. Cengiz T (2023) Improvement of power transformer fault diagnosis by using sequential Kalman filter sensor fusion, International Journal of Electrical Power & Energy Systems, Volume 149, 2023, 109038, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2023.109038.

[22] Dathan B, Ramnath S (2015). Object-Oriented Analysis, Design and Implementation. 10.1007/978-3-319-24280-4.

[23] Ellner S P, Guckenheimer G (2006) What Are Dynamic Models?. Dynamic Models in Biology, Princeton: Princeton University Press, pp. 1-30. https://doi.org/10.1515/9781400840960-004.

[24] Ljung L, Torkel G (1994) Modeling of Dynamic Systems. Prentice Hall Information and System Sciences Series. Englewood Cliffs, NJ: PTR Prentice Hall.

[25] Kim O D, Rocha M, Maia P (2018) A Review of Dynamic Modeling Approaches and Their Application in Computational Strain Optimization for Metabolic Engineering. Front Microbiol. https://doi.org/10.3389/fmicb.2018.01690.

[26] Pavlenko I, Trojanowska J, Ivanov V, Liaposhchenko O (2019) Scientific and Methodological Approach for the Identification of Mathematical Models of Mechanical Systems by Using Artificial Neural Networks. In: Machado, J., Soares, F., Veiga, G. (eds) Innovation, Engineering and Entrepreneurship. HELIX 2018. Lecture Notes in Electrical Engineering, vol 505. Springer, Cham. https://doi.org/10.1007/978-3-319-91334-6_41.

[27] Shumway R H, Stoffer D S (2017) Time Series Analysis and Its Applications. With R Examples. Fourth Edition. Springer International Publishing. ISBN: 978-3-319-52452-8. 2017.

[28] Yang D, Wang X (2020) Unified Modular State-Space Modeling of Grid-Connected Voltage-Source Converters in IEEE Transactions on Power Electronics, vol. 35, no. 9, pp. 9700-9715, Sept. 2020, doi: 10.1109/TPEL.2020.2965941.

[29] Wang Y, Wang X, Blaabjerg F, Chen Z (2017) Harmonic instability assessment using state-space modeling and participation analysis in inverter-fed power systems, IEEE Trans. Ind. Electron., vol. 64, no. 1, pp. 806–816, Jan. 2017.

[30] Costa M, Alpuim T (2010) Parameter estimation of state space models for univariate observations, Journal of Statistical Planning and Inference, Volume 140, Issue 7, Pages 1889-1902, ISSN 0378-3758, https://doi.org/10.1016/j.jspi.2010.01.036.

[31] Durbin, James & Koopman, Siem Jan, 2001. Time Series Analysis by State Space Methods. OUP Catalogue, Oxford University Press, number 9780198523543.

[32] Liu W (2013) Introduction to Hybrid Vehicle System Modeling and Control. 8 March 2013. SBN:9781118308400. DOI:10.1002/9781118407400.

[33] Grewal M, Andrews A (2001) Kalman filtering: theory and practice using MATLAB. New York: John Wiley and Sons. 14. 10.1002/9780470377819.

[34] Liu Y C, Lin C C, Tsai J J , Sun Y N (2013) Model-based spike detection of epileptic EEG data. Sensors (Basel). doi: 10.3390/s130912536. PMID: 24048343; PMCID: PMC3821325.

DEVELOPMENT OF A GENERIC TEST RIG FOR THE DETERMINATION OF THE INFLUENCE OF NON-NEWTONIAN FLUID PROPERTIES ON THE LEAKAGE CHARACTERISTIC OF ROTATING DISPLACEMENT PUMPS

Pascal Moor[1]*, Lara Kerres[1], Maximilian M.G. Kuhr[1], Peter F. Pelz[1]

[1]*Institute of Fluid Systems, Technische Universität Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt*

* Corresponding author: Tel.: +49 6151 1627112; E-mail address: pascal.moor@fst.tu-darmstadt.de

ABSTRACT

The efficiency of rotating displacement pumps is mainly influenced by internal leakage. When designing the pumps for Newtonian fluids, the manufacturers report satisfactory quality of predictive models. For the non-Newtonian fluids however, there are no suitable models that allow the manufacturers to predict the pumps behaviour or efficiency. Many of the manufacturers try to estimate the efficiency for non-Newtonian applications with their well-established models, which have however been developed specifically for Newtonian fluids. The efficiency is therefore overestimated by a significant amount. Especially non-Newtonian fluids with pseudoplastic and viscoelastic properties are of interest since the overlap of these properties has not been well researched. The most important blind spots in literature are the impact and interdependency of the moving boundary within the sealing gaps and the contraction flow for this type of fluid. The Chair of Fluid Systems addresses these blind spots with the development of a test rig for generic studies of the gap flow representative of the characteristic sealing gaps within pumps. The test rig is characterised by its modular construction which allows the efficient investigation of geometric gap parameters and fluid properties.

Keywords: non-Newtonian leakage, pump efficiency, rotating displacement pump, gap flow, contraction flow

1. INTRODUCTION

In order to improve their sustainability, both plant operators as well as component and system manufacturers seek to meet the desired function with best availability and acceptance while minimising efforts, specifically life cycle cost. The desired function for pumps is the supply of a required volume flow at defined pressure. In addition to the manufacturing costs, the life cycle costs for pumps depend mainly on operating costs, which depend on the efficiency of the pump in the specific application.

Rotating displacement pumps, e.g. rotary lobe pumps, gear pumps or screw pumps, are used throughout many industrial applications especially for pumping high viscosity fluids or satisfying high pressure demand. The fluids cover a range from gaseous to low-viscosity fluids, e.g. in the field of oil hydraulics, as well as high-viscosity media, e.g. for the food industry [1].

The achievable efficiency of displacement pumps is limited by internal leakage. The internal leakage mechanisms are inherent to the specific pump types. The clearance gaps between the components represent leakage paths [2]. A detailed prediction of the volumetric efficiency requires the knowledge of the leakage characteristics of a given fluid when passing through these tight sealing gaps. **Figure 1** shows characteristic sealing gaps for rotating displacement pumps.

While the behaviour of the leakage flow when pumping Newtonian fluids has been well understood [3, 4], the assessment of leakage flow of non-Newtonian fluids still presents a challenge. This is due to the complex rheological behaviour of the fluids.

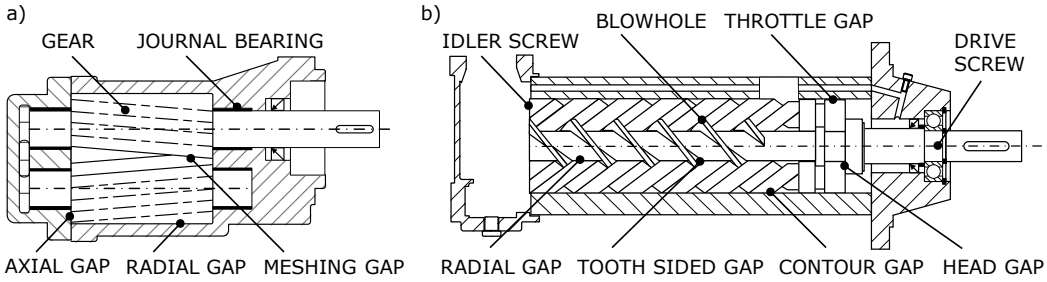


Figure 1: Gaps in a) an external gear pump, b) triple screw pump

The non-Newtonian fluids which are most relevant for the assessment of the gap flows can be categorised as shear-thinning fluids with viscoelastic properties. When experiencing shear stresses, the viscosity decreases due to the shear-thinning properties, which leads to an increase in leakage flow compared to Newtonian fluids with a viscosity similar to the zero-shear viscosity. Therefore, engineering the pumps according to the models for Newtonian fluids significantly underestimates the leakage flow. This effect may be superimposed by viscoelastic effects, which are known to be of importance e.g. for polymer solutions or melts. For the flow through small orifices or piping components, an increase in the pressure difference is documented that is linked to the normal stress effects. Similar effects can be expected for the flow across the contraction from the fluid chamber into the sealing gap, thus reducing the leakage. The superposition of this so-called extensional flow in the entrance to the sealing gap with the Couette-flow and the Poiseuille-flow that are relevant within the sealing gap, has not yet been subject to research.

Due to the lack of suitable models for the estimation of the leakage flow for non-Newtonian fluids, the prediction of the volumetric efficiency is subject to great uncertainty. Many pumps that are operated in many industrial applications day by day are therefore a major contributor for not reaching sustainability goals, especially when considering life cycle cost and emissions. Some manufacturers report cases, of newly developed pumps, being unable to satisfy the pumping requirements. This leads to leads to great financial losses both for the manufacturer as well as the operator.

The few empirical models that are already in use by some of the manufacturers are only validated for specific fluid properties. Slight deviations of the fluid properties or the specifications of the pumps from the validated limits, cannot be handled dependently by these black-box models. Thus, in order to find the optimal specifications for their pumps, the manufacturers are forced to resort to experimental analysis of their specific pumps for the specific applications. This approach allows little carry-over of the findings to future projects, while taking a lot of time and generating great costs. Many of the relevant process fluids do not allow such dedicated experimental work at the pump manufacturers facilities e.g., polymer melts.

In short, there is a great need for (i) precise and dependable models for the prediction of non-Newtonian leakage flow and (ii) a viable database for validating future models. In order to address this issue, the Chair of Fluid Systems proposes to approach the different geometrically complex sealing gaps that are characteristic for the rotating displacement pumps with a generic experiment that also includes the relative motion of the boundaries. Due to the modular structure of the proposed test rig, models for the different gap leakage flows can be derived, thus upgrading the modelling of the pumps for non-Newtonian flows.

2. LEAKAGE OF NON-NEWTONIAN FLUIDS

The variety of fluids, collectively known as non-Newtonian fluids is vast. All of these fluids diverge from the linear and purely viscous behaviour that defines Newtonian fluids [5]. For the research of gap flow in the context of rotating displacement pumps, the pseudoplastic fluids with viscoelastic behaviour and without time-dependencies are of most relevance, since they cover the most frequent applications with non-Newtonian fluids. This kind of behaviour is documented for example for polymer melts or solutions. In order to achieve consistent fluid properties, silicone oils (Polydimethylsiloxane PDMS) are chosen for the experimental work. Silicone oils can be described as polymer melts that are liquid at room temperature. The rheological behaviour of silicone oils is well documented. [6–8]

2.1. Research on leakage in rotating displacement pumps

A literature review shows multiple analytical, numerical and experimental works that cover leakage flows in rotating displacement pumps. These however cover only applications with Newtonian fluids [2, 9–11]. The few works that cover the gap flow of non-Newtonian fluids range from findings in valves and piping components [12] to findings in pumping applications of centrifugal [13] and displacement pumps [14]. Rather than focusing on the modelling of non-Newtonian fluids, these works however focus on deriving effective viscosities for the specific applications. This approach does not allow for a carry-over of the results to different applications, fluids, pump sizes or pump types. Such field studies are therefore too close to the application to allow for extensive, yet cost effective parameter studies which would allow, to derive leakage mechanisms and models. They can be considered as black-box models and do not resolve the isolated gap flows.

Rituraj and Vacca [15] model the non-Newtonian leakage flow within gear pumps in their numerical simulation. The authors differentiate between the leakage paths within these pumps. Even pump specific behaviour such as the deflection of the gears due to the journal bearings can be modelled. The authors show good agreement of the simulation and experimental findings. Viscoelastic properties of the fluids are considered; however, the model fluids and their parameters are subject to confidentiality. The results can therefore not be evaluated or reused.

Research activities at the Chair of Fluid Systems in Darmstadt in the context of developing a type-independent efficiency definition for positive displacement pumps have shown that for Newtonian fluids, the internal leakage for screw pumps is independent of the rotational speed of the pump [2]. Imamoglu [16] reports similar observations for gear pumps, both based on analytical considerations and numerical simulation efforts. These findings are however not expected to be applicable to non-Newtonian fluids, which is one of the research questions to be addressed with help of the experimental test rig that is covered by this paper.

2.2. Research on generic gap flows

Extensive research on generic gap flows of non-Newtonian fluids can be found. These allow for a more differentiated view of the gap flow types that might be of relevance for the characteristic gaps in displacement pumps. A specific gap flow can thereby be further divided into the zones of characteristic flows as shown in **Figure 2**. The entrance and exit of the gap show an extensional flow, while the flow through the gap length with constant gap height can be modelled as a laminar flow.

The most relevant properties for the non-Newtonian fluids in scope for the research project are the shear viscosity and the extensional viscosity. For the type of flow relevant for the research activities, the planar extensional flow is assumed to be representative [17]. For Newtonian fluids the shear viscosity and the extensional viscosity for planar extensional flow are coupled by the following Trouton ratio, for non-Newtonian flow this generally is not the case [18].

$$\mu_{\mathrm{E,planar}} \sim 4\,\mu \tag{1}$$

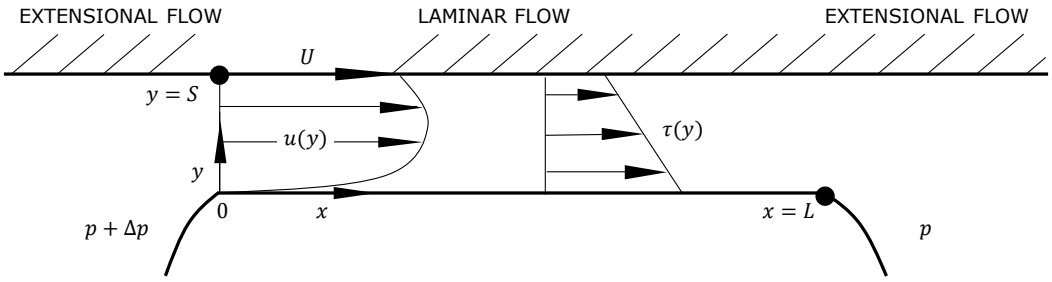


Figure 2: Characteristic gap for the leakage flow in displacement pumps

The shear and strain rates for the flow along sealing gaps are location dependent. In order to determine the actual viscosities, a fully resolved analysis of the non-Newtonian flows would be required. For the complex gap geometries and flow states in displacement pumps this is however not feasible. It is therefore not possible to estimate the local Reynolds number. The transition from laminar to turbulent flow is shifted to lower Reynolds numbers due to the shear-thinning fluid properties, which further increases the difficulties in predicting the leakage behaviour. [19] Therefore, modified or generalised Reynolds numbers are often used. These however are specific to the underlying fluid models. Rütten et al. [19] report the local Reynolds number to be multiple orders of magnitude higher compared to the global Reynolds number. Lacking the detailed knowledge for the superposition for all relevant effects, an apparent viscosity μ_{app} is often used. [18]

Apart from the extensional flow in the entrance and exit of the gap, the flow along the gap length L with constant gap height is also of interest. The laminar flow is expected to be represented by a Couette-Poiseuille flow due to the pressure difference across the sealing gap and the relative motion of the boundaries represented by the pump components surfaces. [9] The interaction of Couette and Poiseuille flow are well researched and documented both for Newtonian and non-Newtonian fluids. For non-Newtonian fluids the flow profiles may not simply be superimposed, since they are not independent from one another. [20] The resulting flow is known as the generalised Couette-Flow. The modelling of Couette and Poiseuille flow is based on the assumption of fully developed flow and may therefore not include all relevant effects, especially for viscoelastic fluids.

The viscoelastic fluid properties have been subject to research especially in the context of orifice flow. Rheological models often describe an orifice of capillary length L and radius R with an effective capillary length $L + n_{\mathrm{corr}}R$ for Newtonian fluids. The approach allows for the consideration of viscous resistance due to the velocity gradient at the entrance of the orifice. Bagley [21] shows that the correction coefficient n_{corr} is a function only of L/R even for non-Newtonian fluids with viscoelastic properties. The Bagley correction thus allows for the characterisation of flows entering a capillary, as may be the case at nozzles or measuring orifices. More than 50 years after the works of Bagley, square edged orifices are still the subject of research activities, which highlights the complexity and relevance of detailed models for small channel flows. [22, 23]

Apart from the works of Bagley [21], the works cited within this chapter do not consider fluids with elastic properties; apart from the simulative work of Rituraj and Vacca on gear pumps (see 2.1), the research centre is not aware of any work that investigates similar contraction flows in the context of moving boundaries.

In addition to the literature research, initial simulative work on the leakage of non-Newtonian fluids has already been carried out during the preliminary project at the Institute of Fluid Systems [24, 25]. Initial results confirm the assumption that the wall movement of non-Newtonian fluids must not be neglected. The test rig presented in this paper is therefore intended to offer data to be used in the validation of simulative efforts.

3. GAP FLOW TEST RIG DESIGN

Even though there has been extensive research on the fundamentals of gap flows of non-Newtonian fluids, the possibility of a transfer to the gap flows within the sealing gaps of rotating displacement pumps is not possible. Therefore, the design process of such pumps is far from straight forward and often requires very cost intensive iterative approaches. Evaluating the finalised pumps in their specific applications is rarely possible and places a big burden on the customers, yet the lessons learned are rarely, if at all transferable to future projects in different applications. With such approaches there is no practical way of isolating the different gap flows and performing a structured optimisation of characteristic geometric and operating parameters. In order to address these issues, the Chair of Fluid Systems at TU Darmstadt has set out to design a generic experiment which will allow the isolated consideration of the different characteristic sealing gaps. The experiment is to be situated in between the application specific considerations based on selected pump types for their specific applications and the research on standard gap flows, where the transfer to real engineering tasks is, as of now, not yet feasible. Based on the experimental results, a model is to be developed which allows the transfer to real engineering tasks. The model is expected to vastly improve the efficiency that is achieved in practical operation as well as the design process for non-Newtonian applications.

3.1. Hypotheses

For the experimental considerations of non-Newtonian gap leakage, a modular generic test rig has been designed. The test rig will aid the research on the following hypotheses:

1. When pumping non-Newtonian fluids, the internal leakage within the pumps can no longer be considered to be independent of the rotational speed. Therefore, the experiments on the generic gap flow should not show that the drag flow for the generic gaps with superimposed contraction flow is of negligible influence.
2. Upon entering into the gap and exiting out of the gap, due to the contraction flow of the non-Newtonian fluid, dissipative losses are present. By considering different characteristic gap shapes, the influence of the geometric parameters on the dissipative losses can be resolved.
3. Based on the results from hypotheses 1 and 2, guidelines for the design of sealing gaps and selection of pump operating parameters can be derived. These guidelines allow the manufacturers to improve their pump portfolio and allows them proactive design for new applications and markets with non-Newtonian fluids.

The results will be the basis for the development of an extensively validated, empirical model which allows the description of the leakage behaviour of non-Newtonian fluids across characteristic gaps within rotating displacement pumps. Furthermore, the experimental results are to be published in a reusable database in order to encourage further research.

3.2. Requirements

In order to be representative of the continuous pumping cycles that are most relevant for applications of the pumps, and in order to ensure quasi-steady flow, the test rig is required to offer continuous operation, batch operation is not desired. Thus, the measuring section needs to provide a continuously moving wall. The hydraulic test field needs to provide a continuous supply of conditioned fluid.
Further requirements for the test rig are a modular construction, ease and unambiguity of the assembly process as well as ease of adaptability for future, more detailed research projects. With these requirements the test rig is expected to provide reliable and repeatable test results even after major disassembly and reconstruction with minimum effort for validation and setup.
For the design of the test rig, the research parameters are specified. The range of zero shear viscosity extends from 0.1 Pas to 1 Pas, the pressure drop at the gap to be investigated will be up to 50 bar. Due to sealing components, the temperature is limited to 120°C.

3.3. Concept for the modular measuring section

In order to achieve a continuously moving wall with constant, yet adjustable surface speed, the measuring section is equipped with a cylindrical surface that is in motion relative to the gap module. In order to reduce the number of dynamic seals and maximise adaptability as well as ease of maintenance, a concept with stationary gap module and rotating convex cylindrical surface was chosen. Hereby care must be taken to keep the geometric similarity to the relevant leakage gaps within the pumps. The result of a conceptual design process is shown in **Figure 3** with an exemplary gap module highlighted in yellow.

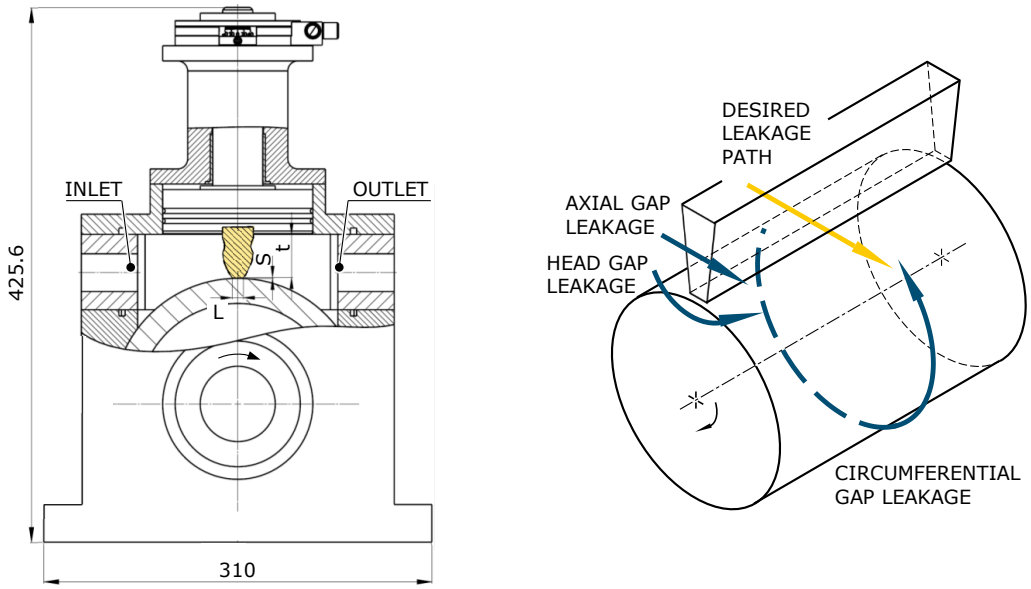


Figure 3: Modular measuring section of the gap leakage test rig (left) and schematic sketch of the secondary leakage paths within the measuring section (right)

The rotating shaft is hollow in order to reduce the rotating masses and thus decrease energy consumption as well as the forces exerted on the bearings of the test rig. By adjusting the rotational speed, the Couette-flow can be set. The pressure differential from inlet to outlet can be provided by a positive displacement pump, thus allowing for a Poiseuille-flow to be imposed.

The housing is made up of three parts and can be split axially. The bearings (not shown in **Figure 3**) are preloaded in order to minimise radial play. Thus, very tight tolerances between rotor and shaft can be achieved, while allowing secondary leakage flows to exit the test rig via drainage holes.

The sealing gaps within the displacement pumps that are the focus of the research activities often have a gap height of less than 10 µm. With the minimum achievable gap height of the secondary leakage paths in the same order of magnitude, the feasible gap heights to be assumed for the experiments are higher than the gap heights within the pumps. The gap module therefore needs to be scaled accordingly in order to be geometrically representative of the sealing gaps within the pumps, while the gap height for the secondary leakage paths is to be minimised, see **Figure 3**.

The head gap leakage in between the inlet and outlet side along the front side of the rotor is to be considered the most critical leakage path for secondary leakage, since it is a dynamic seal with comparatively short gap length. Therefore, tight tolerances are required.

For the analysis of different parameters, the experimental gap height needs to be adjustable. The mechanism for adjusting the height of the gap module is housed in the top plate. The gap module

which is affixed to a piston can be retracted from the rotating shaft using a spindle nut. The spindle nut is in turn affixed to a precision rotary table, thus allowing precise adjustment of the gap height. The piston seals the measurement section and is keyed in order to avoid rotation of the gap module. By variation of the dimensionless geometric and operating parameters, the measuring section allows to characterise the behaviour of the leakage flow as described in the sections before. The inlet and outlet ports are located in an intermediate plate between the three-part housing and the top plate.

In order to achieve visual accessibility for flow visualisation, the top plate with the adjusting mechanism as well as the intermediate plate can be substituted by an intermediate plate and a fixed-height top plate made from a clear plastic (e.g., PMMA). This allows for flow visualisation and even Particle Image Velocimetry (PIV) measurements to be performed [26].

For the characterisation of the operating conditions and fluid properties provided, measurement of the pressure and temperature is required. The sensors can be integrated into the intermediate plate. The temperature of the fluid is not only important to characterise the fluid properties but also for the controls of the fluid conditioning unit which is set to provide an adjustable constant fluid temperature for the experiments.

In order to minimise calibration efforts, the gap module is to include a distance sensor, which can be calibrated prior to being inserted into the measuring section. Thereby the actual gap height for each experiment can be logged. The direct measurement by laser micrometer e.g., is not feasible due to the very low gap height (0.1 … 0.6 mm). Capacitive distance sensors are sensitive to the fluid properties, thus complicating their application to the measuring task at hand. Therefore, inductive sensors e.g., eddy current sensors are favourable [27].

In order to be able to adjust the fluid parameters without impacting the general fluid properties, a suitable test fluid must be chosen. The choice of the test fluid also impacts the applicability of the results to future research, as can be seen by the works of Rituraj and Vacca [15] where the type of fluid cannot be disclosed due to confidentiality, thus drastically hindering future research. The fluid needs to have shear-thinning behaviour, show viscoelastic properties, while not presenting a health and safety risk or being subject to aging effects. A literature review shows a multitude of test fluids within the scope of shear thinning-fluids. The most prominent fluids are silicone oils [19, 6–8], aqueous Xanthan-solutions [12, 18] and Sodium Carboxymethyl Cellulose (CMC) solutions [17]. However, apart from silicone oils, none of the test fluids are able to satisfy the requirements mentioned above. Xanthan-solutions e.g., show changes in their properties in between batches and are subject to ageing effects. For CMC-solutions the literature review shows a shear thickening behaviour at low shear rate and thixotropic behaviour. Therefore, silicone oils have been chosen for the experimental tests. Silicone oils are also often quoted in the relevant technical literature. Silicone oils can be understood as viscoelastic polymer melts with a melting point below room temperature [28]. By a variation of the molecular weight, the zero-shear viscosity can be set. For low molecular mass and low shear rate there may be an area with Newtonian properties, however this is not expected to be of relevance for the experiments.

A characterisation of the leakage behaviour relies on the measurement of the leakage flow. Due to the low electric conductivity of silicone oils, magnetic-inductive flow meters are not applicable, therefore gear or spindle type flow meters can be applied.

The frictional torque as well as the rotational speed at the hollow shaft are measured by a torque measuring shaft in order to characterise the operating parameters.

The hydraulic test environment is required to provide an adjustable, constant volume flow of the non-Newtonian fluid with constant fluid properties. The concept depicted in **Figure 4** is able to provide the required volume flow for continuous operation of the test rig by pumping the fluid from a tank to the measuring section. In order to ensure consistent experiment parameters, a special focus is to be placed on the conditioning of the fluid. Compared to Newtonian fluids this is even more relevant for the non-Newtonian fluids, since, for example the temperature-induced change in viscosity also

impacts the shear stresses and the distribution of these, thus drastically changing the local flow characteristics. The conditioning of the fluids is realised by a heating and a cooling and filtering unit that are designed to control the desired fluid temperature in the tank.
The fluid is pumped by a positive displacement pump in order to achieve the desired pressure drop between inlet and outlet.

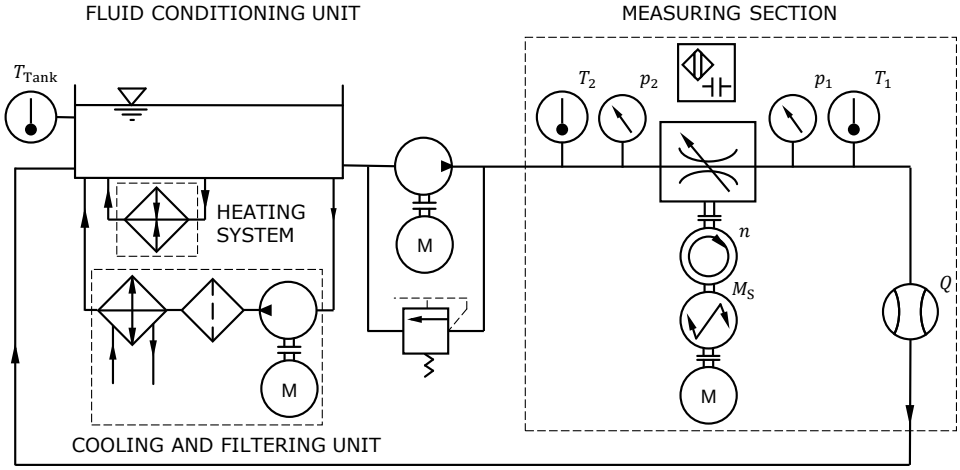


Figure 4: Schematic plan of the hydraulic test field for non-Newtonian fluids

In order to characterise the test fluids prior to the experiments, the parameters, such as shear storage modulus G' and shear loss modulus G'', can be determined by oscillatory rheometry. In accordance with the Cox-Merz-rule, the shear-rate dependent viscosity can then be derived [25]. Utilising the time-temperature-shift, the experimental effort can be reduced [29]. This objective characterisation of the fluid properties allows the reusability of the data obtained by the experiments and is of fundamental importance for the transferability of the findings.

4. CONCLUSION

As could be shown, the literature contains relevant blind spots when trying to develop a systematic model for the design of displacement pumps for non-Newtonian applications. The proposed setup of the test rig is expected to (i) provide important insight into the gap leakage, allowing for the (ii) development a model that will enable manufacturers to design their pumps according to the complex applications. Based on an extensive literature review, the test rig was designed to be a modular platform for future experimental works on the leakage of non-Newtonian fluids. The setup allows for extensive parameter studies to be performed, varying fluid parameters and operation parameters as well as the geometric parameters of the gap modules, in order to resolve the influence of the extensional flow.

ACKNOWLEDGEMENT

We would like to thank the Research Fund Pumps of the VDMA Pumps + Systems for the funding and continuous support of the research project.

NOMENCLATURE

A	area	m²
G'	shear storage modulus	N/m²
G''	shear loss modulus	N/m²
L	gap/capillary length	m
M_S	shaft torque	Nm
n	rotational speed	1/min
n_{corr}	correction coefficient	-
p	pressure	bar
Δp	pressure differential	bar
p_1	outlet pressure	bar
p_2	inlet pressure	bar
Q	volume flow	m³/s
R	radius	m
S	gap height	m
T_1	outlet temperature	°C
T_2	inlet temperature	°C
T_{Tank}	tank temperature	°C
u	velocity	m/s
U	speed of the boundary wall	m/s
μ	viscosity	kg m/s
μ_{app}	apparent viscosity	kg m/s
μ_0	zero-shear viscosity	kg m/s
$\mu_{E,planar}$	extensional viscosity for the planar case	kg m/s
ν	kinematic viscosity	mm²/s
τ	shear stress	N/m²

REFERENCES

[1] Karassik IJ (ed) (2001) Pump handbook, 3. ed. McGraw-Hill, New York

[2] Schänzle C (2021) Wirkungsgradmodell rotierender Verdrängerpumpen auf Grundlage von Ähnlichkeitsbeziehungen, 1. Auflage. Forschungsberichte zur Fluidsystemtechnik, vol 25. Shaker, Düren

[3] Schänzle C, Corneli T, Pelz PF (2018) A Similarity based Efficiency Model of Spindle Screw Pumps. IOP Conf Ser.: Mater Sci Eng 425:12020. https://doi.org/10.1088/1757-899X/425/1/012020

[4] Schänzle C, Störmer N, Pelz PF (2018) Modeling the Efficiency of External Gear Pumps Based on Similarity Considerations. In: BATH/ASME 2018 Symposium on Fluid Power and Motion Control. American Society of Mechanical Engineers

[5] Pahl M, Gleißle W, Laun H-M (op. 1991) Praktische Rheologie der Kunststoffe und Elastomere. Kunststofftechnik. VDI-Verlag, Düsseldorf

[6] Kataoka T, Ueda S (1967) Viscosity of polydimethylsiloxane blends. J Polym Sci A-1 Polym Chem 5:3071–3089. https://doi.org/10.1002/pol.1967.150051209

[7] Venczel M, Bognár G, Veress Á (2021) Temperature-Dependent Viscosity Model for Silicone Oil and Its Application in Viscous Dampers. Processes 9:331. https://doi.org/10.3390/pr9020331

[8] Rudolf M, Boutelier D, Rosenau M et al. (2016) Rheological benchmark of silicone oils used for analog modeling of short- and long-term lithospheric deformation. Tectonophysics 684:12–22. https://doi.org/10.1016/j.tecto.2015.11.028
[9] Sedri F, Riasi A (2019) Investigation of leakage within an external gear pump with new decompression slots: numerical and experimental study. Journal of the Brazilian Society of Mechanical Sciences and Engineering 41:224. https://doi.org/10.1007/s40430-019-1717-8
[10] Vacca A, Guidetti M (2011) Modelling and experimental validation of external spur gear machines for fluid power applications. Simulation Modelling Practice and Theory 19:2007–2031. https://doi.org/10.1016/j.simpat.2011.05.009
[11] Zhao X, Vacca A (2017) Numerical analysis of theoretical flow in external gear machines. Mechanism and Machine Theory 108:41–56. https://doi.org/10.1016/j.mechmachtheory.2016.10.010
[12] Polizelli MA, Menegalli FC, Telis V et al. (2003) Friction losses in valves and fittings for power-law fluids. Braz J Chem Eng 20:455–463. https://doi.org/10.1590/S0104-66322003000400012
[13] Kalombo JJN, Haldenwang R, Chhabra RP et al. (2014) Centrifugal Pump Derating for Non-Newtonian Slurries. Journal of Fluids Engineering 136. https://doi.org/10.1115/1.4025989
[14] Mrinal KR, Samad A (2017) Leakage flow correlation of a progressive cavity pump delivering shear thinning non-Newtonian fluids. IJOGCT 16:166. https://doi.org/10.1504/IJOGCT.2017.086299
[15] Rituraj F, Vacca A (2018) External gear pumps operating with non-Newtonian fluids: Modelling and experimental validation. Mechanical Systems and Signal Processing 106:284–302. https://doi.org/10.1016/j.ymssp.2017.12.042
[16] İmamoğlu D, Ertunç Ö (2018) Analysis of tip leakage in external gear pump
[17] Stock J (2009) Untersuchung der Hochviskosströmung in der Umgebung gleichsinnig rotierender Doppelschneckenelemente mittels dreidimensionaler Laser-Doppler-Velocimetry: Untersuchung der Hochviskosströmung in der Umgebung gleichsinnig rotierender Doppelschneckenelemente mittels dreidimensionaler Laser-Doppler-Velocimetry, Universitätsbibliothek der HSU/UniBwH
[18] Matthies M-O (2002) Einfluß der Knäuelstruktur auf das viskoelastische Verhalten von Polymerlösungen am Beispiel ausgewählter Hydrokolloide. Dissertation, Universität Hamburg
[19] Rütten M (ed) (2019) Verallgemeinerte newtonsche Fluide: Thermische und viskose Strömungseigenschaften. Springer Berlin Heidelberg, Berlin, Heidelberg
[20] Flumerfelt RW, Pierick MW, Cooper SL et al. (1969) Generalized Plane Couette Flow of a Non-Newtonian Fluid. Ind Eng Chem Fund 8:354–357. https://doi.org/10.1021/i160030a028
[21] Bagley EB (1957) End Corrections in the Capillary Flow of Polyethylene. Journal of Applied Physics 28:624–627. https://doi.org/10.1063/1.1722814
[22] Ntamba Ntamba B, Fester V (2012) Pressure Losses and Limiting Reynolds Numbers for Non-Newtonian Fluids in Short Square-Edged Orifice Plates. J Fluids Eng 134. https://doi.org/10.1115/1.4007156
[23] Chowdhury MR, Fester VG (2012) Modeling pressure losses for Newtonian and non-Newtonian laminar and turbulent flow in long square edged orifices. Chemical Engineering Research and Design 90:863–869. https://doi.org/10.1016/j.cherd.2011.11.001
[24] Pelz PF, Ludwig G, Kerres L et al. (2023) Verschleißverhalten bei partikelbeladenen Fluiden und Leckage bei nicht-Newton'schen Fluiden in rotierenden Verdrängerpumpen:

Abschlussbericht für das Forschungsvorhaben FKM-Nr.: 7051800, gefördert aus Eigenmitteln des Fachverbandes Pumpen und Systeme in VDMA, Darmstadt
[25] Mehrnia S, Kerres L, Kuhr M et al. (2022) Rheological modelling of viscoelastic fluid in a generic gap of screw pump. IOP Conf Ser.: Mater Sci Eng 1267:12013. https://doi.org/10.1088/1757-899X/1267/1/012013
[26] Antoniak P, Stryczek J (2018) Visualization study of the flow processes and phenomena in the external gear pump. Archives of Civil and Mechanical Engineering 18:1103–1115. https://doi.org/10.1016/j.acme.2018.03.001
[27] Micro-Epsilon Abstands-Sensoren. https://www.micro-epsilon.de/. Accessed 01.10.22
[28] Mincks LM (2002) Pressure drop characteristics of viscous fluid flow across orifices
[29] Ahmed J (2017) Time–Temperature Superposition Principle and its Application to Biopolymer and Food Rheology. In: Advances in Food Rheology and Its Applications. Elsevier, pp 209–241

SELF-SENSING MICROPUMP WITH GAS BUBBLE DETECTION FOR IMPROVED DOSING RELIABILITY

Kristjan Axelsson[1]*, Mohammad Sheikhsarraf[1], Phillip Höllein[2], Dirk Lewke[2], Martin Richter[1]

[1]*Fraunhofer EMFT, Hansastr. 27d, 80686 München*
[2] *RAPA Healthcare GmbH & Co. KG, Albert-Pausch-Ring 1, 95100 Selb*

* Corresponding author: Tel.: +49 89 54759-454; E-mail address: kristjan.axelsson@emft.fraunhofer.de

ABSTRACT

In micro-dosing system gas bubbles are a common source of dosing errors, by adding sensors to detect these the complexity and costs of the system increases. This paper introduces a novel approach for piezoelectric micropumps to detect gas bubbles without the usage of additional sensors like pressure or medium sensors. The proposed approach uses artificial intelligence (AI) methods directly implemented onto the micropump's high voltage driver circuit. The AI-methods are trained on a dedicated testbench, demonstrating successful gas bubble detection without increasing system costs and complexity and without the usage of additional sensors. This technology enhances dosing accuracy to ensure safer and more reliable micro-dosing applications.

Keywords: Piezoelectric, Micropump, Drug-Dosing, Self-Sensing, Artificial Intelligence, Gas-Bubble Detection

1. INTRODUCTION

In medical technology, biotechnology and the chemical industry, very small quantities of liquid often have to be precisely dosed. Micropumps can be used for this purpose [1]. To ensure high dosing accuracy and reliability, it is important to monitor possible faults and failures. In micro-dispensing technology, small gas bubbles are one of the main disturbances which are unavoidable in practical operation. This failure can lead to dosing errors or even risk the safety of patiens during medication dosing if the required amount of medication is not dosed correctly [2]. There are various options for monitoring the dosing and detecting malfunctions and anomalies, such as gas bubble detection. However, these require additional external sensors, which makes the dosing system more complex and expensive. This paper introduces a new technique for detecting gas bubbles in micropump applications through a novel approach called **Self-Sensing**.

1.1. Piezoelectric diaphragm micropumps

The RAPA micropump is made up of thin stainless steel foils that are structured by etching and bonded using laser welding. Two of these foils contain the passive check valves and are welded on top of a ground plate. A piezoelectric diaphragm is then mounted on top of the stack. The diaphragm consists of a stainless steel foil with a zirconate titanate piezoelectric ceramic (PZT) glued on top (Figure 1).

To achieve a high compression ratio, a patented process called pre-tension (Patentnr. US9410641B2, 2016) is used [3]. This method bends the piezoelectric diaphragm into an upward position, creating the pump chamber. The micropump is then integrated into a microfluidic system, providing both

electrical and fluidic contacts. To actuate the piezoelectric micropump a rectangular voltage is applied. The voltage level is determined by the thickness of the piezoelectric material. During each stroke, one passive check-valve opens while the other one closes. Due to continuous repetition, this results in a constant volume transport through the pump chamber from the inlet to the outlet valve.

The experiments in this publication were carried out at Fraunhofer EMFT using 6 piezoelectric micropumps manufactured by RAPA.

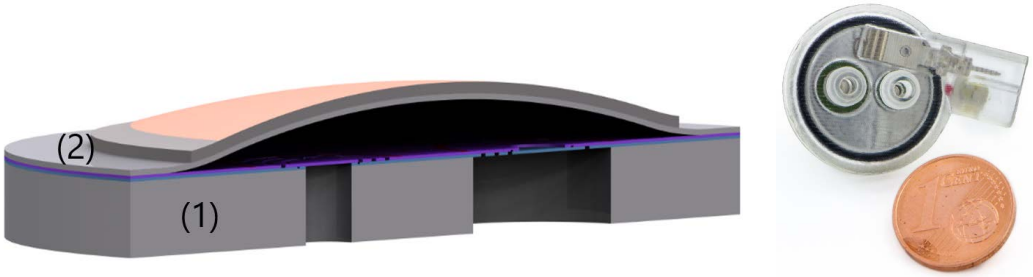


Figure 1: RAPA micropump made up of stacked steel foils with passive check valves (blue and purple layers) bonded by laser welding onto the base plate (1). The actuation is realized by a piezoelectric ceramic glued on top steel foil (2).

1.2. Self-Sensing theory

Piezoelectric actuators use the indirect piezoelectric effect to create mechanical deflection through an electrical control signal. In contrast, piezoelectric sensors use the direct piezoelectric effect to generate an electrical signal from mechanical deformation. The Fraunhofer EMFT intelligent pump utilizes both the indirect piezoelectric effect (actuator) and the direct piezoelectric effect (sensor) to achieve the desired self-sensing effect (Figure 2).

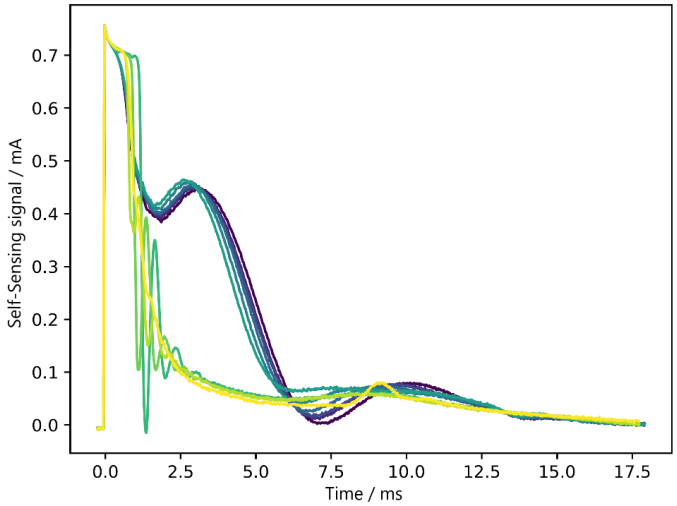


Figure 2: Self-Sensing signal measured with a current-to-voltage-converter amplifier circuit. Blue and dark green curves indicating a normal operation by pumping water. Yellow and light green signals indicating a gas bubble inside the pump chamber.

The Self-Sensing mechanism of piezoelectric micropumps relies on the indirect piezoelectric effect, which causes volumetric displacement in response to pressure p, voltage U, and the piezoelectric coupling factor C_E^*. The coupling factor is also dependent on the applied voltage U. Additionally, the fluidic capacity is denoted by C_p.

$$V(p,U) = C_p p + C_E^* U \tag{1}$$

The equation for the charge with electrical capacitance C_{el} follows the electrical analogy of fluidics.

$$Q(p,U) = C_{el} U + C_E^* p \tag{2}$$

Equation (2) represents the direct piezoelectric effect. To obtain the current I, we differentiate with respect to time, resulting in a four-term equation.

$$I(p,U) = C_{el}\frac{d\mathrm{U}}{dt} + \mathrm{U}\frac{dC_{el}}{dt} + \mathrm{p}\frac{dC_E^*}{dt} + C_E^*\frac{dp}{dt} \tag{3}$$

The charging current of a constant capacitance is described by the first summand. The second summand describes the current resulting from the change in capacitance due to the big signal piezoelectric effect. The third term also describes a large signal effect, namely the change in the d_{31} coefficient at high voltages. The fourth summand describes the signal of interest, which is caused by the direct piezoelectric effect resulting from pressure changes in the pump chamber. Modulation of the charging current occurs when a bubble passes through the pump chamber, caused by a change in viscosity of the medium. Fore- and backpressure changes also affect pressure equalisation in the chamber and therefore the modulation of the loading current by the Self-Sensing signal.

1.3. Idea of a self-sensing micropump

Due to the fluid-mechanical couplings (see section 1.2) of the system, the sensor current modulates the charging current in a variety of ways so that various system states create a characteristic "fingerprint" in the "Self-Sensing" signal. To obtain this additional information, the self-sensing signal is measured by a load-free amplifier circuit.

Due to the high variability of environmental changes and their influence on the self-sensing signal, classical solutions cannot reliably distinguish between them. Artificial Intelligence methods (AI-methods) are therefore used to classify them. For example, by looking at the mean value of the current, some fault conditions, such as bubbles in the pump chamber or changes in back pressure, look the same.

The AI-methods are directly integrated onto the driver circuit of the piezoelectric diaphragm pump. This minimizes the increase in the circuit's footprint and eliminates the need for additional pressure and air bubble sensors, resulting in reduced system costs.

For an application-orientated use of the self-sensing property, the AI-methods are trained with suitable measurement data. This training data is generated at a measuring station specially developed for this purpose at Fraunhofer EMFT (section 2). This allows to generate various error states, such as a change in viscosity, gas bubbles, changes in system pressures, blockages of the fluidic channels. These so generated error scenarios can be clearly assigned to the measured Self-Sensing signal. The AI-methods are trained with the data obtained in this way on a high-performance computer and optimized for runtime and memory usage on Arm Cortex-M4 microcontrollers. The original driver

circuit of the micropump is enhanced with these AI-methods and is therefore able to recognize complex system states and react accordingly. Thanks to RAPA's controlled production process, the variance of the pumps is so low that an AI-method generated for one pump can also be used for others of the same generation design.

Training AI methods and deploying them on microcontrollers has become easier due to the availability of open-source libraries such as TensorFlow, TensorFlowLite, and CubeAI. However, obtaining appropriate training data for a specific problem remains a challenge. To address this, a novel testbench for micropumps was designed and implemented.

2. SELF-SENSING TESTBENCH

To obtain high-quality training data for various environmental scenarios, a suitable testbench was designed and implemented (Figure 3). This testbench generates basic operating conditions and specific failure scenarios, and the corresponding Self-Sensing signal is sampled.

Multiple controllers are required to set the desired environmental scenarios, such as inserting gas bubbles and changing system pressures. The piezoelectric micropump is actuated using a high-voltage signal driver. Pressure controllers and active valves are used to set different pressure levels at the inlet and outlet of the fluidic system. To record these generated operating states and the corresponding self-sensing signal, various types of sensors are required, such as gas bubble detectors and pressure sensors. This information is collected and later used to label the training data.

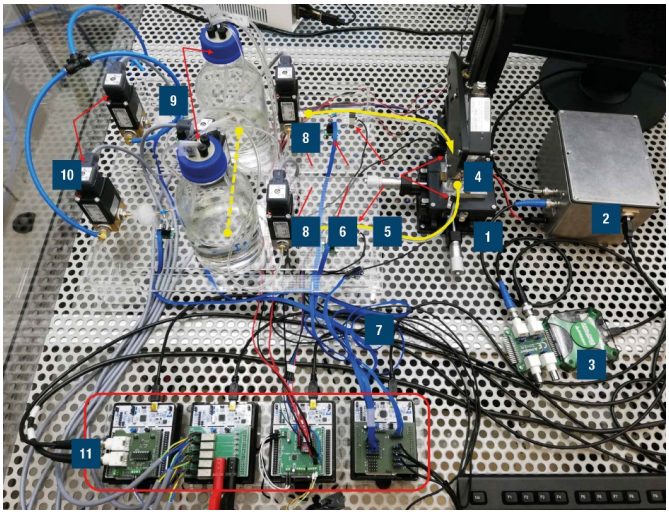


Figure 3: 1) Piezoelectric diaphragm pump 2) Self-Sensing circuit 3) Oscilloscope to measure Self-Sensing signal 4) Optical deflection sensor to measure diaphragm movement 5) Gas-bubble detectors 6) Pressure sensors 7) Temperature sensor 8) 3/2 valve to change medium from water to air 9) Inlet- and outlet-reservoir 10) 3/2 valve to connect pressure controller with inlet or outlet 11) Electronics to control sensors and actuators of the testbench. Below the testbench: Function generator, Piezo-Driver, Pressure controller.

A 3/2-valve is used at the inlet reservoir to switch the medium in the system between water and air. The valve allows a controlled insertion of gas bubbles of varying sizes into the fluidic periphery by regulating the switching time of the valve at the inlet tubing. Additionally, a 2/2 valve is installed between the input and output reservoirs to enable water to flow back from the outlet reservoir to the inlet reservoir (refer to the yellow dashed line in Figure 3). The next step involves generating various pressure scenarios by setting the inlet and outlet pressure using a pressure controller.

After the generation of the desired error states follows the detection of the corresponding environmental configurations and sampling the Self-Sensing signal. Four gas bubble sensors are used to detect bubbles: The gas bubble is detected by the first sensor as it enters the inlet tubing. The second sensor detects the bubble as it enters the pump inlet. The third sensor detects the bubble as it exits the pump chamber, and the fourth sensor detects the bubble as it leaves the system. To measure the applied pressure, four sensors are utilized: One for the inlet reservoir, one for the outlet reservoir, and one each for the pump inlet and outlet.

Finally, the self-sensing signal is sampled by the self-sensing electronic. All this data is later used for training and labelling with the obtained sensor data.

To train an AI-method to detect gas bubbles inside the piezoelectric micropump a suitable dataset for 6 devices was generated. To realize this the fluidic system is filled with water and gas bubbles are inserted. The measurement takes 5 minutes at a pump frequency of 25 Hz, resulting in 7500 pump strokes. The number of pump strokes with gas bubbles inside the pump chamber ranges from 1000 to 3000 cycles.

3. AI-METHODS FOR GAS BUBBLE DETECTION

Models are trained and evaluated on the PC with the usage of Cube-AI to estimate the Flash usage and runtime on the Arm Cortex-M4 microcontroller. In this work we compare different types of AI-methods such as Random Forest, Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN).

In binary classification problems with imbalanced classes caused by rare events, the F1-score metric is crucial. The F1 score (equation 4) is the harmonic mean of precision and recall, with a best value of 1 and a worse value of 0. F1-score is the classification metric in this work to evaluate the accuracy of the AI-methods [5].

$$F1 - score = \frac{2 * Presicion * Recall}{Precision + Recall} \tag{4}$$

The precision ratio represents the number of true positive results that were accurately predicted by the model.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{5}$$

Recall, also known as sensitivity, is a metric that measures the model's ability to correctly identify positive results.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{6}$$

As previously stated, AI training is conducted on a high-performance PC and then transferred to a Cortex-M4 microcontroller. However, a challenge with this approach is the varying implementations of AI model libraries. To address this, the Cube-AI from STMicroelectronics is utilized to directly compare AI methods using the same metric on a PC. The metrics used are memory usage in kilobytes and runtime in Multiply-Accumulate (MACC).

In this work gas bubble data was recorded for 6 micropumps and trained by procedure called cross evaluation. To obtain metrics for a single device, named as test-device, data for training and evaluation is generated from the remaining five devices. The F1-score is then calculated for each of the 6 micropumps and 50 different models (18 Random Forest, 16 Deep Neural Networks and 16 Convolutional Networks). This procedure is then repeated for each of the other 5 devices. The F1-score is shown in Figure 4. The memory usage in kilobytes and the runtime in MACC's for each of the 50 models is shown in Figure 5.

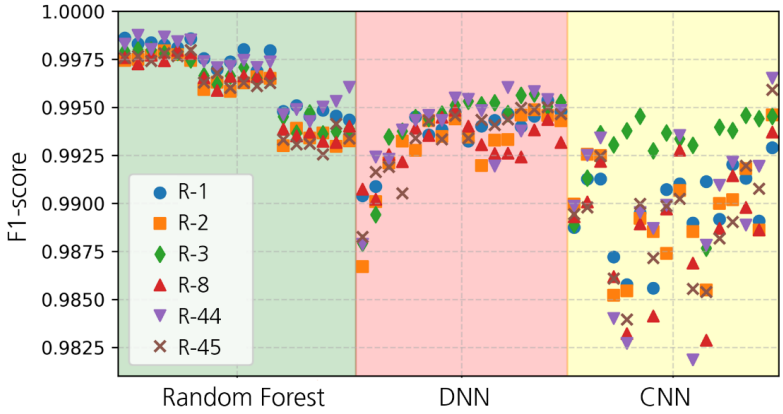


Figure 4: F1-score for Random Forest, Deep-Neural Networks and Convolutional Neural Networks. The Random Forest models shown the highest accuracy (F1-score).

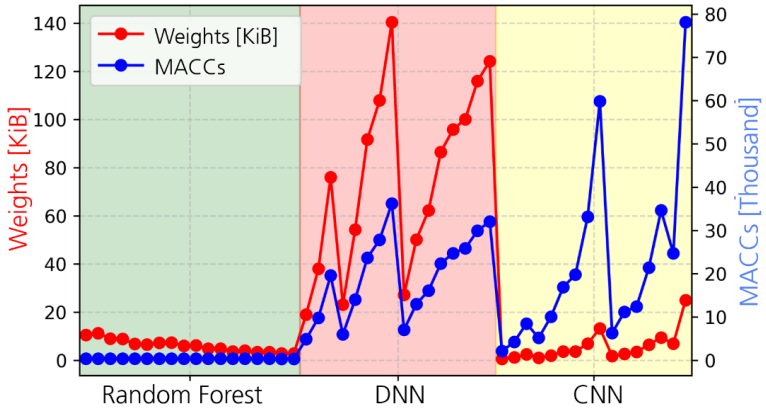


Figure 5: Flash consumption in kilobytes and runtime in MACC's for Random Forest, Deep-Neural Networks and Convolutional Neural Networks. The Random Forest models performs best by considering memory usage and execution time.

3.1. Results

This work demonstrates that AI methods can be used to detect gas bubbles within the pump chamber by analysing the self-sensing signal. Therefore, no additional sensors are needed.

If a gas bubble is detected in the pump chamber, the number of pump strokes can be counted. By multiplying the number of strokes with air by the stroke volume of the micropump, the size of the gas bubble can be estimated. This can be achieved with a small increase in the electronics footprint and a software update to the standard microcontroller for high voltage generation.

This new technology improves the micropump's functionality at a similar cost and size, representing a significant advancement in the safety and reliability of micro-dosing applications.

4. CONCLUSION AND OUTLOOK

This novel use of the sensory properties of the piezoelectric actuator enables the piezoelectric diaphragm pump to realize a self-monitoring feature for micro-dosing systems. This allows for observation of normal operation and detection of disturbances such as gas bubbles, without the need for additional sensors. This concept of gas bubble detection and quantification can improve dosing accuracy and precision in drug delivery applications by compensating for undosed volume caused by gas bubbles.

In addition to detecting and compensating for gas bubbles, a further stage of development is to detect and compensate for fluctuations in backpressure.

This concept enables real-time extraction of information from the piezoelectric actuator's regular actuation signal without the need for additional sensors. It demonstrates how the piezoelectric actuator interacts with its environment. As gas bubbles and backpressure fluctuations, including catheter occlusion, can interfere with drug dosing, this method has the potential to disrupt drug dosing with micropumps to a new level of efficiency, reliability and cost effectiveness.

REFERENCES

[1] A. Bußmann, H. Leistner, D. Zhou, M. Wackerle, Y. Congar, M. Richter and J. Hubbuch, "Piezoelectric Silicon Micropump for Drug Delivery Applications" Applied Science, 2021.

[2] A. Bußmann, C. Durasiewicz, S. Kibler and C. Wald, „Piezoelectric titanium based microfluidic pump and valves for implantable medical applications,“ Sensors and Actuators, 2021.

[3] M. Richter, M. Wackerle and M. Herz, „Method for manufacturing a bending transducer, a micro pump and a micro valve, micro pump and micro valve“. Patent US9410641B2, 2016.

[4] D. Horsch „Untersuchung des struktur-, und fluidmechanischen Verhaltens einer piezoelektrischen angetriebenen Mikromembranpumpe“, München: TUM, 2009.

[5] C. M. Bishop and N. M. Nasrabadi, Pattern recognition and machine learning. Springer, 2006, Vol. 4

Chapter 13

Industrial Control Strategy

SEAMLESS INTEGRATION OF DEVICE AND FIELD DATA INTO THE SYSTEM SIMULATION OF A HYDRAULIC SERVO-PRESS USING AAS

Raphael Alt[1]*, Malte Becker[2], Roland Bublitz[4], Sebastian Heppner[3], Heiko Baum[1], Tobias Kleinert[3], Katharina Schmitz[1]

[1] *Fluidon GmbH, Jülicher Str. 338A, 52070 Aachen, Germany*
[2] *Institute for Fluid Power Drives and Systems (ifas), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany*
[3] *Chair of Information and Automation Systems for Process and Material Technology (IAT), RWTH Aachen University, Turmstr. 46, 52072 Aachen, Germany*
[4] *Parker Hannifin Manufacturing Germany GmbH & Co. KG, Gutenbergstrasse 38, 41564 Kaarst, Germany*

* Corresponding author: E-mail address: raphael.alt@rwth-aachen.de

ABSTRACT

This contribution presents a framework for improving the simulation-based development process relying on seamless data and parameter exchange between general types of components, physical components of a specific system and the respective simulation model. The proposed solution relies on the concept of the Asset Administration Shell (AAS) to leverage the availability and interoperability of the heterogenous assets and to access their proprietary properties. Therefore, three AAS-based solutions are introduced to integrate different asset kinds. They represent component type data and simulation models provided by the component supplier, simulation models of components instantiated in a local simulation environment, and real components in operation with different communication interfaces. The solutions are implemented in a framework and demonstrated successfully through different simulation-based engineering use-cases using a servo-hydraulic press as a reference system.

Keywords: Industrie 4.0, Asset Administration Shell, Digital Twin, Simulation-based Development Process, OPC UA, NFC

1. INTRODUCTION

Model-Based Systems Engineering (MBSE) is playing an increasing role in the development of complex machines and products, with ever more demanding requirements. System simulations can add significant value at all phases of a machine's product life-cycle, from concept validation and component dimensioning, function and logic control development, virtual commissioning, process optimization, condition monitoring, and remote maintenance. However, their effectiveness as a basis for decision-making relies on a sufficiently accurate representation of the system state in the simulation, which depends on the level of detail of the simulation models and the congruence between actual system and component characteristics and model parameters assumed for simulation. Since simulation models, which are usually created during the engineering phase in the beginning of a product life-cycle, typically lack a tight coupling to the actual system and its respective engineering artifacts, there is a growing discrepancy along the life-cycle. The current process of aligning relevant data and parameters of the simulation models with engineering artifacts, technical documents, and the real system in operation is conducted manually. This process requires coordination across multiple

stakeholders, is labor-intensive and prone to errors, and is therefore rarely performed. The poor linkage is caused by the high heterogeneity, non-interoperability, and lack of tool-independent Application Programming Interfaces (API) to access relevant data and information to both, components in the field and in the simulation.

Therefore, the aim of this publication is to enable interoperable and cross-manufacturer automated bi-directional exchange of simulation-relevant parameters between machine components and their associated simulation models according to the principle of digital twins. This eliminates the need for manual parameterization steps, allowing individual development steps in the development process to be flexible, error-free, and time-efficient.

After showing the specific challenges and requirements that emerge from the simulation in this context in section 3, a solution that strongly relies on the Asset Administration Shell (AAS) is presented in section 4. In this approach, the AAS integrates three different asset kinds for facilitating the standardized exchange of asset information in an open and interoperable data synchronization layer - components or models – across a value creation network over multiple phases of the product life-cycle.

After introducing relevant features of the AAS, the solution approaches for the integration of three different asset kinds are presented in section 5. They consider:

- component type data and simulation models provided by the component supplier,
- simulation models of components instantiated in a local simulation environment,
- real components in operation with different communication interfaces.

The feasibility of the implemented framework with respect to the three asset kinds is demonstrated and validated in section 6 through the realization of engineering use-cases of the drivetrain of a hydraulic servo-press provided by Parker Hannifin.

2. DIGITAL TWINS IN MODEL BASED SYSTEMS ENGINEERING

Shorter product life-cycles and increasingly sophisticated product functionalities require the development of methodologies that effectively manage the complexity of multi-disciplinary engineering. At the same time, the flexibility and scalability of the established design methods shall be maintained. MBSE serves as an established approach in this context, for continuous description and analysis of the system to be developed, by utilizing simulation models, from the early phase of conception through the entire product life-cycle. [1]

Within MBSE, simulation models play an important role in analyzing, testing, and optimizing systems in a controlled virtual environment, even before the creation of physical prototypes. This approach not only reduces the costs and time associated with the development of physical models but also enhances the precision of system behavior predictions under varying conditions and life-cycle stages [2]. To serve as a reliable foundation for decision-making in design and optimization processes, the simulation model must accurately reflect the system behavior in different operating points. Conversely, any optimization performed on the simulation model should be reflected in the regarded system. Therefore, a bidirectional data flow for the continuous updating and synchronization of both the model and the current system state is crucial. For machine design and system optimization, simulation models are often represented by lumped parameter 0D or 1D system simulations. These system simulations incorporate the individual behavior of all components which are part of the machine. For these kinds of simulations, the required data flow between the physical machine and the simulation model normally includes the exchange of model parameters as well as input and output signals.

Simulation models can be utilized for different purposes over the entire product life-cycle of a machine. To ensure alignment between the simulation model and real system, occasional synchronization of parameters is essential, especially when changes have been made to either the machine or the simulation model. **Figure 1** illustrates the application scenarios for simulation models as well as the required points for parameter updating.

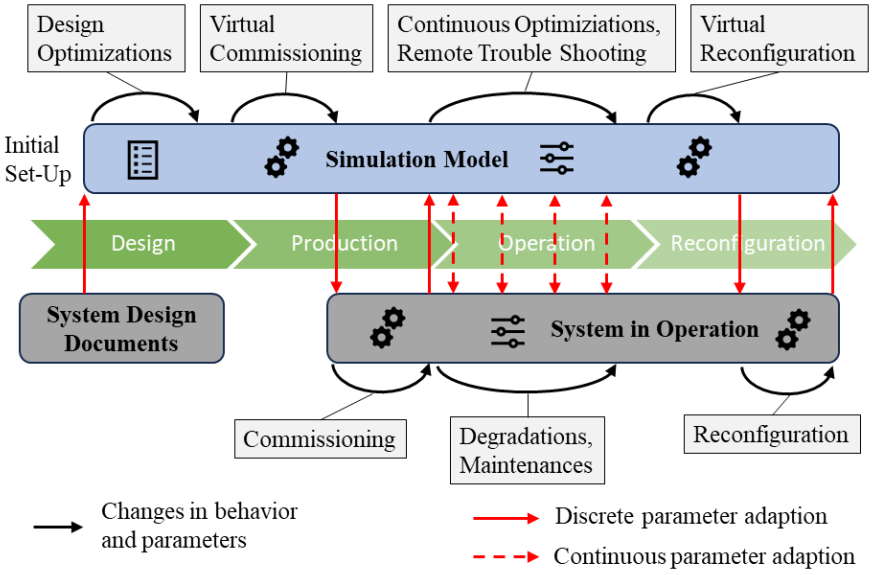


Figure 1: Changes of Machine and Simulation Model Twin across the product life-cycle

Initially, the system simulation model is set-up with a simulation environment, by linking simulation models of individual components. After general concept validation, during the dimensioning process, the simulation component parameters are adjusted to match the sizes and characteristics of the components available on the market. Afterwards, the simulation model is connected to a programmable logic controller, for control and controller design as well as virtual commissioning. The optimal parameters determined are subsequently transferred to the machine and further adjusted during the real commissioning process. During the operational phase, the simulation model is used continuously for monitoring and optimizations, e.g., to improve the energy efficiency. Concurrently, the physical system experiences continuous changes of its characteristics due to degradation, such as wear or contamination. Prior to reconfiguration of a manufacturing process, the intended changes are first prepared and tested using the simulation model, and only then implemented to the physical system, similar to the processes described during commissioning.

The frequent modifications in both, the machine and its model, throughout the product life-cycle underline the need for continuous mutual synchronization. Manual synchronization, as suggested in **figure 2** by the concept of Digital Model, represents the current state-of-the-art in the industry. Manual synchronization and adaption are labor-intensive and prone to errors, leading to outdated models that diminish in quality as a foundation for decision-making. Therefore, widespread adoption of simulation models in industry remains limited despite the many advantages they offer.

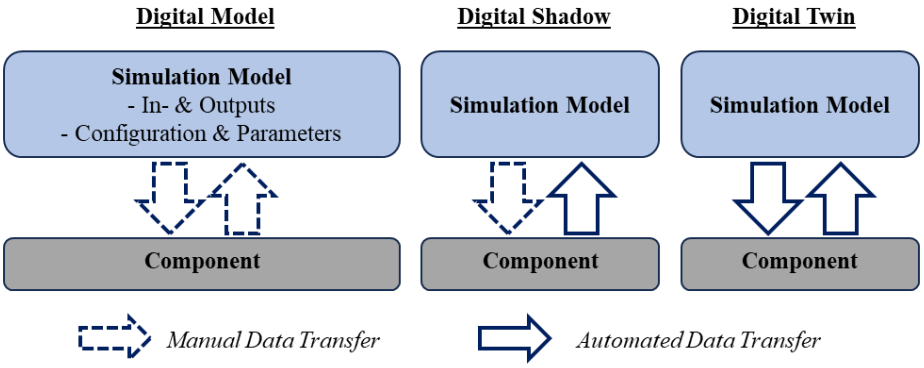


Figure 2: Digital Model, Shadow, Twin, after [3]

Given these challenges, there is a compelling need for a tightly coupled relationship between the simulation model and the system at design and in operation phases, facilitated by automated bidirectional data flow. This becomes even more important when the simulation model is integrated into real-time control loops with the corresponding real system. A digital object, which possesses this level of automated coupling with the real object, is referred to as a Digital Twin [3], see figure 2. Therefore, the implementation of Digital Twins serves to overcome these limitations to fully realize the potential of MBSE and enable efficient, effective, and economical application of digital engineering.

3. CHALLENGES AND REQUIREMENTS FOR USING SIMULATION-BASED DIGITAL TWINS IN ENGINEERING

The evolution from traditional engineering approaches to MBSE has introduced both opportunities and challenges in the engineering landscape. This transition necessitates a reevaluation of the requirements and challenges associated with engineering using Digital Twins.

MBSE offers a holistic approach to engineering, wherein all relevant requirements and aspects are described across various domains and are consistently interlinked. This allows for transparent visibility of the impact of design changes across all modeled systems and processes, enabling conflict identification and resolution. The approach also facilitates iterative design changes throughout the product life-cycle, incorporating feedback loops. Moreover, MBSE allows for the decomposition and composition of the overall system into subsystems, thereby enabling independent and parallel development as well as reusability.

This leads to the following implications and consequences during product development:

- The holistic view of MBSE requires the representation of systems from various perspectives and domains, leading to the involvement of multiple stakeholders and the utilization of different engineering tools.
- The approach considers system composition and decomposition requiring modular subsystems and components with defined system boundaries and interfaces that can separately be developed and then combined into an overall system.
- The involvement of multiple stakeholders and engineering tools results in heterogeneous and distributed data sources with proprietary data formats. For seamless integration of all partial solutions across various engineering client applications, the shared data must increase in availability, consistency, and interoperability with unambiguously defined semantics.

Given these implications, the prerequisites for digital engineering include high availability, continuity, and consistency of all relevant and interconnected data across organizational boundaries, life-cycle phases, and tools. Only on this foundation, development processes can holistically be analyzed, optimized, and ultimately automated.

From the described implications of MBSE, the requirements for simulation models following the Digital Twin approach can be derived as follows:

- Simulation models of components should accurately represent the behavior of the real components, requiring suitable modeling of the physics and accurate parameterization.
- To ensure reusability and scalability of simulation models, a modular and encapsulated design with clearly defined interfaces is required for integration. To obtain flexibility and the ability to adapt system variations, relevant parameters of the simulation model should be accessible from external.
- To facilitate integration into various applications and tools, simulation models should be interoperable with unambiguously defined properties.
- To ensure that multiple stakeholders and engineering applications obtain an equal and most recent system state, the simulation instance should be globally accessible and reflect a unique data source, adhering to the principle of single-source-of-truth.

To fully leverage the capabilities of Digital Twins, tools and applications are needed that enable seamless linking of Digital Twins with their real-world counterparts and integration into various engineering scenarios. This, in turn, imposes similar requirements concerning data availability, interoperability, unambiguity, and consistency for the component's parameters at design stage or of the physical component in operation.

4. SOLUTION APPROACH FOR SEAMLESS SIMULATION-BASED ENGINEERING

The proposed solution aims to support efficient and effective collaboration in engineering across multiple stakeholders within a value creation network. The approach addresses key aspects such as:

- openness and decentralization on both sides, data providers and consumers
- modularity and reusability through encapsulation and interoperability
- high data availability, continuity, and consistency.

The presented approach for seamless simulation-based engineering comprises two parts.

The first part of the solution focuses on creating tool-independent and modular simulation models that represent realistic machine behavior. This is achieved through encapsulated Functional Mock-up Unit (FMU) Co-Simulations. A FMU encapsulates a proprietary simulation model and provides a standardized interface that allows for the integration of simulations into various simulation environments [4]. Thereby the concept supports modularity and interoperability for heterogeneous components, which are provided by different suppliers and might also be created by different simulation tools [5].

The second part of the solution focuses on establishing a globally available data layer and providing interoperable access to relevant aspects of assets under consideration. For this, the standardized model of the Asset Administration Shell (AAS) is used. Improvements in asset data access are the foundation for seamless data exchange between components, their respective documentation and simulation models according to the principle of Digital Twin [6].

4.1 Functional Mock-up Unit Co-Simulation

The Functional Mock-up Interface (FMI) open standard maintained by the Modelica Association aims to simplify the usage and exchange of simulation models and is supported by more than 180

simulations tools [7]. The Functional Mock-up Unit (FMU) for Co-Simulation is an implementation of the FMI and enables the encapsulation of simulation models, including a suitable solver. FMU Co-Simulation provides widely applied technology suitable to enhance the versatile integration of simulation models in the MBSE context, also recommended by the prostep ivip SmartSE project [8]. It provides the following advantages in the presented solution approach:

- The FMI standard allows stakeholders to create simulation models using their preferred tools, while enabling the exchange and integration of these models into systems simulation. Provided the required licenses are available, FMUs remain tool-independent regarding their deployment. This enhances the portability and reusability of models and effectively reduces vendor dependency.
- FMUs enable the encapsulation of components, ranging from standard simulation models, such as hydraulic valves from standard component libraries, to individually developed models incorporating intelligent functions or logic control behavior, such as a motion controller. This enables precisely reflecting the component behavior designed by the component supplier and enables, as black-box entities, to share their specialized knowledge securely.
- FMU Co-Simulations can be integrated into various simulation environments and heterogenous system simulations with minimal configuration effort with respect to model and solver parameterization. This characteristic facilitates a plug-and-play methodology, simplifying the simulation process and boosting its effectiveness and user-friendliness.

Figure 3 shows relevant aspects for the presented approach of FMU simulation model integration and interaction. An FMU for Co-Simulation consists of one zip-file with the extension ".fmu" containing an XML-file with the definition of all variables and parameters that are exposed to the simulation environment as well as a set of C-functions and dynamic libraries provided in either source or binary form, used to initialize with defined parameters and run the simulation [9]. The integration, initialization, and orchestration during system simulation are carried out by the simulation tool. During the integration of each FMU, a Simulation Component is created as a tool-specific wrapper class to bind the FMU to the simulation tool. Upon instantiation, each Simulation Component can uniquely be addressed by a tool-specific identifier (ID). Furthermore, parameters described in the modelDescription.xml can be accessed by the simulation tool and, if provided, via a tool-specific API.

Notably, to introduce a clear distinction, a simulation model file is further defined as **Simulation Model**, whereas in contrast a **Simulation Component** defines the Simulation Model instantiated in a runtime environment. For more details on the coupled Co-Simulation in this context, see [5].

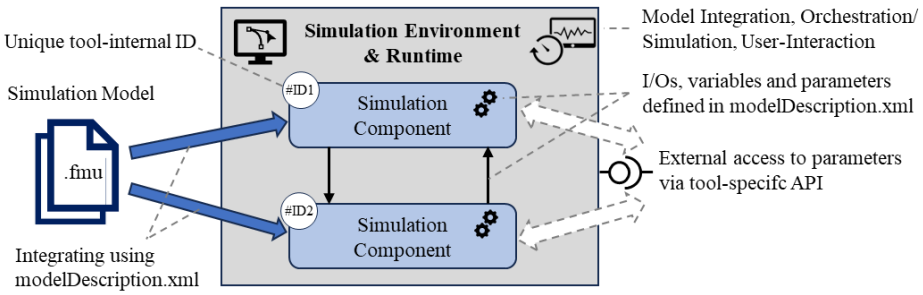


Figure 3: Instantiated Simulation Model and access to its parameters

4.2 Asset Administration Shell

To support MBSE across heterogeneous assets and engineering tools involving multiple stakeholders throughout the entire life-cycle, an open, decentralized and therefore integrative approach for exchanging of information is required. An Asset is any element that represents a perceived or actual value for an organization [6]. In the context of this publication, physical components are considered assets, e.g. a hydraulic cylinder, and its related engineering documents, data, and artifacts, such as simulation models. The presented architecture is based on the concept of the Asset Administration Shell (AAS). The AAS has a technology-neutral meta-model for implementing Digital Twins, as specified by the Industrial Digital Twin Association [6]. One relevant objective of the AAS focuses on enhancing interoperability and availability of properties and services of any asset within value creation networks. Aspects of the AAS relevant to the solution described in this publication are depicted in **figure 4** and briefly described below. The full specification can be obtained from [6, 10, 11].

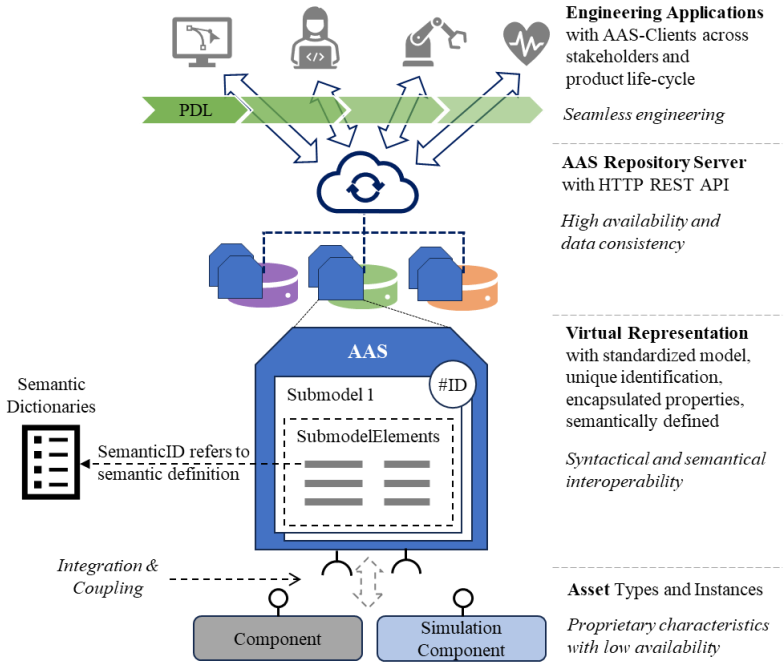


Figure 4: AAS and AAS Repositories for leveraging availability and interoperability of proprietary asset characteristics

An AAS is a virtual representation of an asset, while both are globally uniquely identifiable and referenced to each other. Assets can be of two kinds, Type and Instance. The Asset Type is a generalized classification that encapsulates common attributes and data structures of many Asset Instances, such as general component documentation or simulation model files. The Asset Instance, on the other hand, refers to a unique or individually existing element, such as a manufactured component or an instantiated simulation model. Typically, the AAS of an Asset Instance (AAS Instance) is derived from the AAS of an Asset Type (AAS Type) at a particular point in the product life-cycle to reflect the information relevant to the individual Asset under consideration. To illustrate this difference, an Asset Type could describe a component as found in a catalogue, whereas an Asset Instance would be a concrete component of that type to be shipped to a customer. [6, 12]

In an AAS, the characteristics of an Asset are modeled as SubmodelElements, encapsulating, for example, parameters or files. SubmodelElements are thematically grouped into SubmodelElementCollections and Submodels. Submodels can subsequently be standardized as templates to accommodate a set of relevant properties required for specific use-cases, such as the "Provision of Simulation Models" [13]. Both Submodels and SubmodelElements can be annotated with SemanticIDs. These unique identifiers that refer to semantic descriptions in semantic dictionaries, like eClass [14] or IEC CDD [15], describing the semantics of the content of the annotated element. This allows, in theory, the provision of the Asset's proprietary characteristics not only syntactically but also semantically interoperable.

To further enhance the availability across the entire value chain, including the product life-cycle and value networks, AASs are made available in AAS Repositories hosted on web servers. Using the specified HTTP REST API [10], and assuming the necessary access rights are granted, AASs can be accessed globally via the Internet by multiple AAS clients as part of engineering applications. This concept realizes an open and integrative approach that allows decentralized creation, provisioning, querying, and modification of AASs, to have further access to the respective asset's properties. For more details on implementing the AAS Data Layer through the AAS Server-Client Architecture, readers are referred to [12].

5. AAS-BASED ARCHITECTURE FOR SEAMLESS INTEGRATION OF SIMULATION MODELS IN MBSE

The seamless integration of simulation models into the engineering process and their close coupling with real-world assets according to the Digital Twin approach requires the implementation of three critical interfaces. These interfaces, shown in **figure 5**, are realized through the concept of the AAS and described in the following:

1. Interface for the selecting component types and retrieving of component data and engineering files, such as simulation models provided by the component suppliers.
2. Interface for bi-directional interaction with instantiated simulation models for a specific system.
3. Interface for bi-directional interaction with engineering artifacts and data created for a specific system and with actual components during operation.

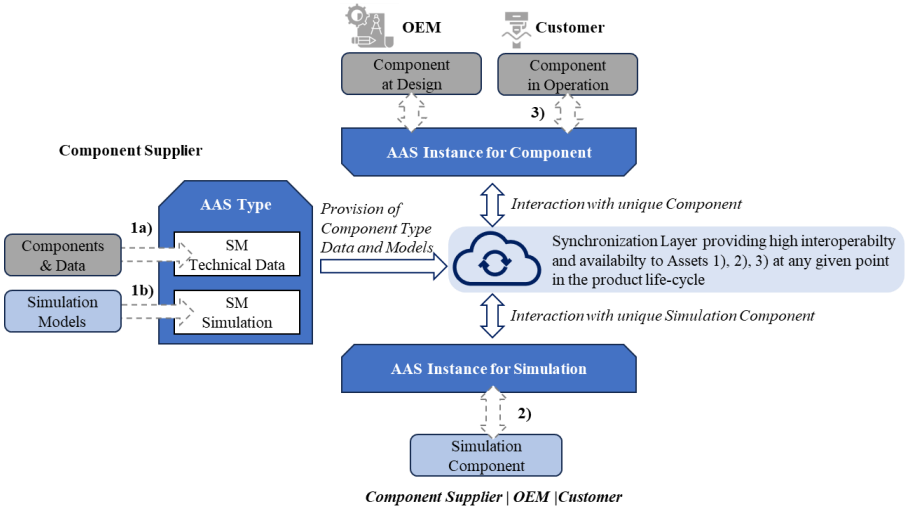


Figure 5: AAS-based Digital Twin synchronization interfaces for MBSE

1) AAS Type for providing component data and simulation model files

Component manufacturers hold valuable data and engineering files for their components, which can significantly support Original Equipment Manufacturers (OEM) in their overall system engineering processes. To facilitate interoperable data sharing for the targeted engineering discipline, the component manufacturer provides its proprietary component data, originating from its database, via AAS Types encapsulated in semantically annotated properties. The manufacturer is responsible for the initial linking of its proprietary data and the properties within the AASs. It is reasonable that due to a significant quantity and diverse range of component configurations, not all AAS Types are pre-generated and stored. Instead, they are generated on-access when required and subsequently made available via the defined AAS REST API from the server.

In the context of this paper, two submodels are particularly relevant and described briefly:

a) The Submodel Generic Frame for Technical Data for Industrial Equipment in Manufacturing [16] (SM Technical Data) specifies both the class of the component and its relevant technical parameters. This information can be used for searching components based on defined specifications, automated sizing processes, or for parameterizing simulation models.

b) The Submodel Provision of Simulation Models [17] (SM Simulation) provides essential information for selecting the provided Simulation Models appropriate to the engineering use-cases considered. Therefore, attributes for defining the considered engineering domain and the simulation purpose, such as sizing, virtual commissioning, or condition monitoring, are clearly defined. Moreover, the Submodel specifies properties for integrating the simulation model file into a simulation environment, such as the file format, the operating system requirements, and recommended solver settings. Within the scope of this publication, encapsulated simulation models are provided as tool-independent FMU Co-Simulation for lumped parameter system simulation (refer to chapter 4). More details on the implementation can be found in [12].

2) AAS Instance for providing access to the Simulation Component

To realize interoperable accessibility and high availability to simulation model files instantiated in local simulation tool environments – the Simulation Components – an AAS Instance is created for each of them to encapsulate relevant parameters. The AAS Instances are then made accessible in an AAS Repository via AAS HTTP REST API.

Figure 6 depicts the proposed modeling concept to provide relevant information for allowing simulation tools to link the Simulation Component and encapsulate its parameters in the AAS. It currently represents a non-standardized extension of the existing SM Simulation specification.

A link between the AAS Instance and the Simulation Component is established by referring to its unique ID in the Reference *LinkedSimulationComponent*. The ID is generated and linked during the instantiation of the Simulation Model. The *Parameter* SECs provide information to the Simulation Component parameters, intended to access from external clients. The Property ID defines the specific parameter identifier. In case of FMUs, the parameter identifiers are defined in the modelDescription.xml. The Reference SemanticReference, refers to a semantic definition via *SemanticID* to ensure an unambiguous definition of the parameter, aiding in identifying and matching corresponding properties in other submodels. Finally, the Property *Value* reflects the current parameter value of the Simulation Component.

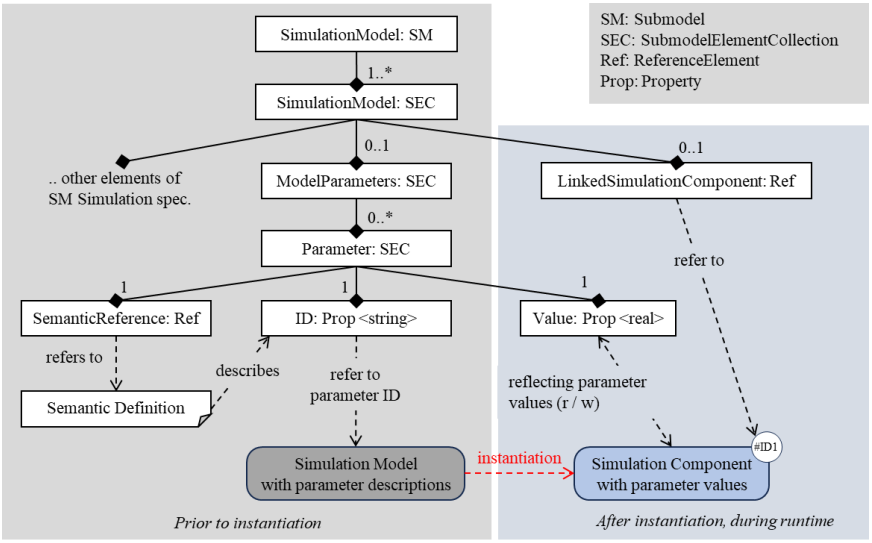


Figure 6: Modelling the coupling between AAS Instance and Simulation Component

Both the ID and the corresponding *SemanticReference* of the Simulation Model parameters are initially modeled in the SM Simulation of the AAS Type by the provider of the Simulation Model. The simulation tool realizes the instantiation of the Simulation Model and provides the corresponding AAS Instance with read and write access to the Simulation Component parameters via the AAS. In this way, proprietary Simulation Component parameters become globally and interoperable accessible with defined semantics for numerous external engineering client applications.

Although the realization described in this publication is implemented for a specific simulation tool using FMU simulation models, the concept is not limited to those and can similarly be applied to others.

3) AAS Instance for providing access to the component data

This interface provides interoperable access to data and parameters of real machine components and systems starting from the initial commissioning, during operation, including reconfigurations. Analogous to the AAS Interface of the Simulation Component, a dedicated AAS Instance is linked to its respective component in order to encapsulate component data and configuration parameters. If no AAS Instance already exists from the previous engineering phase that can be linked to the component instance, a new AAS Instance is created from the AAS Type and made available in an AAS Repository.

As machines consist of many heterogeneous components, different interfaces are used to communicate with the components themselves. **Figure 7** illustrates the coupling between the AAS and components in operation based on different communication interfaces. The controller of the drive system considered in this publication uses Ethernet-based OPC UA communication, which is widely used in the manufacturing industry. All relevant parameters of the controller can be accessed through the OPC UA interface of the device. To link the OPC UA device variables to the AAS, an OPC UA-AAS-gateway is employed. The mapping between proprietary device parameters of the OPC UA namespace and the properties in the AAS Submodels is provided by the component manufacturer. Provided that OPC UA access rights are available for operation, the controller's OPC UA and AAS Instance endpoints are the only specifications required for the linkage configuration. Should parameters require cyclic transfer, the desired timing must be specified.

Other bus systems, e.g., EtherCAT, EtherNet/IP, ProfiNet,and IO-Link can be connected similarly via specific gateways.

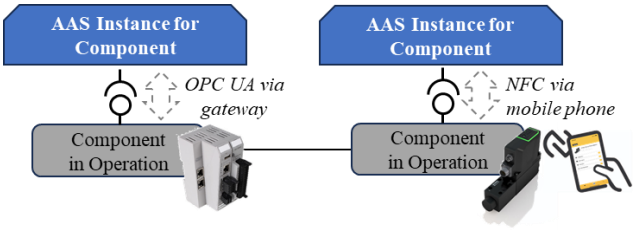


Figure 7: AAS parameter interfaces of the components in operation,
left: continuous transfer via OPA UA and gateway,
right: intermitted transfer via NFC and mobile app

However, not every parameterizable component is permanently accessible via a digital communication interface. Often, only a temporary wired connection to a PC is used for parameterization, for example, during commissioning. Therefore, wireless alternatives such as Bluetooth and NFC are also being used more frequently to enable and facilitate data transfer at certain points.

The valve considered in the use-cases of this publication provides an NFC interface. It can be used via a mobile device to extract all user parameters and additional operating and diagnostic data. The data obtained is accessible via an app on the mobile device and then synchronized to the AAS Instance in the AAS Repository. This temporary connection via mobile phone and NFC bypasses access to machine-critical real-time communication.

In the presented application, only data for device identification (asset ID and article number), the device parameters and condition monitoring information for future analysis of the system are exchanged. Generally, this approach enables not only the adjustment of control parameters, but also the full configuration of the controller, encompassing the connected valves and sensors. Direct manual entry on the machine is henceforth unnecessary.

Like the AAS Instance of the Simulation Component, proprietary and only locally available parameters of the physical component and its configuration become globally and interoperable accessible with defined semantics for numerous external engineering applications.

6. PROOF-OF-CONCEPT AND USE-CASES VALIDATION

The presented concepts are implemented as partial solutions and combined in a framework using the Eclipse BaSyx Python SDK [18]. In the following, common scenarios from different phases of the engineering process are demonstrated using a servo-hydraulic press. The framework's functionality is briefly explained, and benefits are illustrated in different scenarios.

The servo-hydraulic press consists of a drive-controlled pump, four parallel positioning axes, a drawing cushion axis, and a press cylinder. In addition, the press is controlled by a PAC 120/PACHC motion controller [19, 20]. For the proof-of-concept, only one positioning axis of the servo-hydraulic press is considered and set up in the simulation environment DSHplus [21]. The system simulation model is built as FMU Co-Simulation, in which relevant system components are modeled as individual FMUs and then connected at their hydraulic volume node signal interfaces.

6.1 Sizing and function validation

During sizing and function validation, the selection of the components and the initial parameterization in the system simulation is usually done manually by the engineer searching suitable components and their respective parameters from different data sources, e.g., websites and data sheets. This process is error-prone and requires a significant amount of time. To support the selection of suitable components, the framework allows searching among AAS Types in AAS Repositories offered by component manufacturers that match the desired requirements. From the selected AASs, the framework supports the initial creation of a Simulation Component by instantiating a provided Simulation Model as well as transferring parameters to an existing Simulation Component. This leads to a faster and error-free set up of a system simulation model, that closely matches the real properties of the system components. If necessary, the user can carry out simulative investigations and repeat the process with other component parameters. More details on the implementation are provided in [5, 12]. In this scenario, the interfaces 1. and 2. described in chapter 5 are used.

Figure 8 shows the general features of the framework for the component selection and model parameterization. In the regarded use-case, the engineer is looking for a valve that enables the hydraulic press to meet the requirements and full functionality. The system simulation structure (figure 8-1) is derived from the hydraulic scheme. However, in the beginning, the components in the simulation do not include a Simulation Model (figure 8-2a). From basic calculations, the valve dimensions are estimated by the engineer to be in a range between 10 to 40 bar nominal pressure difference and with a nominal flow rate between 50 and 100 l/min. In order to obtain a selection of suitable components, the framework offers the possibility to query an AAS Repository of the component manufacturer using search criteria defined by the user. Search criteria can be added in the query definition section (figure 8-3) via the specification of the semanticID of the regarded property with respective value constraints (min, max, equal). For better usability, separated drop-down lists and checkboxes are provided, such as for defining a specific component type or for targeting AASs that also provide a Simulation Model. However, in both cases, the query criteria are derived similarly and added to the others. Notably, as described in chapter 5, the component type is specified in the SM Technical Data of the AASs and defined via eClass IRDI. The availability of an FMU simulation model can similarly be determined from the SM Simulation.

The AAS Repository section (figure 8-4) lists AASs that comply with the combined query criteria (figure8-4a), and the engineer can select from these for further model initialization via “Initialize Simulation Component”. As described in chapter 5-2, if the selected AAS provides a specific FMU valve Simulation Model, it is integrated and instantiated by the tool to become a Simulation Component. Otherwise, an AAS with a default valve Simulation Model from the simulation tool provider AAS Repository is considered in the same way to obtain a Simulation Component (figure 8-2b). During the integration process, the respective AAS Instance is generated and coupled with the Simulation Component via their unique IDs (figure 8-4b). This AAS Instance serves as the representation of the Simulation Component in the synchronization layer for providing interoperable access to external applications. Based on the semantic annotation of the simulation parameters in the AAS Instance, the Simulation Component can further be re-parameterized via “Sync Parameters from AAS” by selecting other AASs with or without SM Simulation. This process can arbitrarily be repeated until the desired system behavior is archived. In the presented example, the AAS with a default valve Simulation Model was initialized with default parameters of 45 l/min @ 35 bars, then re-parameterized with the AAS of a valve with 50 l/min @ 35 bars. Since the semanticID of the parameter property includes value unit definition, automated unit conversion can also be included during the parameter transfer in both directions. Here, l/min and bar were converted to m^3/s and Pa, as required by the Simulation Component in the simulation tool.

This approach allows both to integrate specific and encapsulated component specific Simulation Models, as well as for considering AASs in the simulation that only contain technical parameters without any Simulation Model. This leaves great flexibility and lowers the entrance barrier of component manufacturers that can only provide very basic AASs of their components. In any case, during the creation of the Simulation Component, the corresponding AAS Instance is generated and linked, providing an interoperable interface for bi-directional parameter transfer.

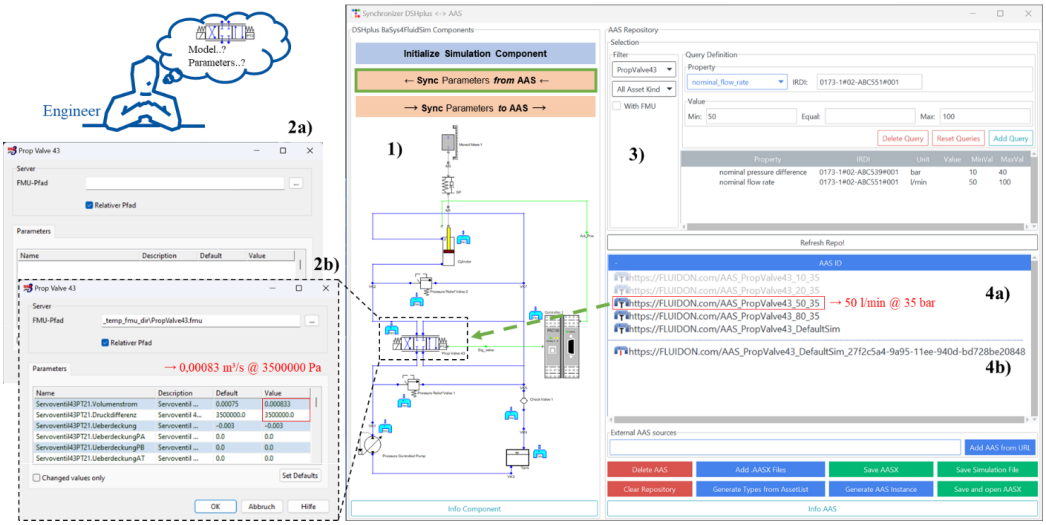


Figure 8: Implemented framework during selection and parameterization

6.2 Commissioning and operation phases

In the virtual commissioning scenario, the commissioning steps are carried out virtually in the simulation environment before the actual assembly of the physical machine. This enables potential errors to be detected and eliminated at an early stage, as well as improves and speeds up the commissioning process of the real machine. Typical steps include, for example, determining control parameters, estimating capacitive lines between components, and analyzing the interaction of subsystems with and without the control hardware. These tests and optimizations can be performed using simulation, for example, via Software-in-the-Loop and Hardware-in-the-Loop. Therefore, it is crucial that the system simulation behaves closely to the real system, for which a solution was presented in the previous chapter. When the machine is set up and put into operation during commissioning, all the parameters that were optimized in the simulation should be transferred to the components of the physical machine. The interfaces 2. and 3. described in chapter 5 are used in the following scenarios.

In the example shown in **figure 9**, the user optimized the parameters in the Simulation Components of the controller and the valve during virtual commissioning (figure 9-1). After the user has selected the Simulation Component and the targeted AAS Instance of the real PAC120 controller by pressing “Sync Parameters to AAS”, the parameters are transferred automatically in three combined steps. First, simulation parameters are updated in the SM Simulation of the AAS Instance of the respective Simulation Component (AAS Instance is not shown in the figure). The parameter values are then transferred, using the SemanticReferences of the Parameters, to the semantically equivalent properties found in various Submodels of the AAS Instance of the PAC120 controller (figure 9-2), which the Parker AAS Repository provides. From there, the parameters are updated in the PAC120 controller via the OPC UA-AAS-gateway (figure 9-3). If required, the operator can now access and further

optimize the parameters via HMI. Similarly, the valve specific settings and parameters are transferred from the simulation to the real component, with the exception that instead of OPC UA, the mobile phone application with NFC interface is used by the operator.

As commissioning is often iterative, parameters that were changed during commissioning should similarly be transferred back from the real component into the simulation environment. According to the principle of Digital Twin, the mechanism presented for transferring component parameters via AAS interfaces is bi-directional, allowing both the component and the simulation model values to be updated automatically when changes are made to either. However, when new parameter values are updated to the physical components, it must be ensured that the machine is in a safe state. Therefore, exchange is initiated manually by the operator in the HMI.

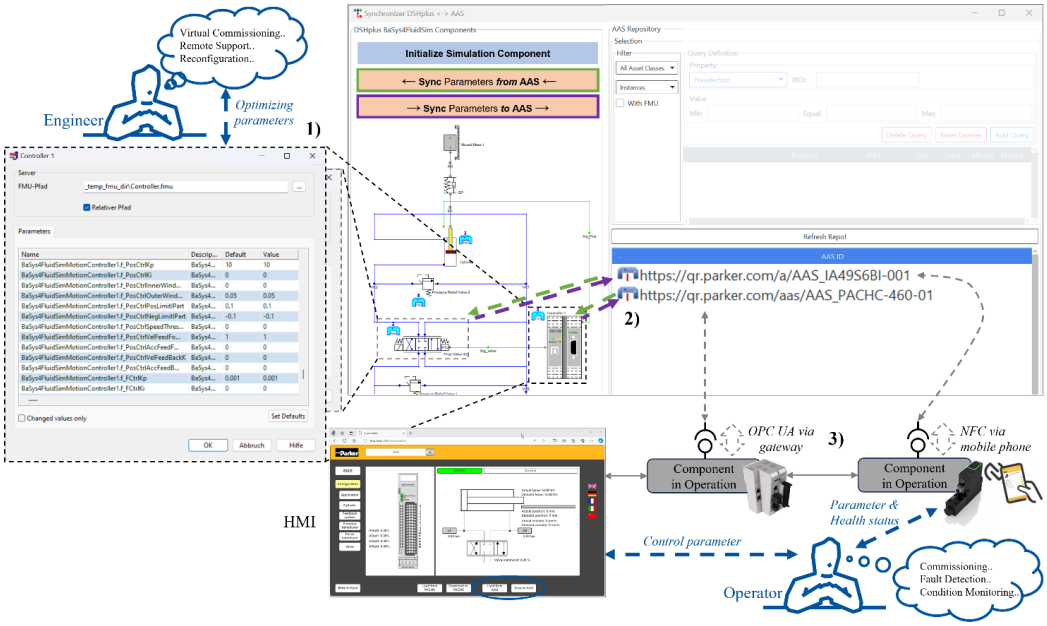


Figure 9: Bi-directional parameter transfer to the field devices for operational phases

In order to always have a profound simulative decision-making basis, and efficiently propagate found optimizations back to the machine, the presented mechanism for bi-directional parameter synchronization between simulation and field is not only important during commissioning, but also crucial to other use-cases among the entire operational phase.

In the case of unexpected irregularities or machine degradation during operation, solutions can be found more effectively and efficiently enabled by the transparency of simulative investigations. For example, when friction in the cylinder of the servo-press changes, the controller parameters need to be re-parameterized to compensate for the changes. Similarly, this applies to scheduled changes and feasibility tests, for example, to answer whether the servo-press used in operation can perform a modified pressing task. In all cases, the solution also improves the possibility of simulation-based remote support by another engineer or even the OEM, provided access rights are granted.

7. CONCLUSION AND OUTLOOK

This publication presents a framework that enables seamless simulation-based engineering across the entire life-cycle and with different stakeholders. Therefore, relevant requirements emerge in Model-Based System Engineering to better handle complex, modular, and multi-domain product development. The identified requirements aim to improve the availability, consistency, and interoperability of the assets involved, aiming to efficiently and seamlessly link all the steps and tools involved in engineering.

For this purpose, an open, decentralized approach based on an interoperable layer for data exchange is established in the framework by considering three relevant interfaces. The interfaces are realized based on the Asset Administration Shell (AAS) concept. The AAS accepts heterogeneous assets and provides interoperable access and better availability to component type data and simulation models of the suppliers, to simulation models instantiated in a simulation environment, and to the components operated in the field with various communication interfaces. The resulting data layer is the basis for enabling simulation-based engineering according to the Digital Twin concept, with automated bi-directional data exchange for the synchronization of system states between simulation models and real components, which is required in numerous use-cases with different engineering client applications across engineering and operation phases.

The concept was successfully implemented in a framework in Python. Thereby, different kinds of components were integrated through the AAS. Amongst these are the component database from Parker, the instantiated simulation models in a simulation environment, and two field devices from Parker using OPC UA and NFC. The basic feasibility was demonstrated by realizing different use-case scenarios on a servo-hydraulic reference system from Parker and a system simulation in DHSplus. The framework supports use-cases such as the initial dimensioning and function validation of the components in the machine, the commissioning process and further optimization measures during operation. It is concluded that the proposed and implemented AAS-based solution simplifies and optimizes the simulation model-based product development for OEMs, enables more reliable simulation-based decision basis in operation for the operator, and offers new business models for component manufacturers by providing easily accessible information.

For widespread use in practice, it must be clarified to what extent and which parameters may be written into an operational machine from the outside. Therefore, role concepts must be developed to solve these safety-relevant issues.

Besides, the concept presented is currently based on an existing simulation structure. In the future, it should be possible to derive and generate the hydraulic simulation structure automatically from the existing machine topology. Furthermore, the approach presented here mainly considers static parameters, whereas access to dynamically changing I/O signals via the AAS would also be relevant for some use cases. Based on this, stand-alone Simulation Components with an independent runtime environment and communication interface can also be realized for more flexibility and tool independence.

8. ACKNOWLEDGMENTS

The content of this paper is created during the research project BaSys4FluidSim under the reference 01IS21071A. The project is funded by the Federal Ministry of Education and Research (BMBF) and coordinated by the German Aerospace Center (DLR). The responsibility for the content of this paper lies with the authors. The authors would like to thank all project participants for their support.

REFERENCES

[1] Gausemeier, J., Dumitrescu, R., Echterfeld, J., Pfänder, T., Steffen, D., Thielemann, F „Innovationen für die Märkte von Morgen. Strategische Planung von Produkten“, Dienstleistungen und Geschäftsmodellen. Carl Hanser Verlag, München, 2019.

[2] K. Henderson and A. Salado “Value and benefits of Model-Based Systems Engineering (MBSE): Evidence from the literature,” Systems Engineering, vol. 24, no. 1, 2021, doi: 10.1002/sys.21566

[3] Fuller, Aidan et al. “Digital Twin: Enabling Technologies, Challenges and Open Research.” IEEE Access 8 (2019): 108952-108971

[4] J., Andreas, et al. “The Functional Mock-up Interface 3.0 - New Features Enabling New Applications.” Proceedings of the 14th International Modelica Conference September 20-24, 2021, Linköping, Sweden (2021), doi: 10.3384/ecp2118117.

[5] M. Becker, S. Heppner, R. Alt, T. Kleinert, und K. Schmitz „IMPROVING FLUID POWER SYSTEM SIMULATION THROUGH AN AAS-BASED SIMULATION FRAMEWORK”, in Proceedings of the ASME/BATH 2023 Symposium on Fluid Power and Motion Control FPMC2023, October 16-18, 2023, Sarasota, Florida, USA

[6] IDTA „Asset Administration Shell Specification - Part 1: Metamodel Schema“, March 2023.

[7] Modelica Association “Functional Mock-up Interface”, https://fmi-standard.org/, visited: 23.10.2023

[8] prostep ivip “Smart Systems Engineering – Collaborative Simulation-Based Engineering Version 3.0, Part E – FMI Best-Practice Guide”, https://www.prostep.org/, visited: 01.12.2023

[9] T. Blochwitz, et al. “Functional Mockup Interface 2.0: The Standard for Tool independent Exchange of Simulation Models”, Proceedings of the 9th International Modelica Conference, September 3-5, 2012, Munich, Germany, DOI 10.3384/ecp12076173

[10] IDTA „Asset Administration Shell Specification - Part 2: Application Programming Interfaces“, April 2023.

[11] IDTA „AAS Specifications“, https://industrialdigitaltwin.org/content-hub/aasspecifications, visited 01.10.2023

[12] S. Heppner, T. Miny, T. Kleinert, M. Becker, K. Schmitz, und R. Alt „Asset Administration Shells as Data Layer for Enabling Automated Simulation-based Engineering “, in 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), Sep. 2023, doi: 10.1109/ETFA54631.2023.10275474.

[14] eClass e.V. “eClass standard”, https://eclass.eu/, visited: 23.04.2023

[15] International Electrotechnical Commission “IEC 61360-4 - IEC/SC 3D - Common Data Dictionary”, https://cdd.iec.ch/cdd/iec61360/iec61360.nsf/TreeFrameset?OpenFrameSet, visited: 24.06.2023

[16] IDTA “IDTA 02003-1-2 Generic Frame for Technical Data for Industrial Equipment in Manufacturing”, August 2022

[17] IDTA “IDTA 02005-1-0 Provision of Simulation Models”, Specification, December 2022

[18] Chair of Information and Automation Systems for Process and Material Technology “basyx-python-sdk”, https://github.com/rwth-iat/basyx-python-sdk, visited 15.12.2023

[19] Parker Automation Controller Series PAC120 Operation Manual, Bulletin MSG11-5715-719/UK, 2021-12-06

[20] Electrohydraulic Control Module Series PACHC Operation Manual Bulletin MSG11-5715-720/UK, 23-02-17

[21] Fluidon GmbH “Simulation, Analyse und Optimierung fluidtechnischer Systeme“, https://www.fluidon.com/dshplus, visited: 12.12.2023

DEVELOPMENT OF AN OPEN AND MODULAR PLATFORM FOR HYDRAULICS TO INCREASE PRODUCTIVITY AND FLEXIBILITY

Marco Genise[1]*, Peer Mumcu[2]

[1]*Bosch Rexroth AG, Partensteiner Str. 23, 97816 Lohr a. Main*
[2]*Bosch Rexroth AG, Lise-Meitner-Straße 4, 89081 Ulm*

* Corresponding author: Tel.: +49 9352 18-6703; E-mail address: marco.genise@boschrexroth.de

ABSTRACT

This document describes the challenges and solution strategy for decoupling software solutions that implement hydraulic control functions from their hardware and unifying them into an open and modular software platform for hydraulics. The challenges we identified primarily concern the software engineering domain, which may be surprising since our starting point and even the resulting software building blocks are all about controlling hydraulic systems. Thus, the key solution elements are based on software engineering principles: computer hardware abstraction and model-based software development.

Keywords: Hardware Abstraction, Model-Based Software Development, Licensing, Digitialization, Software Engineering

1. INTRODUCTION

Over the last decades, hydraulic systems have evolved from purely mechanical to electromechanical systems. With the evolution of digital electronics, software emerged and took over more and more functions that were previously either realized in an electromechanical way, or perhaps not even possible. These software solutions were tightly coupled to their hardware and limited to the hydraulic setups of their actual intended target industrial applications.

This tight coupling of software and hardware resulted in a variety of different software solutions with inconsistent control behavior, user interfaces, and architectures. Subsequently, this inconsistency made it difficult for customers to understand and use all these different solutions in an efficient way. As it concerns Bosch Rexroth Industrial Hydraulics, it was rather difficult to maintain and adapt to new customer requirements.

Thus, the next step in digitalization is to decouple software solutions from their hardware and unify them into an open and modular software platform for hydraulics. Using the principles of software development, hydraulics know-how can be encapsulated in hardware-independent software building blocks that can handle the wide variability of hydraulics systems in a compact manner.

The resulting software building blocks increase usability by allowing users to implement them in their familiar development environment, perhaps even by eventually implementing them on an existing machine control device rather than in a separate motion controller. In addition, these software building blocks can be used throughout the life cycle of a machine, for example in a simulation environment during planning and design, see Figure 1.

In the following, Section 2 describes the challenges to unify these disparate software solutions into a software platform for hydraulics, followed by Section 3, which presents the solution strategy using

software engineering principles. Disadvantages of software-heavy approaches and measures to mitigate them are presented in Section 4. Finally, Section 5 concludes this paper.

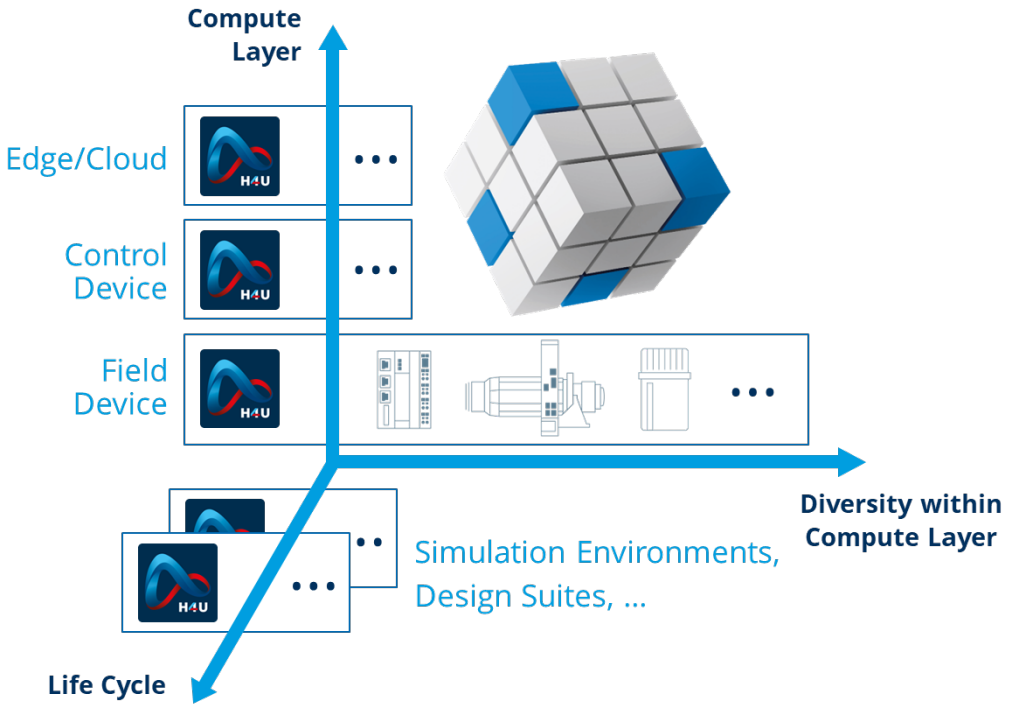


Figure 1: In this figure potential execution environments for software solutions are illustrated. The compute layer represents an aspect for different classes of computer hardware, for example field devices, control devices and edge devices. For each such class of computer hardware there are various realizations, for example the class of control devices comprises many different products from various manufacturers. This diversity within one class of computer hardware represents another aspect. The third aspect concerns the influence of the life cycle stage on the execution environment. For example, during productive use an application software is typically run on a dedicated hardware such as a control device, but during development and testing often virtual simulation environments realized with common PC hardware are used.

2. CHALLENGES

Software is commonly understood as all non-technical-physical functional components of a computer system, typically programs and the associated data. It determines the function of a software-controlled device and how it works.

As mentioned above, current software solutions already provide most of the functionality needed to control hydraulic systems, but they are still specialized for the hardware products they were designed to support. It is then, by definition, trivial to separate the functionality from the hardware. The real challenge is to unify these disparate solutions into a single software platform that can flexibly handle the variability of computer systems and hydraulic systems.

2.1. Variability of Computer Systems

The decoupling of software solutions from the computer system on which they may finally be deployed opens the opportunity of targeting broader applications than today.

A decoupled software solution is no longer limited to being implemented and used within only one compute layer. For control tasks, the possible deployment target is now a question of sufficient performance of the communication channels between the computer systems and the actuators and sensors, regardless of whether it is a traditional field device, a control device or perhaps even an edge or cloud solution. As an added dimension, the same software solution can now also be used in a different environment, such as a simulation environment or a design suite (Figure 1).

The challenge lies in the heterogeneity of (virtual) computer systems. Software can be deployed across several computing layers, in the diversity within each layer, and to other targets along a machine's lifecycle (e.g., simulation environments). Across all of these (virtual) computing systems, we encounter different computer architectures, operating systems, programming systems, and application ecosystems.

2.2. Variability of Hydraulic Systems

A major advantage of hydraulic systems is the tremendous variability of feasible solutions for a given task. Figure 2 shows some possible hydraulic setups to build a pressure and flow-controlled subsystem (p/Q control) with a motor-pump combination. The actuating variable may be the motor speed, pump displacement or even a combination of the two. Having multiple control elements also enriches the simple p/Q control with the ability to perform higher-level functions such as operating point optimization – an added value that can be dynamically provided in a software solution.

This variability is also the greatest challenge when trying to standardize higher-level functionality, such as p/Q control, in a single solution.

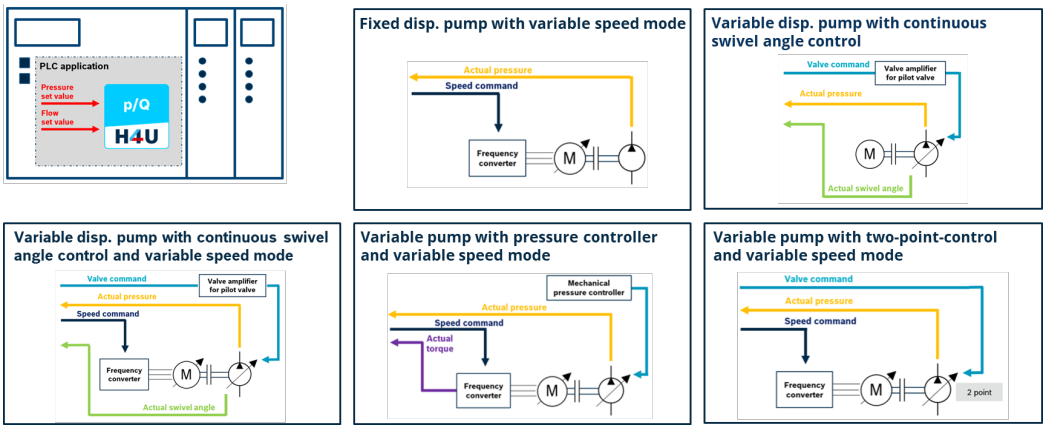


Figure 2: Examples of pressure and flow-controlled motor-pump combinations.

3. SOLUTION STRATEGY

The challenges described above should be primarily regarded as software engineering challenges. This may be surprising since our starting point and even the resulting software building blocks are all about controlling hydraulic systems. However, accepting the increased importance of software makes it less surprising that the key elements of the solution that we propose are based on software engineering principles: computer hardware abstraction and model-based software development.

3.1. Defining Computer Hardware Abstraction Layers

"Separation of concerns" [2] is a fundamental principle of software development that divides an application into distinct sections. Separation of concerns is achieved by encapsulating knowledge within a section of code that has a well-defined interface. This results in a high degree to which parts within one section belong together (strong cohesion) and a low degree of interdependence with parts of other sections (low coupling).

When applying this principle to computer systems, we can design computer hardware abstraction layers by identifying logical layers in data/signal processing, separating these layers from each other (decrease coupling), and making each layer its own compact section of parts that belong together (increase cohesion).

Figure 3 shows the chosen layer architecture, which is common in other industrial standards such as AUTOSAR [3]. The top layer, the device function layer, implements the actual hydraulic control functions in a hardware independent manner and has a well-defined interface based on physical dimensions used in the application domain. The middle layer, the control device abstraction layer, serves as an abstract binding between the hardware-independent device function layer and the actual (potentially virtual) computer system. It provides machine-level abstraction by adapting system calls and converting data types, and transforms this data into the semantic information required by the device function layer.

The most important advantage of this approach is that a simulation environment or design suite can now be treated as a virtual computer system. This allows the device function layer to be used there by providing an appropriate control device abstraction layer, e.g., to enable application testing in a virtual environment.

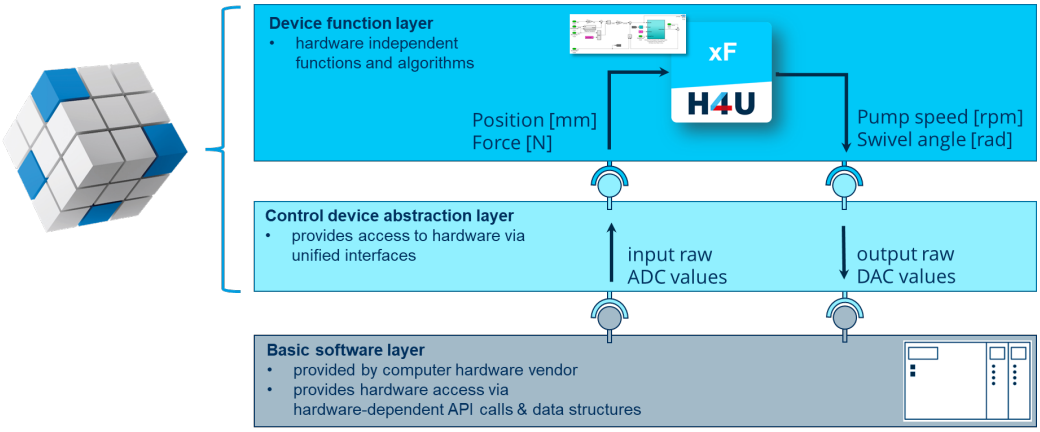


Figure 3: Layer architecture of Bosch Rexroth's Industrial Hydraulics software platform H4U.

The example shown in Figure 3 comprises reading sensors with electrical interfaces, processing the signal through the control device abstraction layer into physical units as input to the device functions, and then reversing this chain to finally output a voltage or current signal to control the actuator. The application in the device function layer is relieved from concerns about how a particular piece of electronics reads voltage signals, what kind of corrections or calibrations need to be made, or how it generates its output signals.

This principle is not limited to the abstraction of analog electrical interfaces, but also applies to sensor and actuator signals transported over digital interfaces or fieldbuses.

3.2. Model-Based Software Development

In software engineering it is common practice to express domain aspects in a language that fits the domain. This is a result of the understanding that general purpose programming languages are often difficult to use as a means of communication between the functional domain and the software engineering domain. In addition to these domain-specific languages, there are entire development methodologies that reflect the domain focus, such as Domain-Driven Design [4].

The idea behind model-based software development is to describe the actual system functions in a language familiar in the system domain to express its structures and processes on a higher level of abstraction. This language abstraction is another fundamental principle of software development. Here, MATLAB Simulink is used to express the system functions in a language that is much closer to the domain of hydraulic control than general-purpose programming languages such as C, see Figure 4.

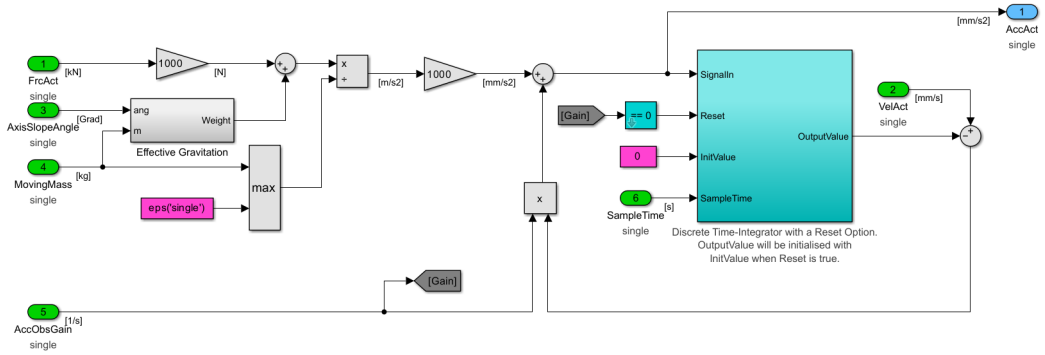


Figure 4: Example: Luenberger observer implementation in MATLAB Simulink.

Since the modeling language is a formal language, the models can be translated into a semantically equivalent program in another formal language, such as IEC 61131-3 Structured Text or C. This principle is used to compile the device function layer into program code for different development environments like TIA Portal, TwinCAT or ctrlX PLC, as shown in Figure 5. This process is commonly known as code generation.

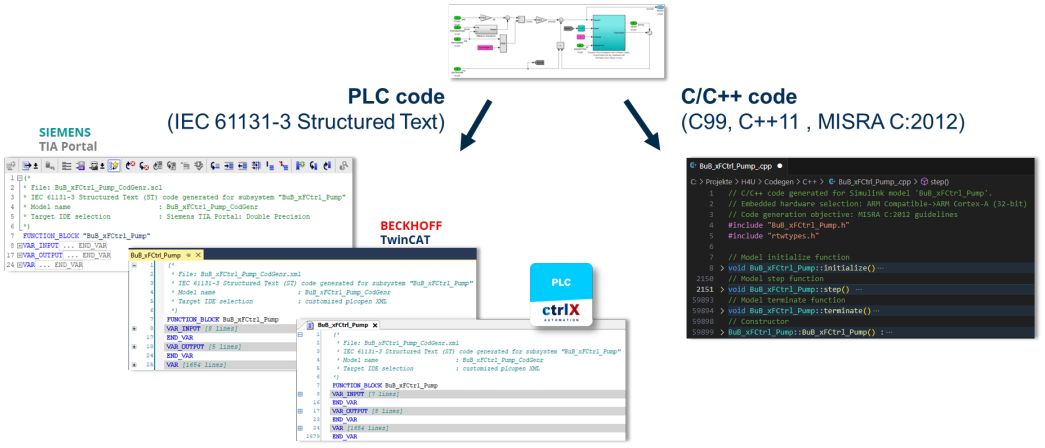


Figure 5: Compiling model to target specific programming languages.

Model-Based Hydraulic Functions Are Software

As mentioned above, model-based software development still results in programs. Therefore, these programs are subject to the principles of software development. To manage the complexity of the hydraulic control function, separation of concerns is used to separate layers, as shown in figure 6. In our implementation to control position and force for a differential cylinder in a closed hydrostatic transmission, the control elements are decoupled using separate layers: a control-layer, the inverse (flow → linear motion) transformation layer and the inverse (rotary motion → flow) transformation layer, representing control, cylinder, and motor-pump units. Any modification in the axis arrangement will concern the arrangement of motor-pump units and will therefore restrict software variations to the last functional layer [1].

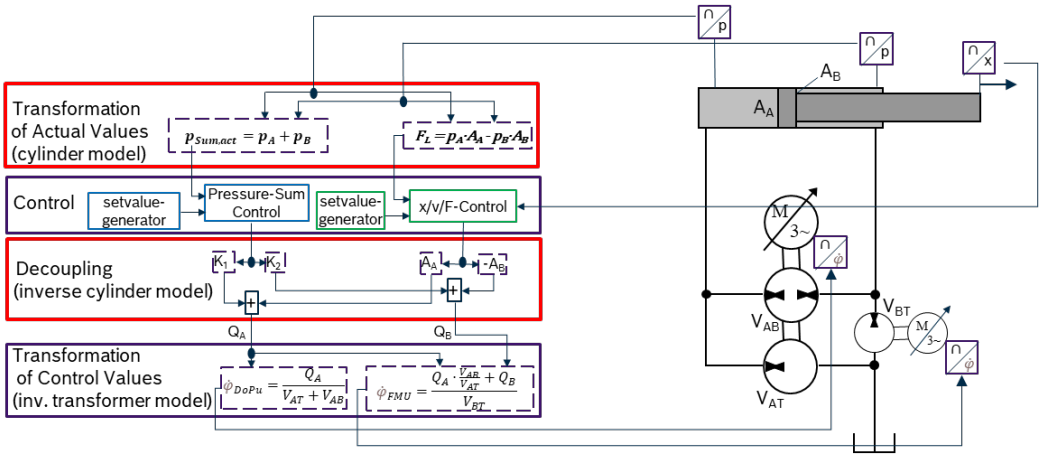


Figure 6: Example: software structure of a position- and force-controlled application [1].

4. DISADVANTAGES OF SOFTWARE-HEAVY APPROACHES

By introducing software-heavy approaches into hydraulics development, we not only gain a lot of efficiency advantages, but we also have to deal with new challenges, namely intellectual property theft and general IT security issues.

By decoupling software functions from their hardware, another fundamental property of software emerges: the marginal cost of replicating software is essentially zero. To avoid economic damage, additional measures must be taken to mitigate the uncontrolled distribution of such software solutions. Because software is usually easily modified, it is also necessary to detect modifications and protect integrity. Otherwise, the proper functioning of the software solution cannot be guaranteed.

Typically, these measures are implemented using cryptographic primitives. However it should be clear that the success and residual risk of any of these measures depends on the level of trust available in the target computer system.

4.1. Cryptography Primer

This section provides a basic introduction to cryptography, which is helpful in understanding the measures taken to mitigate the disadvantages of software-heavy approaches.

A cryptographic primitive is a well-established, low-level algorithm that is used to construct more complex cryptographic systems. These algorithms include one-way hash functions and encryption functions.

A one-way hash function is a function that takes arbitrary data of any length and outputs a short, fixed-length hash value. A key property of a one-way hash function is that it is unfeasible to find input data that matches a known hash value (also called pre-image resistance).

Encryption functions create an unintelligible blob of data (also called ciphertext) from the actual data in order to ensure its confidentiality. The operation of an encryption function usually depends on a piece of auxiliary information called a key. Without knowing the key, it is extremely difficult to decrypt the resulting ciphertext back to the original input data.

In public-key cryptography – also called asymmetric cryptography – two keys are used, one for encryption and one for decryption. Depending on the application, one of the two keys can be made public, that is, not kept secret. For example, if a sender, Alice, wants to send a message to Bob, Alice uses Bob's public encryption key, which is made available to anyone who wants to encrypt a message and send it to Bob. Alice then encrypts the message using Bob's public key before sending the ciphertext to Bob. Only Bob can decrypt the ciphertext using his private decryption key, which remains secret.

4.2. Integrity Protection

Digital signature mechanisms are usually used to detect modifications and protect integrity of a software building block. A digital signature is essentially a hash value that is encrypted with the private key of an asymmetric cryptographic algorithm. The hash value is generated by a one-way hash function given the software building block as input data. This digital signature is distributed with the software component.

When the software starts, it computes the hash value of itself, decrypts the digital signature with the public key, and compares the hashes - and refuses its operation if the hashes aren't equal, meaning the software building block has been modified.

4.3. Mitigation of Uncontrolled Distribution

A common solution to mitigate uncontrolled distribution is to link the software application to a physical reference point using licensing mechanisms. For example, a physical reference point, such as a unique hardware identifier, is placed in a file. This file is then digitally signed as described above. This file and its digital signature make up the license file that is provided to the software application.

At startup, the software verifies the validity of the license file by proving that it has not been modified. It then compares the physical reference point stored in the license file with the one provided by the computer system. If either of these steps fails, the software will refuse its operation.

5. CONCLUSION AND OUTLOOK

In conclusion, this paper provides an exploration of the challenges in developing a capable software platform for hydraulic control functions. Recognizing the pivotal role of software engineering principles in overcoming these challenges, our proposed solution strategy hinges on two fundamental pillars: computer hardware abstraction and model-based software development.

However, the proposed transition to a software-heavy approach also introduces challenges such as intellectual property theft and security concerns. The outlined mitigation measures serve as initial safeguards. Continuous research and development are imperative to fortify these measures, ensuring the integrity and controlled distribution of such software solutions.

Looking ahead, the envisaged software platform holds the potential for deployment on cutting-edge computing architectures, including edge and cloud solutions, extending its reach beyond traditional

hardware constraints. This expansion not only ensures scalability but also positions the hydraulic control functions to operate seamlessly in distributed and connected environments, opening the possibility for regular updates, which may provide optimizations and support for new hydraulic components. Furthermore, the proposed approach establishes a fertile ground for applying data analytics. By consolidating data from diverse sources, the software platform can enable its users to leverage analytics for various purposes, for example performance optimization and resource utilization.

The outlook for the proposed software platform extends beyond the immediate challenges addressed in this paper, envisioning a future where hydraulic control functions not only meet the demands of today but also adapt and thrive in an ever-evolving technological landscape. The integration of edge and cloud computing, coupled with internet-enabled updates and data analytics, positions the software platform as a dynamic and future-ready solution, driving efficiency, sustainability, and innovation in hydraulic systems.

REFERENCES

[1] Bonefeld R, Gellner T, Brandt L, Bauer K (2024) Precise hydrostatic Cylinder Drive with increased Pressure Level for Industrial Applications. 14th International Fluid Power Conference. Dresden, March 19-21, 2024, Dresden, Germany

[2] Dijkstra, Edsger W (1982). On the role of scientific thought. In: Selected writings on Computing: A Personal Perspective. Texts and Monographs in Computer Science. Springer, New York, NY

[3] AUTOSAR (AUTomotive Open System ARchitecture). https://www.autosar.org

[4] Evans, Eric (2003). Domain-Driven Design: Tackling Complexity in the Heart of Software. Addison-Wesley, Boston

PHYSICAL IMPLEMENTATION OF A DISTRIBUTED, AGENT-BASED CONTROL FOR FLUID SYSTEMS USING OPC UA

Tobias C. Meck[1]*, Oscar L. Lefemmine[1], Kevin T. Logan[1], Peter F. Pelz[1]

[1]*Chair of Fluid Systems, Technische Universität Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt*

* Corresponding author: Tel.: +49 6151 1627126; E-mail address: tobias.meck@tu-darmstadt.de

ABSTRACT

Conventional control strategies for fluid systems often rely on local control of the system's components, like pumps and valves. Here, communication between the control units is non-existent, which can have a negative impact on the energy efficiency. Distributed control is a promising alternative where so-called agents are assigned to components. These agents are autonomous units with individual goals. They can perceive and influence their environment through sensors and actuators. Furthermore, they are able to share information with each other. This leads to an increased energy efficiency while maintaining the positive aspects of local control, such as a low implementation effort and high robustness. The concrete methods are the subject of current research and are typically only verified in simulations. For a thorough evaluation and broad acceptance in industry, an assessment of the methods when facing real systems is crucial.

In this work, we therefore focus on the physical implementation of distributed control. We examine a simple fluid system with a centrifugal pump and a valve. A valve agent measures its volumetric flow rate and communicates this information to a pump agent via Wi-Fi and OPC UA. The pump agent has the goal of achieving a target flow by using a PI controller and adjusting the rotational speed. The results are promising and easily scalable to more complex systems and control methods.

Keywords: Multi-agent systems, OPC UA, Microcontrollers, Control, Robustness

1. INTRODUCTION

When it comes to the energy efficiency of fluid systems, such as the water supply in buildings or whole cities, expectations usually do not meet reality. Society demands special attention to be paid to the matter. This is once again confirmed when looking at the recently revised Energy Efficiency Directive 2023/1791 from the European Union, which puts "Energy Efficiency First" [1].

Still, conventional control strategies for fluid systems are usually not tailored for efficiency. They use *local control* of components, mostly valves, to reach certain set points for the pressure or the volumetric flow rate, e.g., using PID controllers. In this way, the introduced energy from pumps gets dissipated due to throttling losses.

These losses can potentially be avoided if additional information of the system, such as the current state or the topology, is used [2]. This information is used to decide if it is possible to reduce the introduced energy, e.g., by reducing the rotational speed of pumps, rather than throttling excessive energy, e.g., by shutting down valves. In *central control*, the information from all components and available sensors is collected and possibly enriched by a system model in a central controller. Using techniques from mathematical optimization, it is then even possible to calculate optimal operating modes with respect to efficiency, as shown in [3].

While this method theoretically offers the greatest possible improvement in this dimension, Logan et al. [2] list further requirements to control approaches that have to be taken into account. Looking at fluid systems as a vital part of critical infrastructure or production systems in industry, it becomes clear that the robustness of the control is of great importance. Perturbations in the operating phase, such as unforeseen load cases or wear of components up to component failures, should have a minimal influence on the overall function of the system. Central control systems can cope with slight perturbations, but pose a single point of failure and often rely on sufficiently accurate models. Local control, on the other hand, can in most cases still maintain a basic functionality, even if failures occur.

Another aspect worth considering is the implementation effort of control systems. Fluid systems are getting increasingly complex and individual. Modelling is therefore cumbersome and accompanied with various uncertainties, for instance regarding the calculation of pressure losses. Over time, there might also be changes or extensions of the system, requiring further adaptions. These reasons serve as a possible explanation why local control is still preferred in a lot of applications.

Aiming at combining the advantages of local and central control, Logan et al. [2] focus on *distributed control*. Here, the local controllers are designed as agents, which can access the sensors and actuators of the connected components and thus perceive and influence their environment. In contrast to local control, they are also able to exchange information through communication. Furthermore, they have individual goals. For a pump agent, this might be the desire to minimize the pumps input power.

The concrete methods differ with regards to the decision rules of the agents, i.e., how they manipulate the actuators based on measurements and acquired information due to communication to reach their goals. The authors compare approaches from optimization theory, machine learning and game theory in simulations. The latter approach, referred to as *market mechanism*, is particularly promising. Here, a virtual budget is assigned to the agents, which is used to trade volumetric flow rate guarantees. Similar to a real market, one expects to reach an efficient allocation without explicit system knowledge. As this method therefore does not rely on system models or extensive training, it is highly flexible and transparent, which is beneficial for the acceptance in practical applications.

Regarding the acceptance, it is also crucial to proof the validity of the concepts using experiments and real systems. The step from simulations to experiments often offers additional insights and is necessary for an overall evaluation. A first step towards a validation of the methods controlling real systems was presented in [4]. While the simulation of the fluid system was exchanged by the real counterpart, the agent system was still simulated on a central machine. Nevertheless, the results show a significant increase of the energy consumption compared to the simulation, which does not consider dynamics, such as limited opening rates of valves. Replacing the simulation of the agent system introduces additional complexity, such as limited computing resources of edge devices and communication latencies. In this work, we focus on establishing the necessary framework that enables an all-embracing validation of distributed control methods for fluid systems. We apply the results to a minimal example of distributed control that is outlined in the next section.

2. CONTROL TASK

As a model system, we consider the water supply of a residential building. In order to have a sufficient pressure on the higher floors, it is necessary to install pumps, which are usually placed in the basement of the building. To control the volumetric flow rate, the different floors are equipped with valves.

This model system can be scaled down to a test rig, that is shown in **Figure 1**. Water is pumped by a booster station from a tank to five floors, which are equipped with valves and flow sensors. The available pumps are centrifugal pumps used in heating applications. They possess integrated frequency converters which allow an adjustment of the rotational speed n. It is possible to measure

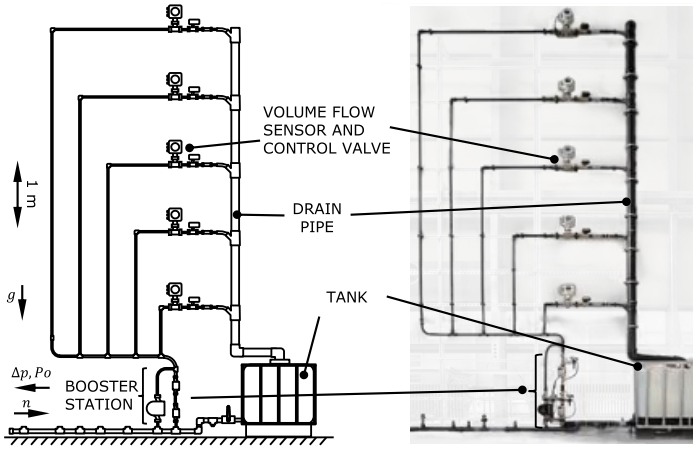


Figure 1: Schematic (left) and photograph (right) of the modular test rig [5].

the pressure difference Δp and the electrical power consumption Po of each pump. From the different floors, the water then returns to the tank through a drain pipe. Because of the modular design of the test rig, it is possible to realise different topologies. For a detailed description of the test rig we refer to [5].

In this work, we only consider the simple case where a single pump is placed in the booster station and all valves except for the valve in the first floor are shut. The control task is to fulfil a target flow rate by using a PI controller, which is assigned to the pump. This serves as a minimal pump agent. The counterpart of the pump agent is the combination of the valve and the flow sensor on the first floor, which constitutes a valve agent. As the actual flow rate is not directly accessible to the pump agent, it needs to acquire this information through communicating with the valve agent.

3. DEVELOPED FRAMEWORK

3.1. Hardware

To be able to realize the proposed control task, it is first and foremost necessary to physically represent the agents with appropriate hardware. For this purpose, ESP32-S3-DevKitC-1 [6] development boards are used as a basis, which combine an ESP32-S3 microcontroller with additional peripherals to form a programmable PCB board. This allows to program and process the internal logic of the agents. The microcontrollers are equipped with 512 KB of SRAM, up to 16 MB of Flash and have a maximum clock speed of 240 MHz. Additionally, they include Wi-Fi capabilities, which can be used for the physical layer of the agent communication. However, as this requires an additional access point which poses a single point of failure, an Ethernet interface is connected to the development board which can be used as an alternative. To be able to connect the sensors and actuators of the agents to the general-purpose input/output (GPIO) pins of the microcontroller, further components, like connectors, voltage dividers and Digital-to-Analog-Converter (DAC) modules, are installed. With this, input signals from 0-10 V/ 4-20 mA and output signals from 0-10 V are possible. The final hardware is shown in **Figure 2**. A distinction between pump agents and valve agents is done solely on the software-side, which enables a straightforward addition of other agent types.

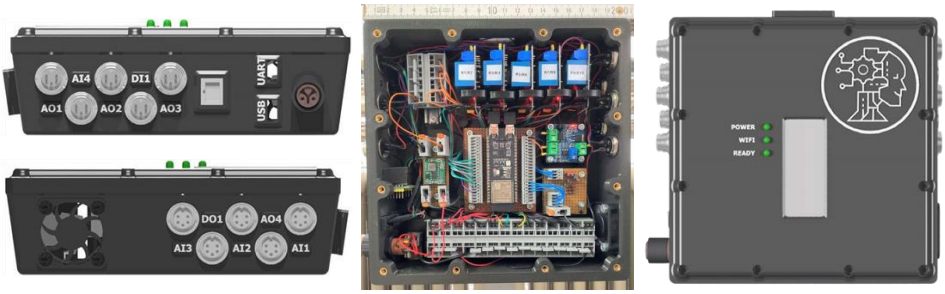


Figure 2: Side view (left) and top view with open (middle) and closed lid (right) of the hardware implementation of the agents

3.2. Software

The software of the agents has to perform different tasks. On the one hand, the control algorithms have to be carried out, which are specific to the subtype of distributed control being considered. For the example in the scope of this work, this reduces to a discrete PI controller for the pump agent. Associated with this are also the sensor readings and the actuator control.

On the other hand, the software needs to manage the connection and communication of the agents. This normally involves additional aspects, such as the discovery of other agents in the network and their topological relationship, i.e., if another agent is connected upstream, downstream or in parallel. We consider this information given. As a possible extension, network discovery can easily be achieved by performing mDNS or ARP scans.

OPC UA

For the communication of the agents, we employ the platform independent machine-to-machine standard OPC UA (*Open Platform Communication Unified Architecture*) [7]. OPC UA offers a wide range of features, such as security settings, that partially extend the scope of our work. We therefore refer to the official reference [7] for additional information.

In OPC UA, not only the raw data transport is considered, but rather the transport of information through an extensive information modelling framework. In essence, this information model is a graph consisting of nodes and references between them [8]. Different node types are available, e.g., object nodes and variable nodes. Object nodes represent physical or abstract elements of the system and may contain further nodes. Variable nodes represent values (e.g., sensor data) that can be read or written and have an associated data type. The nodes have different attributes, such as unique node ids, descriptions or time stamps [7].

For the lower-level transport layer, different protocols are available, of which the binary TCP-based protocol used here is the most common [8]. Depending on the chosen protocol, two different communication models, namely client-server and publish-subscribe (PubSub), can be applied. We focus on the client-server communication model, although OPC UA PubSub is a promising alternative for applications with very high latency requirements, as demonstrated in [9].

In the client-server-model, servers offer services to clients, such as reading or writing of variable nodes values. The question if a particular component should be implemented as a client or server depends greatly on the use case. As advocated by Rinaldi [10], assigning a server *and* clients to components might be beneficial. In our use case, we implement every agent that has access to data that needs to be shared, e.g., from sensors, as a server. If an agent needs to access data from other

agents, it additionally needs to implement a client for each of those agents.

For the servers, a set of nodes has to be defined. The valve agent's server contains a variable node for the volumetric flow rate, which gets updated through readings of the respective input. An additional variable node is the valve opening. Writing to this node allows clients to directly influence the valve agent's output. Similarly, the pump agent's server contains nodes for the pressure difference, the rotational speed, the on/off state and the volumetric flow rate. The latter is updated whenever the valve agent communicates new values. Additionally, the pump agent's server contains a PI controller object node, which in turn contains variable nodes for the set point and the gains as well as an activity node that can be used to enable and deactivate the control. As described above, these node sets constitute basic information models.

For the data exchange, we make use of a subscription concept defined by the OPC UA standard. With this, it is possible to make servers send out notifications at a predefined rate, if the value of a variable node has changed. This especially facilitates the reading of constantly changing values. In our example, the pump agent subscribes to the value change of the volumetric flow rate node of the valve agent's OPC UA server at a publishing rate of 10 Hz.

To implement the OPC UA concept, the open source OPC UA stack open62541 [8] is used. The stack is written in the C programming language and offers support for FreeRTOS [11], which is a well-established operating system for ESP32 microcontrollers. As a basis for our software, a GitHub project [12] by GitHub user *cmbahadir* was used, which implements a simple OPC UA server based on the open62541 stack. Additionally, the project implements the necessary Wi-Fi connection and a time synchronisation via a NTP server.

Control Dashboard

For the purpose of logging and supervising the control process, which includes visualisation, starting the PI control or manually adjusting the pump's rotational speed, a central computer is used. As for instance the PI control could also be enabled directly from the start in the microcontroller program, this does not impair the idea of a distributed control but rather simplifies the conduction of experiments.
The central computer hosts an HTTP server based on the Python framework Flask [13] with a website based on the JavaScript library Highcharts [14] shown in **Figure 3**. This acts as a control dashboard. Two additional OPC UA clients for the valve and pump agents, based on the Python library opcua-asyncio [15], share data with the HTTP server. In this way, reading information (e.g., the current volumetric flow rate) and writing information (e.g., the gains of the PI controller) of the OPC UA servers is possible from any computer or smart device in the same network without additional software. For the visualization of sensor values, the aforementioned subscription concept is again applied.

The final framework is summarized in **Figure 4**.

4. RESULTS

To prove the functionality of the presented framework, the control task described in section 2 needs to be assessed. For this purpose, two different experiments are conducted. In the first experiment, the pump is set to its maximum rotational speed before starting the PI control. As a set point, a volumetric flow rate of 1 m^3/h is specified. In a second experiment, the pump is set to its minimum rotational speed before starting the PI control to reach a set point of 3 m^3/h.

The controller is parametrized for the first set point starting with the well-known Ziegler-Nichols method with a critical gain of 0.38 and a critical period of 6.7 s. Afterwards, the proportional and

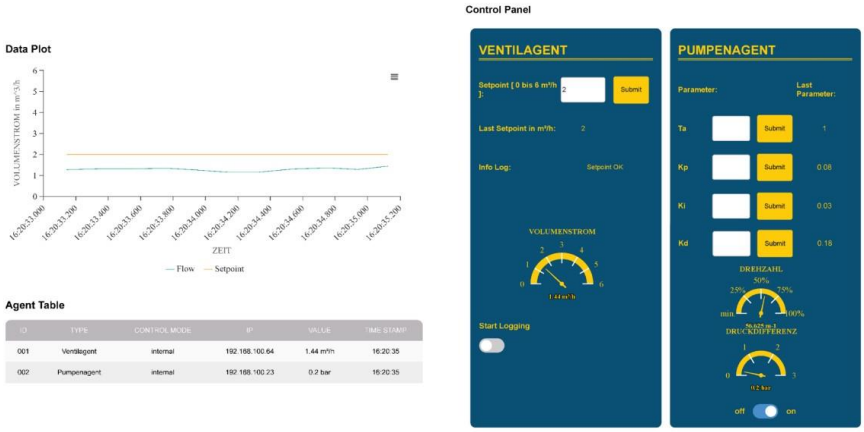


Figure 3: HTML page based on the JavaScript library Highcharts [14], which serves as a control dashboard.

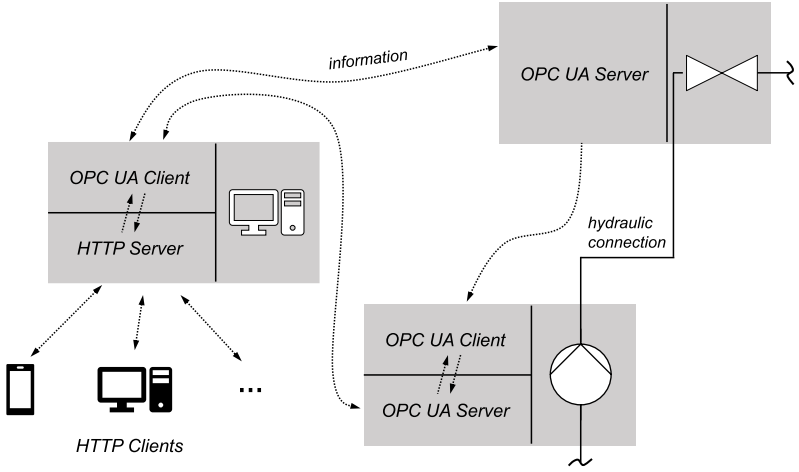


Figure 4: OPC UA-based distributed control framework.

integral gain are slightly adjusted for a better controller performance. The sampling rate is set to 2 Hz. As shown in **Figure 5**, the control starts after approximately 20 s and is able to reach the set points quickly. However, the volumetric flow rate signal from the valve agent is highly noisy, which leads to an even higher fluctuation around the set point due to the PI control. Worth noting is also the apparent increase of the volumetric flow rate in the first experiment before the set point is reached. A possible explanation is that the PI controller, having a different subscription than the logging computer, receives a value below the set point, causing an increase of the rotational speed of the pump.

Looking at the sample mean after sufficient time, one can conclude that the control at least on average reaches the set point with sufficient precision. **Table 1** shows the numerical values of the mean and the standard deviation of the volumetric flow rate signal.

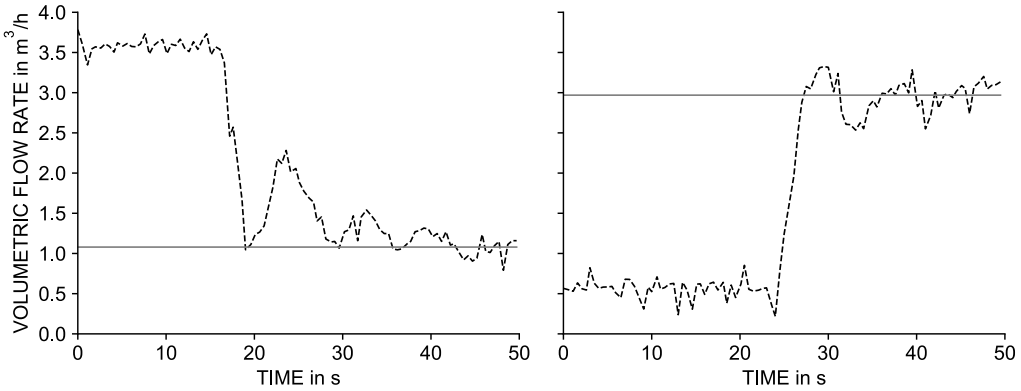


Figure 5: PI control starting from maximum speed with a setpoint of 1 m^3/h (left) and from minimum speed with a setpoint of 3 m^3/h (right). The gray line indicates the mean volumetric flow rate after 40 s. The flow measurements are taken from the valve agent's OPC UA server.

Table 1: Sample statistics of the flow rate in m^3/h (100 samples per experiment at 2 Hz).

Experiment	Set point	Mean, last 10 seconds	(Corrected) standard deviation, first 10 seconds	(Corrected) standard deviation, last 10 seconds
1	1	1.0795	0.086	0.126
2	3	2.969	0.098	0.161

Another important aspect of the presented framework is the performance of the communication. As defined by the OPC UA standard, a *sourceTimestamp* is applied by an OPC server at the data source to a variable value at each change, and a *serverTimestamp* denotes the time when a server receives the value. The difference between these two timestamps can be used to approximate the overall latency of the communication. In our example, the valve agent's OPC UA server assigns a sourceTimestamp when performing a reading from the flow sensor. When the pump agent's OPC UA client receives this value through the subscription, the variable value gets written to the corresponding variable node alongside its sourceTimestamp. The serverTimestamp, on the other hand, gets set to the current time. For a meaningful comparison, the clocks of both agents need to be sufficiently synchronized, which is achieved by a NTP server. Still, slight clock drifts might be possible. It is also important to note that additional computation steps of the microcontrollers might further increase the total latency. For both experiments, the resulting differences are shown in **Figure 6** and **Table 2**. While the mean and median latency are acceptable, especially considering the connection over Wi-Fi, large outliers can be detected. Still, the results validate the use of a sample rate of the PI controller and the logging of 2 Hz.

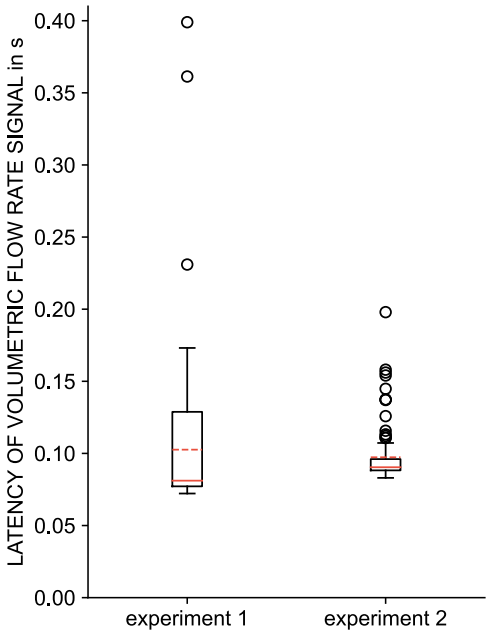


Figure 6: Boxplots of the difference between server and sourceTimestamps of the volumetric flow rate as a measure of latency. Sample size = 100. Maximum whisker length = 1.5 · IQR (Interquartile range). The dashed lines represent the mean and the solid lines the median of the data.

Table 2: Sample statistics of the difference between server and sourceTimestamps of the volumetric flow rate as a measure of latency in seconds. Sample size = 100.

Experiment	Mean	Median	Maximum	(Corrected) standard deviation
1	0.103	0.081	0.399	0.048
2	0.097	0.090	0.198	0.019

5. CONCLUSION

In this work, we introduced a framework for the validation of distributed control algorithms. To physically represent the agents, microcontrollers with appropriate peripherals are used. The communication between the agents is achieved by Wi-Fi and OPC UA. In this context, a OPC UA server is assigned to each agent. Also, an additional OPC UA client is assigned to an agent for each agent that it needs to communicate with. The additional integration of a central computer facilitates logging and supervising of the distributed control. The framework was assessed using a test rig, which represents the water distribution in a residential building. In the test setup, a pump agent employs a PI controller to reach a target volumetric flow rate. As the pump does not have direct access to a flow rate sensor, it needs to acquire the information from an upstream valve agent. This is achieved by subscribing to the value changes of the volumetric flow rate node in the valve agent's OPC UA server address space. Though the results show a high fluctuation of the sensor readings and harsh outliers in the communication latency, the control task can in general be achieved satisfactorily.

To tackle the latency of the communication, Ethernet connections are a natural alternative to Wi-Fi. These in turn pose additional questions, e.g., regarding data transmission topologies.

As pointed out by the manufacturer of the microcontrollers, the integrated ADCs are sensitive to noise, which explains the high fluctuations of the volumetric flow rate readings. To mitigate this issue, the manufacturer suggests the usage of bypass capacitors, which could be examined in a next step.

The developed framework is readily scalable to more sophisticated distributed control tasks. The server-client principle of the agents allows new agents to enter the network in a simple "plug-and-play" manner. However, further considerations have to be taken into account. For one, additional clients pose additional memory requirements and computational complexity for the microcontrollers. This might be unfeasible for systems, where an agent has to be able to communicate with a large number of peers. In such settings, OPC UA PubSub is worth considering. OPC UA PubSub is especially relevant for one-to-many, many-to-one or many-to-many communication settings, but does not reliably ensure the data transmission. Looking at the different variants of distributed control, particularly the market mechanism, a reliable one-to-one communication might however be important.

Especially the more sophisticated distributed control methods require richer information models than those that are considered in this work. In a first step, predefined node sets from the OPC Foundation might be used. Further research could aim at developing an information model to fully describe agents, irrespective of their type. Taking the market mechanism as an example again, the information model should also be capable of mapping complex communication patterns, such as negotiations, to make the framework universally applicable.

ACKNOWLEDGMENT

The authors would like to thank Manuel Rexer and Niklas Puff for their valuable input on designing the hardware solution. Results were obtained in project no. 22514N/1, funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) approved by the Arbeitsgemeinschaft industrieller Forschungsvereinigungen "Otto-von-Guericke" e.V. (AiF).

REFERENCES

[1] European Union (2023) Directive (EU) 2023/1791 of the European Parliament and of the Council

[2] Logan KT, Stürmer JM, Müller TM et al. (2022) Comparing Approaches to Distributed Control of Fluid Systems based on Multi-Agent Systems

[3] Müller TM, Knoche C, Pelz PF (2022) From Design to Operation: Mixed-Integer Model Predictive Control Applied to a Pumping System. In: Trautmann N, Gnägi M (eds) Operations Research Proceedings 2021. Springer International Publishing, Cham, pp 318–324

[4] Logan KT, Stürmer M, Müller TM et al. (2023) Multi-Agent Control of Fuid Systems – Comparison of Approaches.

[5] Müller TM, Leise P, Lorenz I-S et al. (2021) Optimization and validation of pumping system design and operation for water supply in high-rise buildings. Optim Eng 22:643–686. https://doi.org/10.1007/s11081-020-09553-4

[6] Espressif Systems ESP32-S3-DevKitC-1. https://docs.espressif.com/projects/esp-idf/en/latest/esp32s3/hw-reference/esp32s3/user-guide-devkitc-1.html. Accessed 03 Feb 2024

[7] OPC Foundation OPC Unified Architecture. https://reference.opcfoundation.org/. Accessed 04 Feb 2024

[8] open62541 open62541 documentation. https://www.open62541.org/doc/master/. Accessed 04 Feb 2024

[9] Pfrommer J, Ebner A, Ravikumar S et al. (2018) Open Source OPC UA PubSub Over TSN for Realtime Industrial Communication. In: 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA).

[10] John S. Rinaldi OPC UA Client vs. Server. https://www.rtautomation.com/rtas-blog/opc-ua-client-vs-server/. Accessed 04 Feb 2024

[11] Amazon Web Services FreeRTOS. https://www.freertos.org/fr-content-src/uploads/2018/07/FreeRTOS_Reference_Manual_V10.0.0.pdf. Accessed 05 Feb 2024

[12] cmbahadir opcua-esp32. https://github.com/cmbahadir/opcua-esp32. Accessed 05 Feb 2024

[13] Pallets Flask. https://flask.palletsprojects.com/en/3.0.x/. Accessed 05 Feb 2024

[14] Highsoft AS Highcharts. https://www.highcharts.com/. Accessed 05 Feb 2024

[15] FreeOpcUa opcua-asyncio. https://github.com/FreeOpcUa/opcua-asyncio/?tab=readme-ov-file. Accessed 05 Feb 2024

Chapter 14

Pneumatics

SIZING OF PNEUMATIC DRIVES UNDER ENERGY EFFICIENCY ASPECTS

Matthias Doll[1]*, Rüdiger Neumann[1], Wolfgang Gauchel[1]

[1]*Festo SE & Co. KG, Ruiter Str. 82, 73734 Esslingen*

* Corresponding author: Tel.: +49 711 347-57380; E-mail address: matthias.doll@festo.com

ABSTRACT

The correct sizing of pneumatic drives plays a central role when it comes to energy efficiency. While there are simple design formulas for force-based tasks such as pressing or clamping in order to size the drive efficiently, there is no such easy methodology for motion tasks.

Up to now, the sizing of pneumatic drives has mainly been experience-based or simulation-based. A tool from Festo [1] now enables formula-based sizing without simulation, which directly provides the optimum piston diameter and other components. The approach behind this is based on the natural frequency of the pneumatic drive.

The main drawback of this method is that it is only applicable for horizontal installation positions. Based on more recent findings (which arose in a joint project with the TU Dresden [2]) and based on numerous simulations and measurements, this formula has now been extended so that it can also be used with external forces and thus also for a vertical installation position.

Keywords: pneumatic drives, energy efficiency, sizing

1. INTRODUCTION

Pneumatic drives are used in numerous applications in industrial machines. They are known for their simplicity, robustness, high power density, compact design and high forces. They are controlled by a switching valve in combination with exhaust air throttles to adjust the speed. The energy consumption only depends on the internal volume of the cylinder. In the age of global warming, the efficiency of drive systems becomes increasingly important.

The simplest way to increase the efficiency of the drives is the sizing of the cylinder in order to reduce its volume. **Dimensioning** generally describes the process of assigning system components (such as cylinders, valves, etc.) to a given application. In addition to the application parameters such as stroke, force, transition time and moving mass, several additional factors like end position energy, vibrations, requirements on the motion profile (velocity, acceleration), guide load, lateral forces and special requirements like ATEX or LABS-free etc. are important.

Once a cylinder has been found that can solve the application, additional requirements can be realized by selecting a suitable series and additional options. However, even the sizing process of the required piston diameter is not trivial for many applications and therefore often results in oversized drive systems. Other reasons for oversized drive systems include:

- Safety factors and robustness issues
- Storage costs for spare parts (reducing variants, one size for all applications)
- Usage of cylinders as construction elements (mechanical stability)

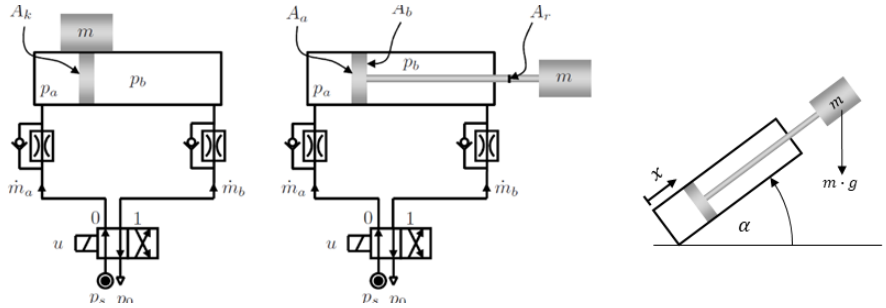


Figure 1: Standard pneumatic system using a rodless cylinder (*left*) and cylinder with rod (*middle*), standard switching valves and exhaust flow throttles for speed adjustment. *Right:* orientation.

In principle, a distinction must be made in the **dimensioning** between force and movement tasks.

Force tasks are applications in which a force plays the central role within the process, e.g. in pressing processes, rolling, joining, etc. They are relatively easy to design, as the specified force F is in a direct, algebraic relationship with the piston area A_k and the pressure difference Δp and therefore can be used for calculating the effective area:

$$F = A_k \cdot \Delta p. \tag{1}$$

Motion tasks are applications that require a movement within a specified time T_f. They are typically described by the moving mass, stroke, transition time, and a pressure level. For motion tasks an easy relationship as in (1) cannot be represented. In contrast to controlled electric drives, the movement profile and therefore the acceleration profile cannot be specified directly. Instead, it results from the selected system components and the throttle values. The aim of the design here is to achieve a defined transition time. Differential equations are used to describe the system behaviour [3]. The determination of the transition time as well as the design process itself is therefore often simulation-based in combination with a variation of the throttle positions and the drive size. An important aspect in the design for motion tasks is the **kinetic energy** in end position, as this must be below a specified limit so that the drive can be operated permanently without destruction. This boundary condition determines the feasibility of an application with a selected system. Mixed forms of force and movement tasks are also frequently used.

The **energy consumption** of a pneumatic actuator can be calculated directly from the air consumption [4]. In standard operation (switching valve + exhaust air throttling), it depends solely on the selected size (and the dead volume) and is independent of the throttle position and the resulting transition time. This makes the **sizing** particularly important for reducing energy consumption. Ultimately, it leads to the central question:

"What is the smallest possible cylinder to fulfil the motion task?"

1.1. Starting point

We consider standard pneumatic drives consisting of a cylinder with exhaust air throttles and an adjustable pneumatic cushioning system and a switching valve as shown in **Figure 1**. The design methodology focuses on motion tasks (with additional force requirement), i.e. on the sizing of the system components for a specified application given by: mass, stroke, transition time, supply pressure, external force and a permissible end position behaviour (compliance with the maximum permissible residual energy). The focus here is on energy efficiency, so that a design methodology for medium to large drives is required, as small or short stroke drives require only little energy in absolute terms.

2. STATE OF THE ART

In addition to simulation-based design methods, there are efforts to map the non-linear and complex processes of the dynamic model in algebraic formulas for a sizing purpose. Even if the force and acceleration profiles result from the application and can only be specified in a highly simplified form, there are a number of methods that attempt to do just that. These include force equilibrium method [5], minimum cylinder [3], exergy equilibrium method [6], operating point analysis [7].

In principle, a distinction can be made between two different assumptions regarding the motion profile: constant acceleration and constant speed.

2.1. Assumption of constant acceleration

The following approaches fall into this category: force equilibrium method [5], exergy equilibrium method [6], minimum cylinder [3]. The common idea behind those methods is that if we search for the smallest cylinder which fulfils the given application (without considering the kinetic energy at end stop) we must apply the maximum acceleration all the time. The result is that the speed increases and reaches its maximum at the end stop. The above-mentioned methods differ in their physical derivation but lead to very similar results and equations. In [5] and [6] constant friction forces are assumed for a cylinder whose diameter is not known yet. Therefore, iterations may be necessary to adjust the friction force in dependence on the chosen piston diameter. In [3] the constant friction is replaced by a friction model using Coulomb and viscous friction. The result is similar than in the other equations, but due to the viscous friction the velocity will reach a constant value if the stroke length is long enough.

2.2. Assumption of constant velocity

A constant velocity can also be directly specified in the known methods: force equilibrium method [5], exergy equilibrium method [6] and operating point analysis [7]. As shown in [5], the results are very similar but not equal.

Both the approaches with constant acceleration and constant velocity have the same drawback: the velocity at stroke end and therefore the kinetic energy is not taken into account. This means that the applicability of those methods seems to be restricted - but also suitable - only for short stroke cylinders with low payloads. For longer strokes (and therefore higher velocities) shock absorbers have to be used. Short stroke cylinders with low payload however will not need much energy, s.th. energy efficiency considerations are less important.

2.3. Assumption of a damped system behaviour at stroke end

In contradiction to the previous mentioned methods no direct assumption on the trajectory is made. Only the assumption of realising a velocity at end stop which yields in a kinetic energy value below the given limit. The only approach using this assumption is published in [8,3] and makes use of the eigenfrequency of the pneumatic system. In [5] some extensions of the eigenfrequency approach were made: an additional external spring force was integrated, a polytropic change of state was assumed, and both cylinder chambers were treated independently. In [9] a similar formula – the dimensionless mass – is derived by similarity considerations and is used for a special control strategy.

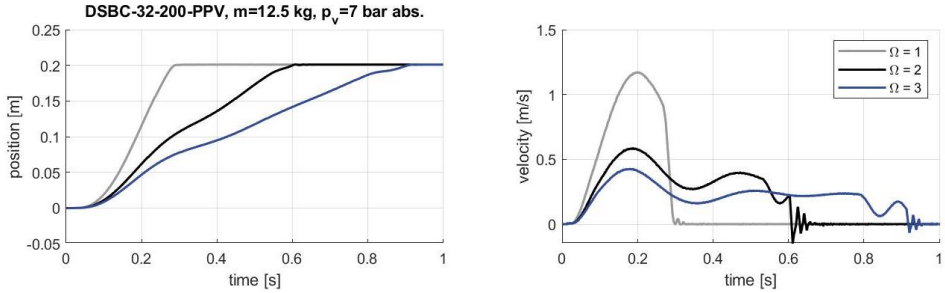


Figure 2: Measurements of the dynamic behaviour of a cylinder in dependence on the exhaust flow throttle and the resulting PFR values.

3. THE PNEUMATIC FREQUENCY RATIO (PFR)

The pneumatic frequency ratio (PFR) was published in [8,3] and expresses the ratio of the pneumatic eigenfrequency ω_0 and the application specific frequency $\omega_f = 2\pi/T_f$ which is given by the transition time T_f. Therefore, it expresses the ratio of the speed (expressed in terms of ω) the system would be able to do (ω_0) and the speed the system achieved within the application (ω_f):

$$\Omega = \frac{\omega_0}{\omega_f} = \frac{T_f}{\pi}\sqrt{\frac{A_k p}{mL}} \tag{2}$$

This formula depends on the effective area of the cylinder A_k, the load m, the transition time T_f, the supply pressure level p and on a characteristic length L. The characteristic length L includes the stroke length l_z and an additional length l_t resulting from the dead volume: $L = l_z + l_t$. The idea behind this formula is that the PFR characterizes the system behaviour, such that all systems with the same PFR behave similar. In [8] an optimal $\Omega = 1$ is determined by evaluating simulation data of optimal movements. Optimal movements mean that it moves as fast as possible but with the restriction of having a damped system behaviour at stroke end such that the kinetic energy at the end stop does not exceed the permissible value. For those movements a PFR of $\Omega = 1$ has been determined. **Figure 2** shows the dynamic response of a system in dependence on the PFR. Optimal behaviour is achieved with $\Omega = 1$ as no oscillations at stroke end are visible in the velocity signal.

3.1. Calculation of the required piston diameter

Equation (2) can now be used for several tasks: characterizing a system by using (2) directly, calculating an optimal diameter d of a cylinder for a given application (with $A_k = \pi/4 \cdot d^2$), calculating the necessary supply pressure level and to estimate the transition time T_f for a given system. Both most important inversions are listed below:

$$d = \frac{2\pi\Omega}{T_f}\sqrt{\frac{mL}{\pi p}} \tag{3}$$

$$T_f = \pi \cdot \Omega \sqrt{\frac{mL}{A_k p}} \tag{4}$$

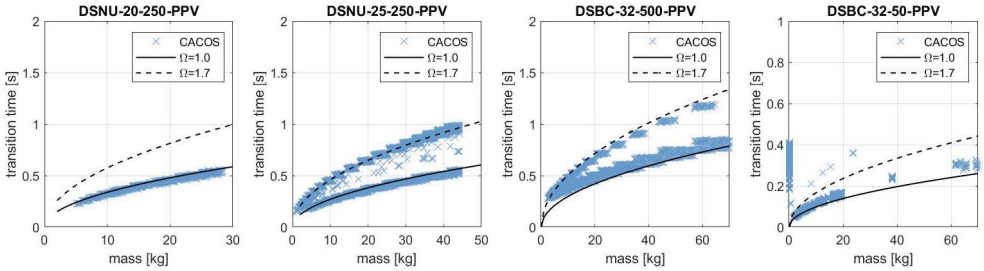


Figure 3: Approximation of the transition time using the PFR for several cylinders in horizontal orientation.

As in [8], once again we simulated a whole bunch of systems and applications by varying all the parameters (piston diameter $d \in \{20,25,32,50,80\}$ mm, stroke length $l_z \in \{50,100,200,250,500\}$ mm, pressure p, load $m \in [0,200]$ kg). Additionally, the exhaust flow throttle and the throttle of the pneumatic cushioning system (PPV) were varied in order to find a suitable combination. Useful combinations lead to a damped system behaviour which fulfil the condition for the kinetic energy at stroke end. Only those results are used afterwards, leading to optimal results (having no oscillations within the pneumatic cushioning system). The simulation study was made with CACOS (a simulation tool of Festo) and contains approximately 5 Mio. Simulations and 27.000 optimal results. **Figure 3** shows the comparison of those optimal simulations of several systems and their corresponding approximation of the transition time T_f using (4). The PFR reflects quite good the simulation results for all types of cylinders in horizontal orientation. In most the graphics there is a second optimality curve visible for a PFR of $\Omega \approx 1.7$ which can be approximated using (2) respectively (4), too.

Figure 4 shows again the results for one system (DSBC-32-200-PPV of Festo SE & Co. KG) in different orientations. Again, the PFR matches quite good for horizontal orientation, but for vertical orientation there is an obvious mismatch between the simulation result and the PFR. This is due to the gravitational forces which are not included in (2).

3.2. Calculation of the required sonic valve conductance

Once the piston diameter is specified, the sonic conductance of the valve can be calculated using a mean value of the required mass flow [8]:

$$C_v = K_C \frac{A_k L}{T_f} = K_C \cdot \pi^2 \Omega^2 \frac{m L^2}{T_f^3 p} \tag{5}$$

Thereby K_C is a constant which must be determined [8]. Concerning energy efficiency however, the valve conductance does not have a direct influence. It is important, that the sonic conductance is great enough to fulfil the required pressure dynamics of the application. Strictly speaking, (5) does not give the necessary conductance of the valve, but the necessary total conductance consisting of the valve, hose, and inlet port of the cylinder. The valve must be chosen, s.th. the overall conductance is met.

This dimensioning method is also part of the tool “Pneumatic Sizing” of Festo SE & Co. KG for an adjustable cushioning and is available online at: https://www.festo.com/x/pneumatic-sizing [1].

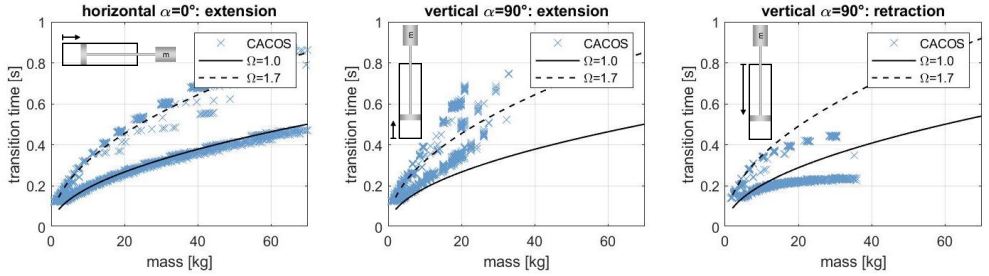


Figure 4: Approximation of the transition time using the PFR for horizontal and vertical orientation of a DSBC-32-200-PPV of Festo SE & Co. KG.

4. EXTENSION OF THE PFR FOR APPLICATIONS WITH EXTERNAL FORCES

As shown in [5] and [8], the application of PFR (2) is limited to horizontal load cases, as gravitational forces cannot be fully compensated by an adapted dimensioning of the cylinder using (2). This holds also for other external forces like forces resulting from the process e.g., for pressing or clamping and for increased friction forces due to external guidance.

Therefore, another simulative study was carried out for several systems by varying not only the load, and the throttles, but also an external force. The external force is a counter force, s.th. a positive value indicates a force against the direction of motion. **Figure 5** shows the results of all optimal solutions for a DSBC-32-200-PPV cylinder in dependence on the external forces. It shows, that for each curve of constant external force the transition times can be approximated by adjusting the PFR. As the external force increases, the transition time of an optimal movement decreases and therefore results in a higher value for Ω according to (2). This however would indicate that the system is oversized (in the sense of PFR), but due to the external force it is not the case: it is well sized because all the data points are optimal solutions. As for the PFR Ω, this should be reflected by the extended version, too. In other words: all optimal solutions should have the same value for Ω_{ext} - independent of the applied force F. The new equation for an extended PFR Ω_{ext} should also satisfy the following condition: $\Omega_{ext}(\mathrm{F} = 0) = \Omega$. The following ansatz function was used for finding an extended version of the PFR:

$$\Omega_{ext} = \Omega \cdot f(F) \tag{6}$$

The function $f(F) = \Omega_{ext}/\Omega$ is now identified using the measured values for Ω shown in **Figure 5**. For $\Omega_{ext} = 1$ it is the reciprocal of the measured Ω-values and is illustrated in **Figure 5** on the right in blue. We found out that its slope equals to: $-1/(A_k p)$, such that the searched function for extending the PFR can be stated as:

$$f(F) = 1 - K \cdot \frac{F}{A_k p} \tag{7}$$

Thereby we introduced an additional factor $K \approx 1$ for adjustments. This function equals to one when no force is applied and becomes zero for $K = 1$ and $F = A_k p$ which is the maximum force of the driving chamber. Using this function (7) and (6) the extended PFR can be stated as:

$$\Omega_{ext} = \Omega \cdot \left(1 - K \cdot \frac{F}{A_k p}\right) = \frac{T_f}{\pi} \sqrt{\frac{A_k p}{mL}} \cdot \left(1 - K \cdot \frac{F}{A_k p}\right) \tag{8}$$

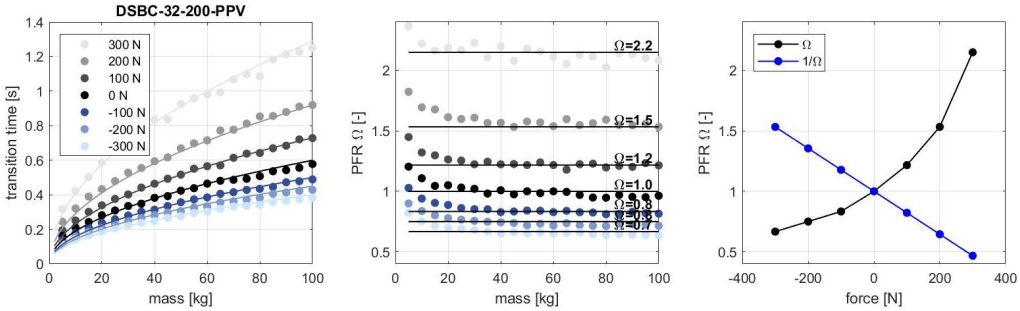


Figure 5: Simulation results (*dots*) of optimal movements in dependence of the load m and the external force F in comparison to the PFR approximation (*solid line*). *Left*: transition time, *middle*: PFR, *right*: PFR in dependence on the external force.

For $F = 0$, we get $\Omega_{ext} = \Omega$, for $F = A_k p$, we get $\Omega_{ext} = 0$. In this case, the external force equals the maximum force of the cylinder, such that it won't move. All in all, we found out, that the force F should not exceed 50% of the maximum force of the cylinder to guarantee a proper motion. As for the PFR, the extended version (8) can now be used for several purposes by inversion. Especially for dimensioning the piston diameter, the following inversion is of interest:

$$d_{ext} = \left(\Omega_{ext} + \sqrt{\Omega_{ext}^2 + 4\left(\frac{T_f}{\pi}\right)^2 \frac{KF}{mL}} \right) \Bigg/ \left(2T_f \sqrt{\frac{p}{4\pi \cdot m \cdot L}} \right). \tag{9}$$

For validating the simulation data, the transition time is also a useful inversion:

$$T_f = \pi \cdot \Omega_{ext} \sqrt{\frac{mL}{A_k p}} \cdot 1 \Big/ \left(1 - K\frac{F}{A_k p}\right). \tag{10}$$

Figure 6 shows again the optimal simulation data, but this time in comparison to the equations of the extended PFR (8) and (10). On the left we see that the transition times are approximated quite well by using (10). And the graph in the middle proves that all the measured transition times correspond to an extended PFR of approximately $\Omega_{ext} = 1$ independent on the external forces. There is a slight deviation for lower loads at high speed of the piston. This might be a result of viscous friction forces or of pressure losses which increase with the velocity. Additionally, **Figure 7** shows further validation data for a lower supply pressure and for other piston diameters and stroke lengths. The approximation (8) matches quite good also for these combinations.

4.1. Applications with gravitational forces

As shown in the previous section, the extended version of the PFR reflects quite good the influence of external forces. In this section we will use and test it with systems in vertical orientation. The force F needed in (8) therefore evaluates to:

$$\begin{aligned} F &= m \cdot g \cdot \sin\alpha\,, \qquad \text{for extension} \\ F &= -m \cdot g \cdot \sin\alpha\,, \qquad \text{for retraction} \end{aligned} \tag{11}$$

Thus, the applied force varies with the used mass and is also dependent on the direction of movement. **Figure 8** shows the same data set of **Figure 4** but now in comparison with the extended version of PFR. It shows that (10) now approximates the transition times quite well also in vertical orientation (α =+90°) for the extension stroke as well as for the retraction stroke.

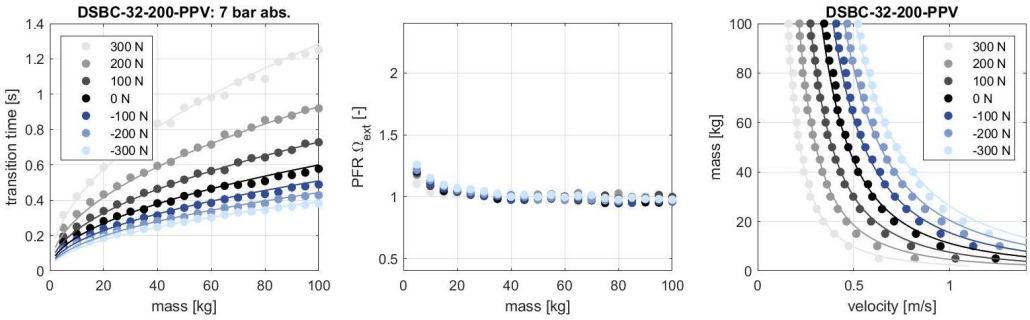


Figure 6: Simulation results (*dots*) of optimal movements in dependence of the load m and the external force F in comparison to the extended PFR approximation (*solid line*). *Left:* transition time, *middle:* extended PFR, *right:* mass-velocity diagram.

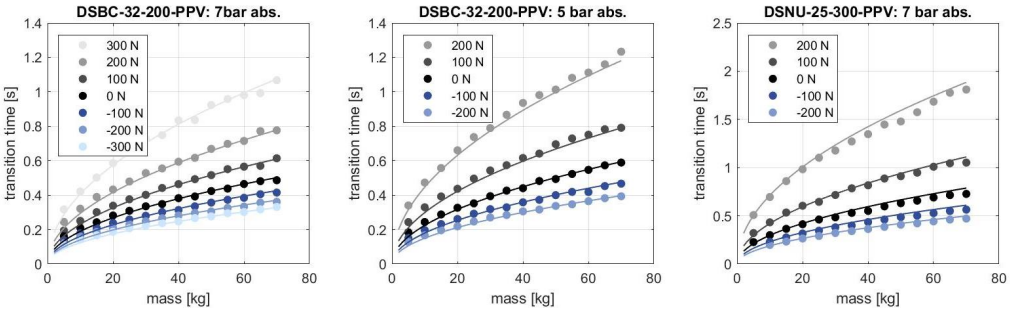


Figure 7: Validation data of the extended PFR for different diameters, lengths, and pressures.

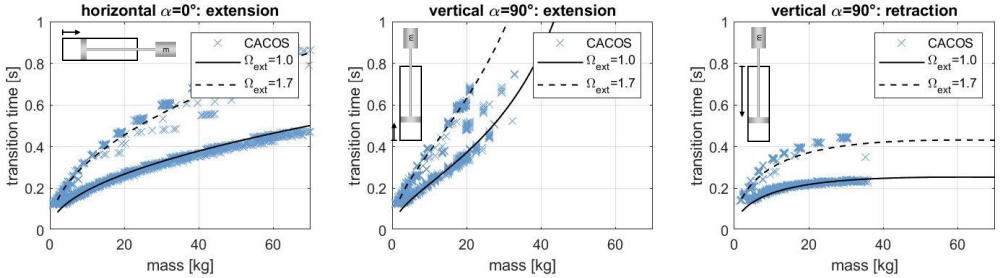


Figure 8: Approximation of the transition time using the extended PFR with external forces for horizontal and vertical orientation of a DSBC-32-200-PPV.

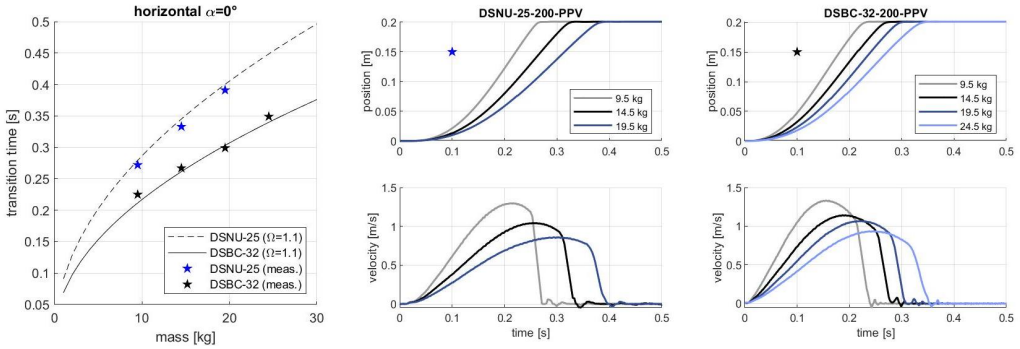


Figure 9: Measurements of two different cylinders with different loads and comparison of the resulting transition times with the PFR approach.

5. VALIDATION AND COMPARISON TO OTHER APPROACHES

In this section we validate the PFR and extended PFR formulas by measurements. Thereby all the measurements were done by adjusting the exhaust flow throttle as well as the throttle of the pneumatic cushioning system, s.th. an optimal behaviour is achieved. That means the fastest possible transition time with the condition of having nearly no oscillation at end stop.

5.1. Validation of PFR

Figure 9 shows the measurements for a horizontal installation position of two systems with several loads in comparison to the calculated transition time of the PFR approach using (4) (same as (10) with $F = 0$). It shows that for both systems the optimal transition times match quite good the calculation if we use $\Omega = 1.1$ instead of $\Omega = 1.0$. This is a learning from all the measurements done, that in the experiments the PFR is always a little greater than in simulation, and may be caused by idealized conditions within the simulation and mismatches concerning friction forces, pressure losses etc. The curves in principle however match quite good and can be taken for sizing the actuators just by adjusting the PFR.

5.2. Validation of extended PFR

Figure 10 shows the measurement results for a DSBC-32-200-PPV cylinder for the extension stroke in three different orientations ($\alpha = 0°, \alpha = 90°, \alpha = -90°$). Again, for the calculation of the transition times using as (10) and (11) an $\Omega_{ext} = 1.1$ was assumed. As could be seen, the measured transition times match quite good the predicted values of the PFR approach.

5.3. Comparison of the extended PFR to other approaches

In this section we compare the PFR-sizing approach with the constant acceleration and constant velocity approaches for exemplary applications. **Table 1** shows three applications and the corresponding piston diameters by using different approaches.

The **first application** is a typical application which is based on speed. For specifying the required piston diameters two different approaches were used. The first is the constant acceleration approach (acc.) which yields half the diameter as in comparison to the PFR approach. The results of both the systems are shown in **Figure 11** on the left. Thereby using the PFR sizing method yields a reduced velocity at stroke end, whereas a smaller cylinder leads to very high velocities which would indicate an infeasible solution or would lead to the necessity of using external shock absorbers.

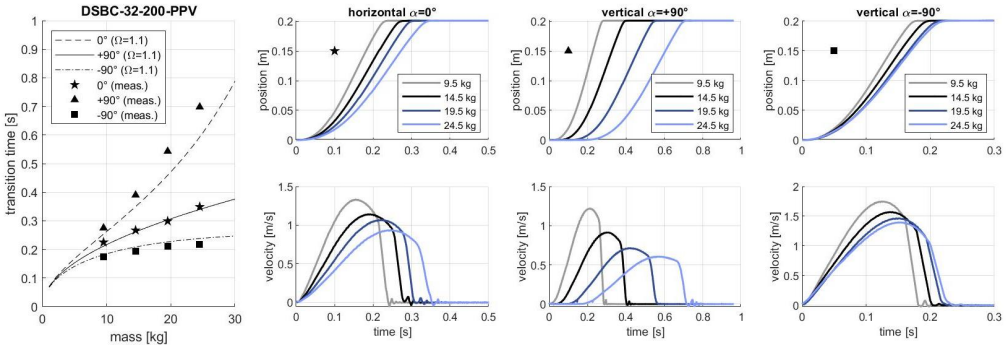


Figure 10: Measurements of a DSBC-32-200-PPV cylinder for the extension stroke in different orientations and comparison of the resulting transition times with the PFR approach for $\Omega_{ext} = 1.1$.

The **second application** is a vertical application with speed requirements. As in application 1, the result is similar, that with the const. acc. and const. vel. approach the cylinder is chosen too small such that a proper motion is not possible. However, when the requirements on the speed are relaxed – as in application 3 - the sizing result is similar.

This is shown in **application 3** which is the same as application 2 but with a very large transition time of $T_f = 7\ s$. As the required dynamics is very low the piston diameter then is mainly influenced by the gravitational and friction forces and thus lead to similar results between both approaches. As could be seen in **Figure 11** on the right, the cylinder needs several seconds before it starts to move. This is caused by the exhaust flow throttle which is adjusted to achieve the large transition time.

Table 1: Applications and calculated/**selected** piston diameters.

		Application 1		Application 2		Application 3	
load [kg]	m	12.5 kg		22.5 kg		22.5 kg	
stroke [mm]	l_z	200 mm		300 mm		300 mm	
transition time [s]	T_f	0.25 s		0.45		7 s	
orientation, dir. [°]	α	0° ext.		90° ext.		90° ext.	
		acc.	PFR	vel.	PFR	vel.	PFR
diameter [mm]	d	13.8 / **16**	30.6 / **32**	22.1 / **32**	39.3 / **40**	22.1 / **25**	21.0 / **25**
corresponding PFR	Ω_{ext}	0.49 / 0.57	1.1 / 1.15	0.15 / 0.73	1.1 / 1.13	2.3 / 5.3	1.1 / 5.3
impact velocity [m/s]	v_e	**1.7**	**0.09**	**0.65**	**0.09**	**0.08**	**0.08**
transition time [s]	T_f	0.26	0.24	0.52	0.42	6.8	6.8

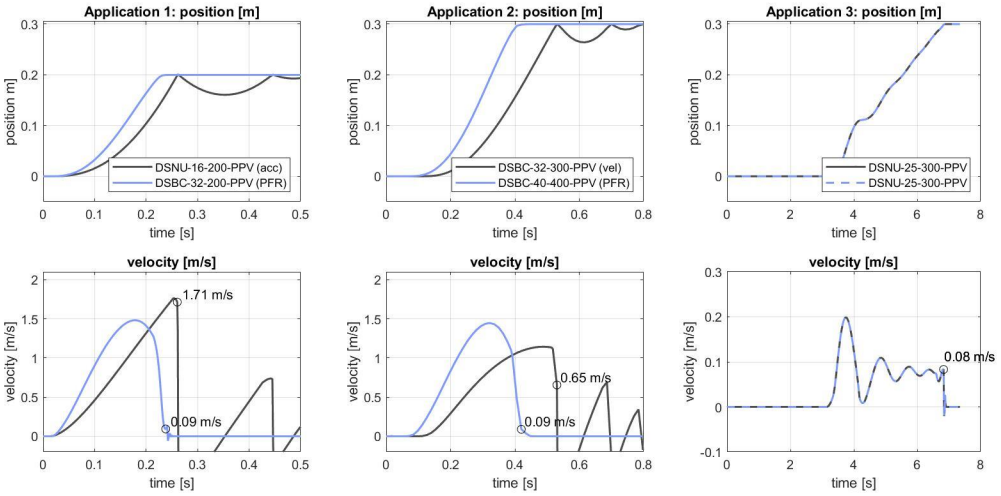


Figure 11: Comparison of different sizing approaches on basis of simulation results for three different applications.

6. CONCLUSIONS

The sizing of the system components is a simple measure to save energy. The choice of piston diameter is crucial here. However, many approaches for sizing only work for short drives or small moving masses because the end position speed is not taken into account or would otherwise lead to greatly oversized drives. In view of an energy-efficient design, however, medium, and large drives play a more important role. For these drives, however, the dynamic behaviour must also be considered, especially in the end position.

The PFR approach takes into account the dynamic properties of the drive and the end position behaviour and can be used for evaluation, design, determination of the transition time or also for calculating a reduced supply pressure for oversized drives. The biggest disadvantage, that the PFR is only valid for a horizontal installation position, was eliminated in this article by a corresponding extension of additional forces. Both approaches of PFR are simple algebraic equations that reflect the essential relationships, but do not represent exact solutions but lead to feasible systems.

In practical applications, the necessary value for Ω can be somewhat larger, as additional pressure drops, dead volumes, hose dynamics or frictional forces from external guides are present. Differences between the two directions of movement are also common due to the different surface ratios. Typical values here are in the range $\Omega \in [1.1, 1.6]$.

Ultimately, the Ω is a design factor that can be determined based on measurements and simulations, so that it can also be used for other cylinders and damping systems.

The presented approach will be used in a future version of the sizing tool of Festo. It has thus reached maturity, so that pneumatics is also equipped for the future in times of climate change.

NOMENCLATURE

A	area	m^2
d	piston diameter (bore size)	m

F	force	N
g	gravitational acceleration	m/s²
L	generalized length	m
l_z	cylinder length	m
l_t	dead length	m
m	mass	kg
p	pressure	Pa
T_f	transition time	s
α	angle of orientation	rad
Ω	pneumatic frequency ratio	-
ext	Index indicating the extension	-

CACOS	Computer Aided Cylinder Optimisation System, a simulation tool of Festo SE & Co. KG
DSBC	Standards-based profile cylinder (ISO 15552) of Festo SE & Co. KG Type code: "DSBC-diameter-length"
DSNU	Round cylinder (ISO 6432) of Festo SE & Co. KG Type code: "DSNU-diameter-length"
PFR	Pneumatic Frequency Ratio
PPV	adjustable pneumatic cushioning system of Festo SE & Co. KG

REFERENCES

[1] https://www.festo.com/x/pneumatic-sizing, Festo SE & Co. KG, 10.01.2024

[2] Boyko V, Nazarov F, Gauchel W, Neumann R, Doll M, Weber J (2024) Comprehensive Application-Based Analysis of Energy-Saving Measures in Pneumatics [Manuscript submitted for publication]. Int. J. of Fluid Power.

[3] Doll M (2016) Optimierungsbasierte Strategien zur Steigerung der Energieeffizienz pneumatischer Antriebe. Shaker Verlag, Aachen

[4] Doll M, Neumann R (2014) How big is the efficiency of pneumatic drives? An experiment provides clarity! 9th IFK, March 24-26

[5] Boyko V, Hülsmann S, Weber J (2021) Comparative analysis of actuator dimensioning methods in pneumatics. Proceedings of the ASME/BATH 2021, Symposium on Fluid Power and Motion Control, FPMC2021; October 19-21, Virtual, Online

[6] Rakova E, and Weber J (2016) Exonomy analysis for the selection of the most cost-effective pneumatic drive solution. Proc. of the 9th FPNI Ph.D. Symposium on Fluid Power. FPNI2016-1518. DOI 10.1115/FPNI2016-1518

[7] Vigolo V, De Negri V. J. (2021) Sizing optimization of pneumatic actuation systems through operating point analysis. J. Dyn. Syst. Meas. Control Vol. 143 No. 5. DOI 10.1115/1.4049170

[8] Doll M, Neumann R, Sawodny O (2015) Dimensioning of pneumatic cylinders for motion tasks, International Journal of Fluid Power 16:1, 11-24 DOI: 10.1080/14399776.2015.1012437

[9] Krytikov G, Strizhak M, Strizhak V (2017) The synthesis of structure and parameters of energy efficient pneumatic actuator, Eastern-European Journal of Enterprise Technologies, 1/7 (85), 38–44. DOI: 10.15587/1729-4061.2017.92833

FEASIBILITY STUDY AND EXPERIMENTAL VALIDATION OF A NOVEL COMBINED THROTTLING APPROACH

peer reviewed

Christian Reese*, Olivier Reinertz[1], Katharina Schmitz[1]

[1]*Institute for Fluid Power Drives and Systems (ifas),RWTH Aachen University, Campus-Boulevard 30, D-52074 Aachen, Germany*

* Corresponding author: Tel.: +49 241 80 47727; E-mail address: christian.reese@ifas.rwth-aachen.de

ABSTRACT

In the field of pneumatic automation, downstream throttled pneumatic drives are commonly used for motion tasks due to their cost-effectiveness, durability, and robustness. However, this type of system is often regarded as being inefficient. Consequently, many researchers have focused on developing more efficient control strategies. This paper presents the results of an experimental investigation of three distinct pneumatic circuits: a novel combined throttling approach, the control by a commercially available downstream throttle with quick exhaust function and conventional downstream throttling. A metric for comparing energy savings adjusted to changes in cycle time is introduced for objective evaluation. The findings highlight a significant reduction in normalized air consumption with the novel circuit compared to both state of the art control schemes.

Keywords: Pneumatics, Energy Savings, Sustainable Systems

1. INTRODUCTION

Currently, downstream throttling is the primary method used to control the speed and cycle time of pneumatic drives. In the typical circuit configuration, the drive chamber is pressurized with the full supply pressure using a 5/2 directional control valve. The opposing chamber employs downstream throttling to regulate the speed by building up a pressure that opposes the movement, referred to as 'back pressure' below. With this setup, the cylinder always consumes the maximum amount of compressed air due to its compressibility and pressurization of the actuator chamber to full supply pressure.

In upstream throttling, the driving cylinder chamber is pressurized only to the extent required for the load to move at the desired speed, while the opposing chamber is vented to the ambient atmosphere. Due to the inherent resistance of the venting paths, only minimal back pressure occurs. As a result, upstream throttling can represent a highly efficient implementation of pneumatic movement. Provided that a further supply of compressed air is prevented when the end position is reached it consumes less compressed air than downstream throttling due to the lower actuator pressure. This crucial feature can be achieved by using a 5/3-way valve with a closed neutral position.

A key difference between the two systems is the back pressure in the exhaust chamber at the end of the movement, which has a major influence on pneumatic end-cushioning. End-cushioning is essential to prevent damage resulting from high-speed piston impacts and plays a crucial role in the dimensioning of pneumatic actuators concerning their cycle time [1][2]. Inherently, upstream throttled drives exhibit low back pressure, which can have an adverse effect on end-cushioning function depending on the load condition. This often results in unacceptably longer cycle times when employing upstream throttling.

Numerous researchers have proposed control strategies with the aim of enhancing the efficiency of systems employing pneumatic drives. Systems designed to maximize the utilization of the expansion energy within compressed air have been examined in references such as [3], [4], [5], and [6]. One especially noteworthy example is the Festo motion terminal ([7][8]), which uses four proportional valves arranged in a Wheatstone bridge configuration. However, it's worth noting that these methods do add complexity to the system and increase the procurement costs and the workload during commissioning, which could pose challenges for broader implementation. Other simpler approaches to energy savings have also been proposed by some researchers, such as the intermediate storage of compressed air, as suggested in [9][10], or the reuse of compressed air via a rapid traverse and feed circuit [11]. The design of these approaches is highly application specific and therefore may have limited transferability. A promising approach for optimizing the efficiency of pneumatic drives has already been proposed by the authors [12], which involves combining downstream and adaptive upstream throttling. In the scope of this paper, a novel configuration of this concept is investigated.

2. INVESTIGATED SYSTEMS

The left side of **Figure 1** depicts the switching scheme originally presented at FPMC2020 [12]. This approach aims to combine the advantages of downstream and upstream throttled systems. Downstream throttles control the cylinder speed, while upstream proportional throttles adaptively regulate the air supply. If it is possible to achieve satisfactory pneumatic end-cushioning and, consequently, acceptable cycle times with a lower drive pressure compared to downstream throttling, there is significant potential for compressed air savings. The use of a 5/3-way valve with a closed neutral position is necessary in both circuits. Furthermore, the switching of the valve to the neutral position must be timed with the cylinder reaching its end position to prevent further unnecessary filling of the driving cylinder chamber after the motion has been completed.

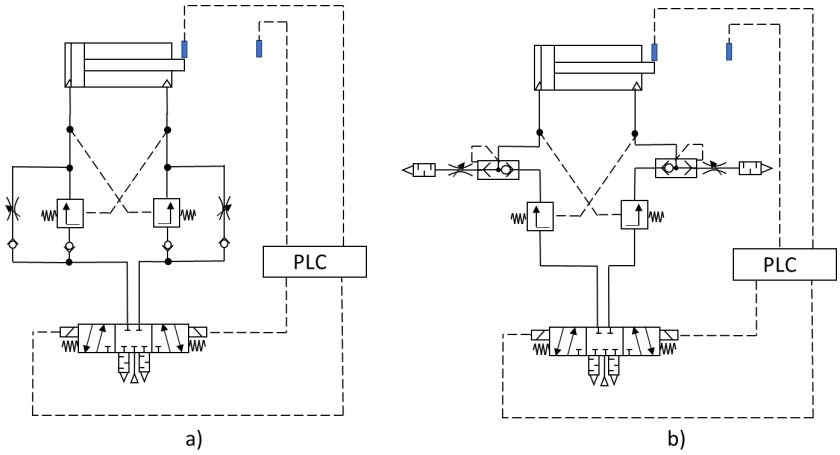


Figure 1: Adaptive upstream throttling configurations: a) original circuit and b) circuit with quick exhaust

The alternative design of the adaptive upstream throttling in Figure 1 (right) is based on the exhaust air flow being discharged into the environment by means of a quick exhaust valve. In contrast, in the original circuit diagram on the left side of **Figure 1**, the exhaust air flows through the directional control valve. The quick exhaust valve combines the logic structure of both check valves connected in parallel in the original switching scheme, resulting in a more compact design while simultaneously reducing pressure losses, especially when using longer pneumatic tubes.

The use of valve terminals is established in pneumatic systems, combining several pneumatic valves in one module. This leads to a reduction in installation and cabling effort. However, the use of this technology also increases the need for longer pneumatic tubes to reach the actuators. Therefore, this paper aims to also investigate the impact of tube sizes on the dynamics of pneumatic drives and air consumption. **Figure 2** illustrates the systems investigated in this paper. In addition to the conventional downstream throttled system a), the effects of quick exhaust on such systems is considered by means of the system b) to enable an objective comparison of all systems.

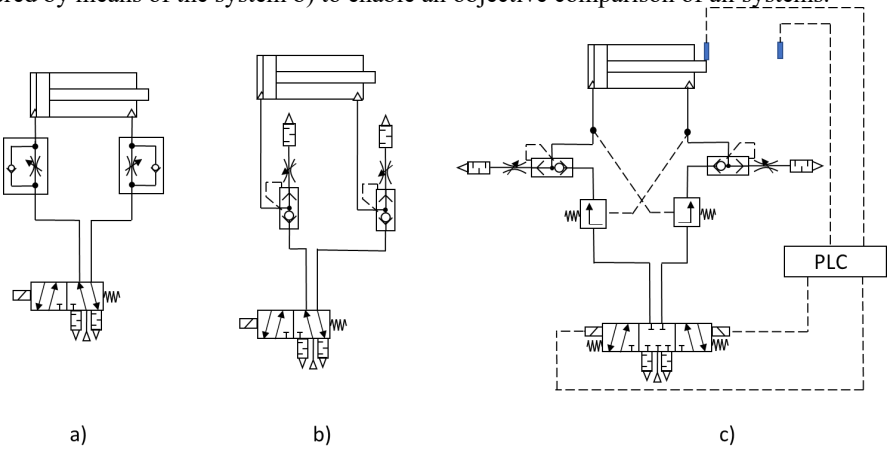


Figure 2: Investigated Systems: a) downstream throttling (DT), b) downstream throttling with quick exhaust (DTQ) and c) adaptive upstream throttling with quick exhaust (CTQ).

3. DESIGN OF THE ADAPTIVE UPSTREAM THROTTLING CONCEPT

Figure 3 shows the mechanical design of the concept with quick exhaust valves, whose circuit diagram was presented in Section 2. All the necessary components are combined in a common housing. The concept divides the circuit branches for each cylinder chamber into two levels, with the upper level shown in Figure 3 through a sectional view. To achieve a more compact design, upstream proportional controllers with a sleeve-spool design were chosen. All the required fluid pathways are integrated into the housing, including pressure signal holes for actuating the upstream controllers.

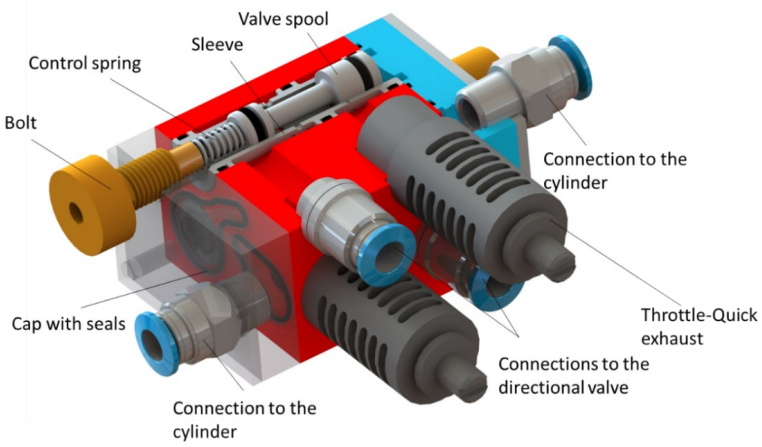


Figure 3: Mechanical design of the novel switching scheme

The manufactured prototype, as depicted in **Figure 4**, is used for the experimental tests conducted in this paper. The housing is designed from aluminium, while the cap is 3D printed using MSLA technology. The block's design is symmetrical, allowing for the use of the same cap shape on both sides. The model of the quick exhaust throttles is ASV 310 F-01-06 S from the SMC company, which were modified to be glued into the corresponding housing's bores, with the main functionality remaining unchanged. The same exhaust throttle model was later employed as a standalone component to test the impact of quick exhaust on downstream throttled drives (see **Section 4**).

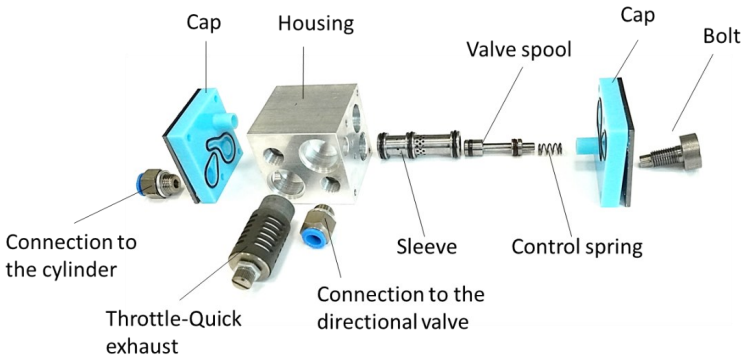


Figure 4: Exploded view of the manufactured components

3.1. Metering Edge Design

An essential aspect of controller design involves the careful selection of an appropriate metering edge geometry, since it has big implications on the metering characteristics of the valve and subsequently on the closed loop stability. **Figure 5** shows two possible implementations of the desired metering edge geometry.

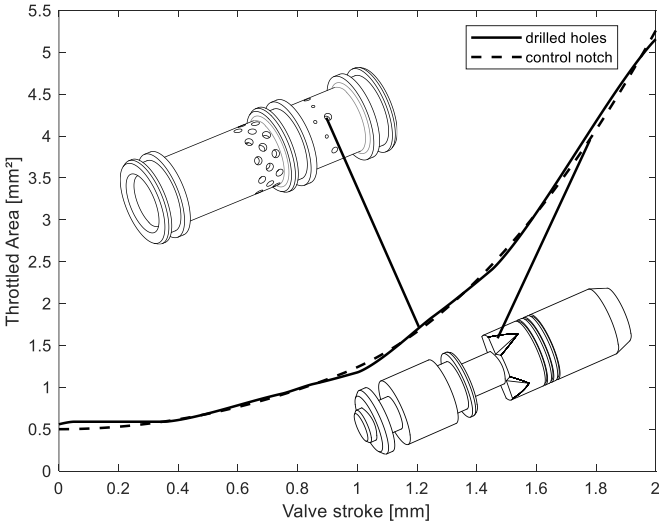


Figure 5: Metering edge design

To achieve satisfactory system performance, it is crucial to maintain closed-loop stability while ensuring sufficient conductivity to generate the necessary back pressure. The authors demonstrated that a linear opening characteristic is inadequate for reliable operation across a broad range of parameters [13]. Consequently, an alternative design approach incorporating a non-linear control

notch geometry was proposed. This design method utilizes a gain-scheduling approach, where controller gains are adjusted based on the steady-state error at different operating points, using a linear model of the system dynamics. A comprehensive explanation of this design method exceeds the scope of this paper; therefore, readers are referred to prior work [13] for more in-depth information. As mentioned earlier, a sleeve-spool design was selected to implement the metering edge of the controller due to its advantages in achieving a more compact design. Consequently, the opening geometry of the control notch design was approximated using drilled holes of varying sizes in the sleeve, as illustrated in Figure 5. The valve spool gradually exposes these holes to the airflow, thereby increasing the flow cross-section in the desired manner.

4. EXPERIMENTAL VALIDATION

For the experimental validation the test rig shown in **Figure 6** was used. The test rig allows for the air consumption and dynamic behaviour of the cylinder to be measured. Depending on the system, the 5/3 directional control valve is controlled as a 5/2 directional valve if the adoption of the neutral position is not required by the system. In this way, different systems with and without shut-off can be examined without change in the test rig components.

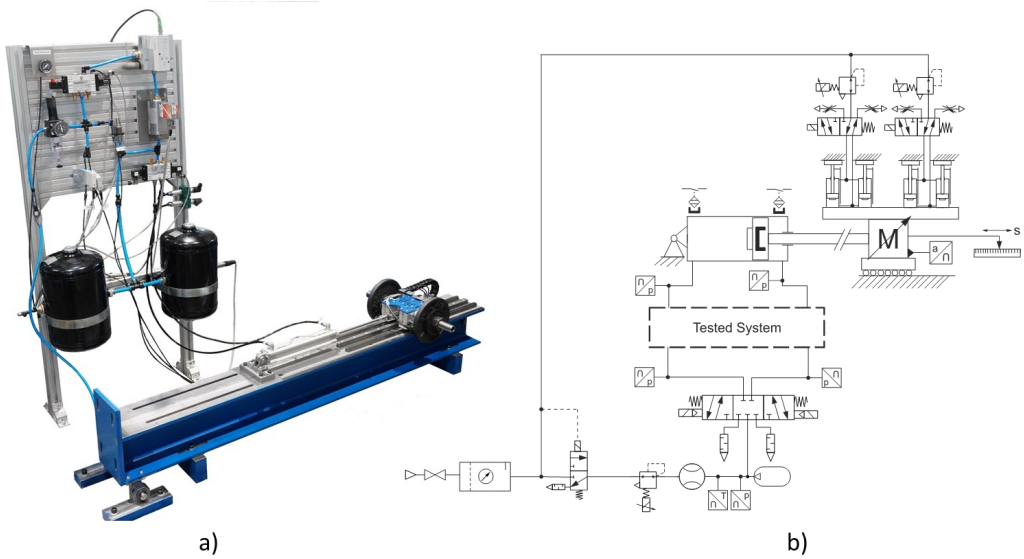


Figure 6: Test rig: a) implementation and b) circuit

4.1. Test plan

The investigated system parameters are listed in **Table 1** below. Three cylinder sizes were tested with varying tube lengths. The conventional downstream throttling (DT) was dimensioned using the sizing tool provided by a pneumatic components manufacturer [14]. It should be noted that with longer tube lengths, the tube diameter must be increased with such system due to the higher resistances in the return line. This can impact the proper functioning of the downstream throttle. In contrast, the systems with quick exhaust (DTQ and CTQ) do not experience an increase in downstream resistance and will therefore be tested with only one tube diameter. This approach allows for the evaluation of the effect of lower dead volume and higher upstream resistance on air consumption and cycle time. Furthermore, during the experimental investigation, all systems utilized the air cushion to its

maximum loading capacity. Consequently, the minimal achievable cycle time was measured for each system, thus allowing for an objective comparison.

Table 1: Test matrix with the corresponding tube diameters

		DT			DTQ/CTQ		
		Tube length [m]			Tube length [m]		
Cylinder size [mm]	Mass [kg]	1	4	6	1	4	6
Ø 25 x 250	5	Ø 6	Ø 6	Ø 6	Ø 6	Ø 6	Ø 6
Ø 32 x 200	10	Ø 6	Ø 8	Ø 8	Ø 6	Ø 6	Ø 6
Ø 32 x 320	10	Ø 6	Ø 8	Ø 8	Ø 6	Ø 6	Ø 6

4.2. Dynamic Behaviour

In this section, an exemplary comparison of the dynamic behaviour of all investigated systems using a cylinder with dimensions Ø 32 x 200 is provided. Quantitative results for all systems and cylinder sizes will be presented in Section 4.3. **Figure 7** shows the cylinder's speed and pressure curves within its chambers using the different systems listed in Table 1, with a used tube length of 1 meter.

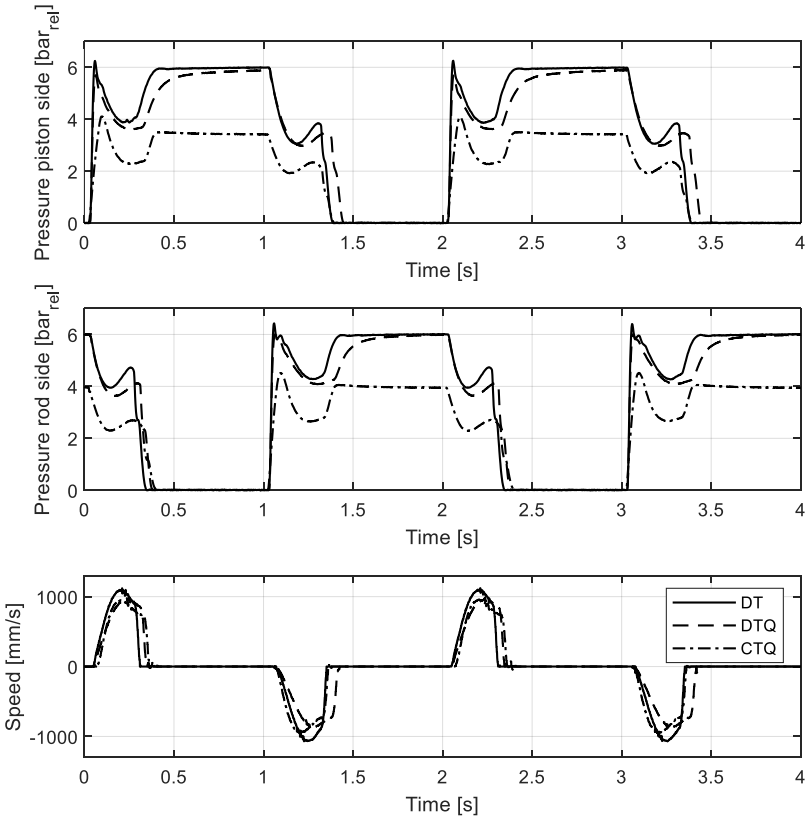


Figure 7: Dynamic behaviour: Ø 32 x 200, M = 10 kg and tube length = 1 m

The Combined Throttling Circuit (CTQ) exhibits lower back pressures and, consequently, lower driving pressure during cylinder movement compared to the other systems. Additionally, in systems

utilizing the directional valve as a 5/2 directional valve (DT and DTQ), the pressure in the active chamber rises to the supply pressure (6 bar_{rel}), whereas in CTQ, this is prevented by using the center position of the 5/3 directional valve as a shut-off. It is worth noting that the pressure rise at the end of the movement is much slower when using the quick exhaust throttle (DTQ) compared to the standard downstream throttling (DT), which indicates higher upstream pressure losses. Surprisingly, the cylinder experiences a lower start-up acceleration with the quick exhaust throttle (DTQ) than with the other systems. This may be attributed to higher upstream pressure losses resulting from the compact design of this commercially available valve, combined with the characteristic high backpressures associated with downstream throttling. On the other hand, CTQ exhibits a very rapid start of the movement especially on the return stroke, possibly due to the overall lower backpressure.

All the systems exhibit slower pressure build-up with a longer tube, as shown in **Figure 8**, due to the higher dead volumes. However, due to the larger tube diameter used in conventional downstream throttling (DT), an even higher driving pressure can be observed compared to Figure 7. Notably, the quicker start of the novel circuit (CTQ) on the return stroke remains observable despite the longer tubes.

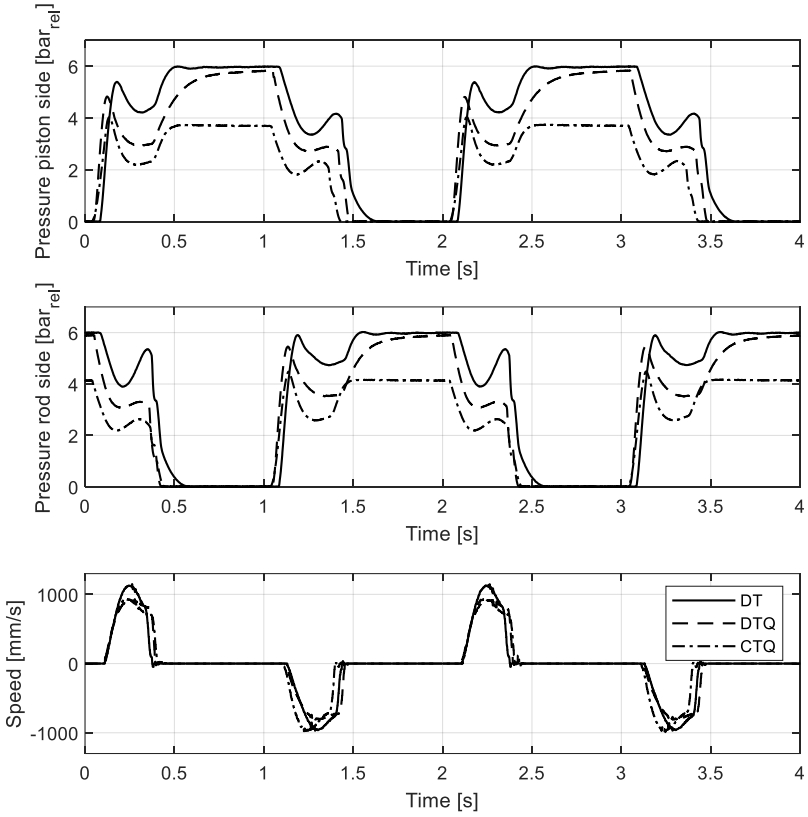


Figure 8: Dynamic behaviour: Ø 32 x 200, M = 10 kg and tube length = 6 m

4.3. Evaluation of the results

In this section, the experimental results are summarized. To facilitate a fair comparison between the systems, it is essential to account for the interaction between an increase in cycle time and potential compressed air savings. Therefore, in addition to absolute performance metrics such as cycle time

and air consumption, this paper will also present a normalized air consumption adjusted for changes in cycle time. For that, the sizing guideline for downstream throttled drives proposed by Doll et al. will be used [2], which is based on a "pneumatic frequency ratio" Ω. The value for Ω can be calculated using (1) and relates the reciprocal of the eigenfrequency (t_0) of the piston to the transfer time (t_f). A well dimensioned cylinder exhibits Ω values between 1,1 and 1,7.

$$\Omega = \frac{t_{f,DT}}{t_0} = \frac{t_{f,DT}}{\pi} \sqrt{\frac{A_p \cdot p_s}{M \cdot (L + 2 \cdot L_{dead})}} \tag{1}$$

Furthermore, the air consumption can be approximated with (2), with α being the cylinder area ratio.

$$m_{Air,DT} \approx \frac{A_p \cdot (L \cdot (1 + \alpha) + 2 \cdot L_{dead})}{R \cdot T_0} \cdot p_s \tag{2}$$

By substituting equation (1) into (2), an expression for the compressed air consumption as a function of the transfer time can be obtained (3).

$$m_{Air,DT} = \frac{M \cdot (L \cdot (1 + \alpha) + 2 \cdot L_{dead}) \cdot (L + 2 \cdot L_{dead})}{R \cdot T_0} \cdot \left(\frac{\Omega \cdot 2 \cdot \pi}{t_{f,DT}} \right)^2 \tag{3}$$

Assuming, that neither the operating temperature nor Ω change, the relationship between compressed air requirement and cycle time in equation (4) results. Here $m_{Air,DT}{}^*$ corresponds to the compressed air consumption at a different transfer time $t_{f,DT}{}^*$. Furthermore, we will use the complete cycle time instead of just the transfer time, under the assumption that the ratio of transfer time to dead time of the downstream throttled drive remains relatively constant at different pressure levels.

$$\frac{m_{Air,DT}{}^*}{m_{Air,DT}} = \left(\frac{t_{f,DT}}{t_{f,DT}{}^*} \right)^2 \approx \left(\frac{t_{cycle,DT}}{t_{cycle,DT}{}^*} \right)^2 \tag{4}$$

Finally, the normalized air consumption can be calculated as the ratio of the air consumption of a given system $m_{Air,i}$ to the corrected air consumption of a downstream throttled drive at the corresponding cycle time $m_{Air,DT}{}^*$.

$$Normalized\ m_{Air,i} = \frac{m_{Air,i}}{m_{Air,DT}{}^*} = \frac{m_{Air,i}}{m_{Air,DT}} \cdot \left(\frac{t_{cycle,DT}}{t_{cycle,i}} \right)^2 \tag{5}$$

Figure 9 to **11** display the results of the different cylinder sizes with the measured Ω for the downstream throttling. In Figure 9, all the systems used the same tube diameters, despite varying tube lengths as indicated in Table 1. As expected, both systems without a shut-off function (DT and DTQ) exhibit similar air consumption for all tube lengths. Interestingly, the normalized air consumption is higher for DTQ and CTQ with the shortest tube length, primarily due to their longer measured cycle time. However, this trend reverses for longer tube sizes, as the advantages of quick exhaust on cycle time become more pronounced.

The differences in the absolute air consumption between DT and DTQ in Figure 10 are solely caused by the variances in tube diameter. Thus, this difference can be understood as savings resulting from reduced dead volumes. The combined throttling circuit (CTQ) exhibits remarkable results in terms of normalized air consumption for this cylinder size, especially with longer tubes. Therefore, it can be concluded that the novel system is particularly advantageous with this specific cylinder size. This can be partially attributed to its faster dynamic behaviour at the beginning of the return stroke, as demonstrated in Figures 7 and 8.

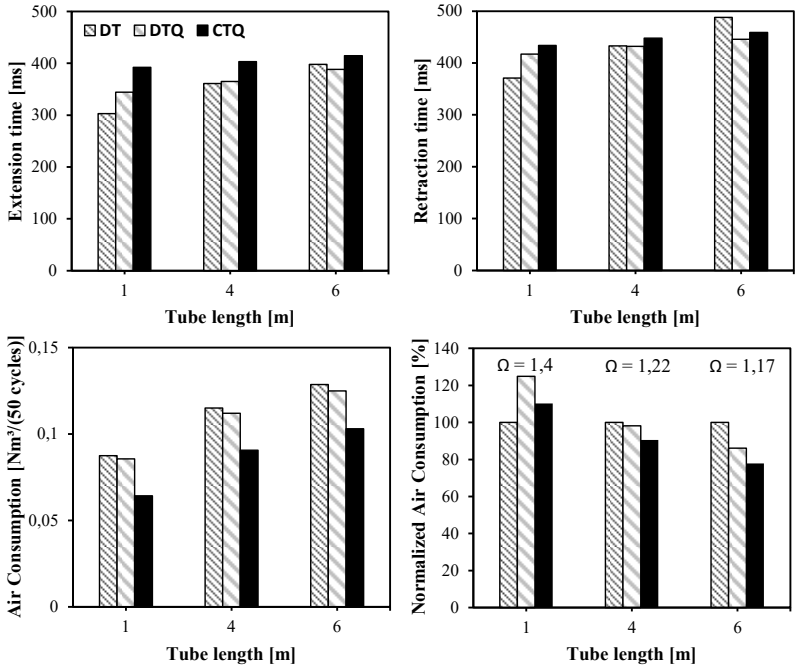


Figure 9: Measured results for Ø 25 x 250 with M = 5 kg

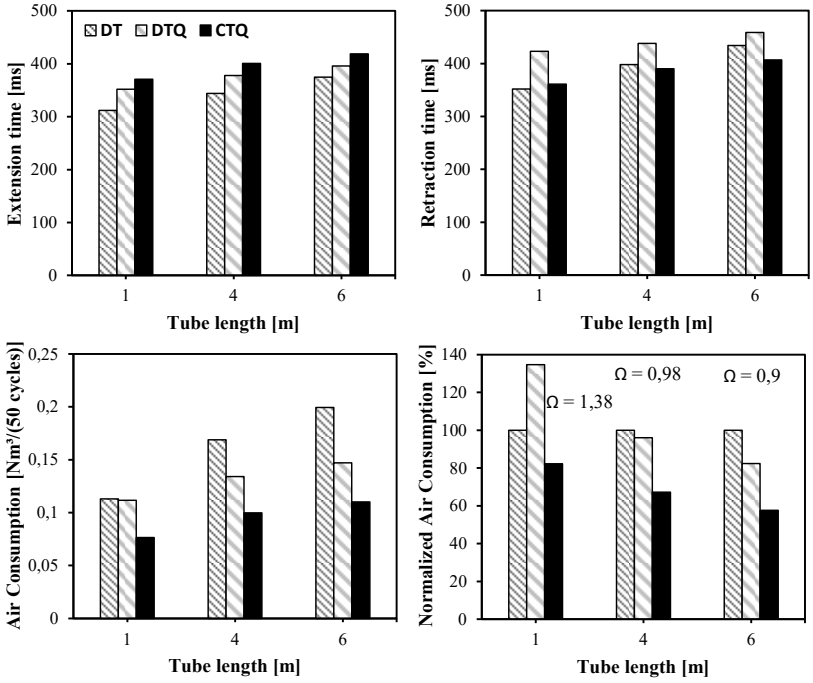


Figure 10: Measured results for Ø 32 x 200 with M = 10 kg

Figure 11 displays the results for the largest cylinder. Interestingly, despite the compressed air savings resulting from the smaller dead volumes used in DTQ compared to DT, DTQ performs worse than

DT in terms of normalized air consumption. One possible explanation is that the relative savings due to dead volume become lower with the larger cylinder and no longer outweigh the increase in cycle time. The normalized savings of the novel combined throttling approach (CTQ) remains largely unchanged with different tube sizes.

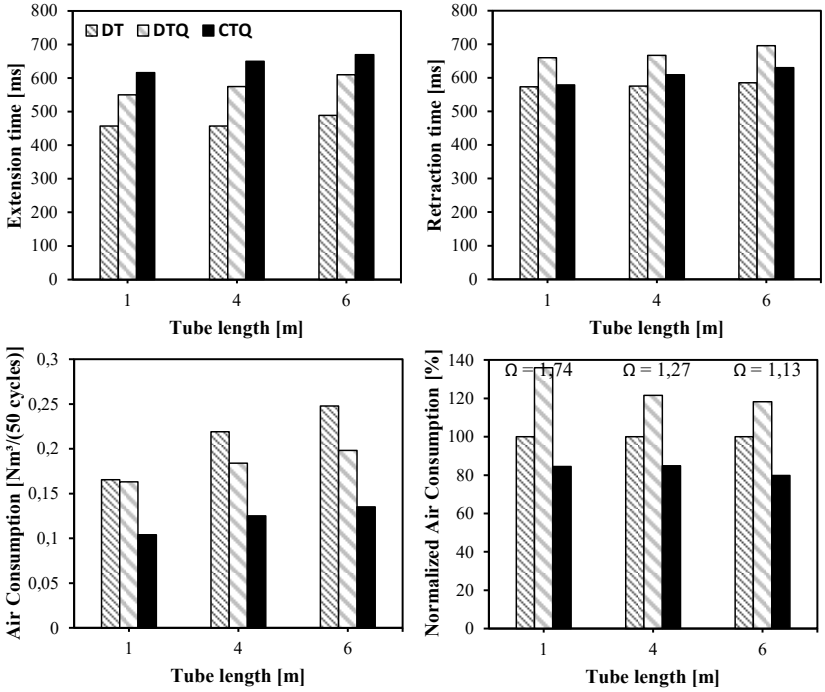


Figure 11: Measured results for Ø 32 x 320 with M = 10 kg

5. CONCLUSION AND OUTLOOK

The paper presents an experimental investigation of three different pneumatic circuits: a novel implementation of a combined throttling switching scheme, a state of the art downstream throttling with quick exhaust, and a conventional downstream throttled drive. Furthermore, a metric for comparing energy savings adjusted to changes in cycle time was introduced, allowing for an objective comparison between the systems. The results demonstrate a significant improvement in normalized air consumption with the novel circuit compared to conventional downstream throttling, although a high variation in performance was observed with different cylinder sizes. The results also indicate that the use of downstream throttling with quick exhaust can be beneficial in some applications, but its performance varies widely depending on the ratio of cylinder size to dead volumes in the lines, with the system achieving the highest normalized savings with smaller cylinders and longer lines.

In future work, the cost-effectiveness of the energy-saving circuits should be analysed. Additionally, the use of a shut-off function, implemented using a 5/3 valve in combination with the downstream throttle with quick exhaust, should be further investigated. Furthermore, exploring the performance of these systems with a broader range of cylinder sizes and moving masses could lead to a deeper understanding of the opportunities and limitations associated with the use of such circuits.

ACKNOWLEDGMENT

The IGF research project 21381 N / 1 of the research association Forschungskuratorium Maschinenbau e. V. – FKM, Lyoner Straße 18, 60528 Frankfurt am Main was supported from the budget of the Federal Ministry of Economic Affairs and Climate Action through the AiF within the scope of a program to support industrial community research and development (IGF) based on a decision of the German Bundestag. The authors would like to take this opportunity to express our sincere thanks for the funding. Furthermore, the authors would also like to thank the institute for "Digital Additive Production" (DAP) of RWTH Aachen University for their support in the production of the prototype.

NOMENCLATURE

A_p	Piston area	m^2
L	Cylinder stroke	m
L_{dead}	Dead volume expressed as cylinder stroke	m
M	Moving mass	kg
m_{air}	Air mass	Nm^3
p_s	Supply pressure	Pa
R	Specific ideal gas constant of air	J/(kgK)
t_{cycle}	Cycle time	s
t_f	Transfer time	s
T_0	Temperature at standard conditions	K
α	Cylinder area ratio	-
Ω	Pneumatic frequency ratio	-
PLC	Programmable Logic Controller	
DT	Downstream Throttling	
DTQ	Downstream Throttling with Quick-Exhaust	
CTQ	Combined Throttling with Quick-Exhaust	

REFERENCES

[1] Nazarov F, Weber, J (2022) Sensitivity Analysis and Robust Optimization of an Integrated Pneumatic End-Position Cushioning - 13th International Fluid Power Conference, June 13-15, 2022, Aachen, Germany

[2] Doll M, Neumann R, Sawodny O (2015) Dimensioning of pneumatic cylinders for motion tasks - International Journal of Fluid Power, 2015, Vol. 16, No. 1, 11–24, http://dx.doi.org/10.1080/14399776.2015.1012437

[3] Harris P, Nolan S, O´Donnell G (2014) Energy optimization of pneumatic actuator systems in manufacturing – Journal of Cleaner Production, 2014, Vol 72, 35-45, https://doi.org/10.1016/j.jclepro.2014.03.011

[4] Raisch A, Hülsmann S, Sawodny O (2018) Saving Energy by Predictive Supply Air Shutoff for Pneumatic Drives – Proceedings of the European Control Conference (ECC), June 12-15, 2018, Limassol, Cyprus

[5] Raisch A, Sawodny O (2019) Modeling and Analysis of Pneumatic Cushioning Systems Under Energy-Saving Measures – IEEE Transactions on Automation Science and Engineering, 2019, Vol 17, 1388-1398, doi: 10.1109/TASE.2019.2955806

[6] Doll M, Neumann R, Sawodny O (2011) Energy Efficient Use Of Compressed Air In Pneumatic

Drive Systems For Motion Tasks – Proceedings of the International Conference on Fluid Power and Mechatronics (FPM), August 17-20, 2011, Beijing, China

[7] Rager D, Doll M, Neumann R, Berner M (2018) New programmable valve terminal enables flexible and energy-efficient pneumatics for Industry 4.0 – 13th International Fluid Power Conference

[8] Neumann R, Rager D, Doll M (2016) Verfahren zum Betreiben einer Ventileinrichtung, Ventileinrichtung und Datenträger mit einem Computerprogramm – 2016, Patent DE 10 2016 206 821 A1

[9] Harimoto G, Senoo M, Fujiwara Y (2021) Fluid Circuit of Air Cylinder – 2021, Patent US 11,118,606 B2

[10] Hepke J, Weber J (2013) Energy saving measures on pneumatic drive systems – Proceedings of the 13th Scandinavian International Conference on Fluid Power (SICFP), June 3-5, 2013, Linköping, Sweden

[11] Heitmann M, Hergenröther T (2019) Fluidrückführvorrichtung für einen doppeltwirkenden Zylinder und Verfahren zum Betreiben eines solchen Zylinders – 2019, Patent DE 10 2019 121 433 A1

[12] Reinertz O, Schmitz K (2020) Optimized Pneumatic Drives through Combined Downstream and Adaptive Upstream Throttling – Symposium on Fluid Power and Motion Control (FPMC), September 9-11, 2020, Bath, United Kingdom

[13] Reese C, Reinertz O, Schmitz K (2022) Energy Savings through Adaptive Upstream Throttling and Supply Shut-Off on Downstream Throttled Drives – Chemical Engineering and technology, 2022, Vol 46, 137-146, https://doi.org/10.1002/ceat.202200381

[14] N. N, Pneumatic Sizing, Website, https://www.festo.com/eac/en_us/PneumaticSizing/

REINFORCEMENT LEARNING-BASED PID CONTROLLER DESIGN FOR MASS FLOW

Moritz Allmendinger[1]*, Michael Erhard[1]

[1]*Research and Development, Bürkert Fluid Control Systems, Christian-Bürkert-Straße 13-17, 74653 Ingelfingen*

* Corresponding author: Tel.: +49 7940 10 96 269; E-mail address: moritz.allmendinger@burkert.com

ABSTRACT

This paper demonstrates a model-based control synthesis strategy based on artificial neural networks and reinforcement learning. For this purpose, both the determination of the systems' response as well as the training of the neural network are transferred to a virtual environment. The neural network acting independently but interacting with a conventional PI controller, is optimized in order to achieve the predefined control target. The definition of the control target and the evaluation of the control response are compared with each other in the time domain enabling a flexible integration and adaptation to a wide range of possible requirements. The results illustrate the feasibility of control synthesis based on a virtual trained neural network. Considering variations and uncertainties for the control target and the environment, the neural network should become more robust and suitable for real systems with inherent deviations between each other.

Keywords: Reinforcement Learning, Simulink, Mass Flow Control, PID

1. INTRODUCTION

Fermentation is the conversion of organic substances with the help of bacteria, fungi or cell cultures in admixture with enzymes. The process is used in a variety of industries, from food and beverage to pharmaceutical and chemical, as well as in research laboratories or pilot plants. To achieve optimum fermentation results, precise control of all process parameters is essential. The type and concentration of nutrients, the temperature, the oxygen content, and the pH-value in the fermenter are crucial parameters for this. Up to 4 fermentation gases are used inside the fermenter to control the fermentation process: oxygen (O_2), nitrogen (N_2), carbon dioxide (CO_2) and air (21% oxygen and 79% nitrogen). Reproducible processes and repeatable product quality require precise control of these gases. For this reason, mass flow controllers (MFC's) are used to control the gas supply and exhaust.

Proportional-integral-derivative (PID) controllers are the most widely used control algorithms in the industry. The control strategy of MFC's for gaseous media used in fermentation processes is also based on the proven PID approach. Although classical PID controllers require only a small set of parameters to set up the control action, successfully tuning them can be a challenging task, especially in the presence of dominant and state-dependent nonlinearities. This results in a time-intensive optimization of the PID parameters to achieve an overall stable and robust control performance.

This paper examines whether neural networks and reinforcement learning (RL) can simplify standard PID design and enhance PID control performance. Against this background, the control action of a conservatively parameterized standard PID controller is extended by a further component that originates from a neural network. For this purpose, both the determination of the systems' dynamic response as well as the training of the neural network are moved to a virtual environment. This

approach needs a dynamic simulation model of the system modeling all significant physical effects of the MFC: nonlinear characteristic curves including actuator hysteresis and event-driven changes in system dynamics with respect to system state.

The paper is structured as follows: section 2 provides an overview about reinforcement learning, the used agent, and the training methods. The description of the simulated application is presented in section 3. Section 4 explains the training setup for the agent used in reinforcement learning. The results obtained are shown in section 5 and finally summarized in section 6.

2. REINFORCEMENT LEARNING – BASICS AND FRAMEWORK

2.1. Reinforcement Learning

Deep Learning requires labeled data to train a neural network. However, it is important to note that a neural network will never perform better on the test data than on the training data. This problem can be avoided with RL, which is based on trial and error.

The basic elements of RL are an agent and an environment. The agent observes a continuous-in-value state $s(t_k)$ of the environment at a discrete time t_k. Based on the recent policy $\pi(t_k)$, the agent performs a continuous action $a(t_k)$ out of space A of possible actions. This leads to a new state $s(t_{k+1})$ of the environment. The reward $r(t_{k+1})$ received for the new state informs the agent about the quality of its action so that it can adjust its policy accordingly. The agent's goal is to maximize the cumulative reward.

The breakthrough of RL came in 2013 when the Deep Q learning algorithm developed by Deepmind outperformed the best players of the Atari game [1]. Since then, the RL approach has made further progress. A variety of new techniques and algorithms still exist today. Research is also increasingly focusing on using RL in an industrial control context. Two relevant contributions are described below, with overlapping content in some areas.

As a nonholonomic system with measurement noise and external disturbances, controlling a robotic system is challenging. In their work [2], Gheisarnejad & Khooban demonstrate a possibility of how a RL agent (here the Deep Deterministic Policy Gradient (DDPG)) can assist a conventional controller in trajectory control. Therefore, the control action of the PID controller is manipulated by a superimposed action selected by the DDPG. During training sequences, this approach can compensate for uncertainties and disturbances in the trajectory control to ensure more robust control. The advantages of this approach are highlighted by a comparison between the RL-based PID controller and the previously used controller. Siraskar's article [3] investigates the use of the DDPG algorithm as a flow control algorithm with nonlinear valves. The trained DDPG is compared to a PID controller. The data indicates that the DDPG controller more accurately tracks the command signal compared to the PID controller. In contrast, the PID controller shows superior performance in suppressing disturbance variables.

A literature review suggests that combining PID control with RL-based control enhances the performance of this hybrid control system. This approach has the potential to improve the control performance for various problems, such as regulating gas mass flow. An RL-based PID controller could combine the positive results of the agent's control performance with the suppression of disturbances by the PID controller, thus also improving the results obtained by Siraskar. With reference to both publications and the DDPG used therein, this RL algorithm is also considered suitable for the problem dealt with here.

2.2. DDPG

Lillicrap et al. [4] developed the Deep Deterministic Policy Gradient algorithm for solving problems in environments with continuous-in-value states and action spaces. This previously mentioned algorithm combines the Deterministic Policy Gradient and Deep Q-Learning agents. The DDPG is based on the Deterministic Policy Gradient, as it can be used for continuous action spaces, and is extended by the techniques of the replay buffer and the updating of networks from Deep Q-Learning.

2.3. Training

There are two ways to train an RL algorithm. One way is to train it against a real environment. Therefore, the RL algorithm has to be executed on the intended real-time hardware and must be connected to the environment physically. The advantage of this approach is the existence of a realistic environment, ideally the environment into which the agent will later be integrated. The training with a real environment should cover all effects that may occur during the subsequent operation. In doing so, the agent is optimally trained for the later field of application. Computing power is often the main problem with this training method. Compared to a high-performance computer, embedded systems have significantly less memory and computing power on the target hardware. Another disadvantage is the possible damage of the hardware used during training. It is unavoidable that hardware used in the system is subjected to high loads due to the large number of training sessions or operating points within limit ranges, which lead to increased wear and tear.

For this reason, the model-based training approach is becoming increasingly popular. Compared to the previously described approach, the model-based training strategy relies on a simulated environment. This solves the problems that occurred during training with the real system – limited computing power together with increased wear and tear and sometimes even potentially dangerous conditions depending on the application. The performance of the computing hardware used for training can be easily adapted to the requirements of the training algorithm reducing overall training time. In addition, there is no hardware wear in a simulation model. Both effects jointly reduce the costs of training. In the model, the environment can always be parameterized in the same way making a reproducibly testing and comparison of the several trained networks easy. In this way, parameters and states in the model can be flexibly isolated and evaluated. This is only partially possible when training on the real system.

MATLAB and Simulink in Release R2023a are selected as simulation environment. The Reinforcement Learning Toolbox of MathWorks makes several implemented RL agents available, including the DDPG algorithm. The architecture of the model used for simulation of controller, agent and environment is shown in **Figure 1**. The plant model is described in section 3, RL setup, agent and controller parameterization are discussed in section 4.

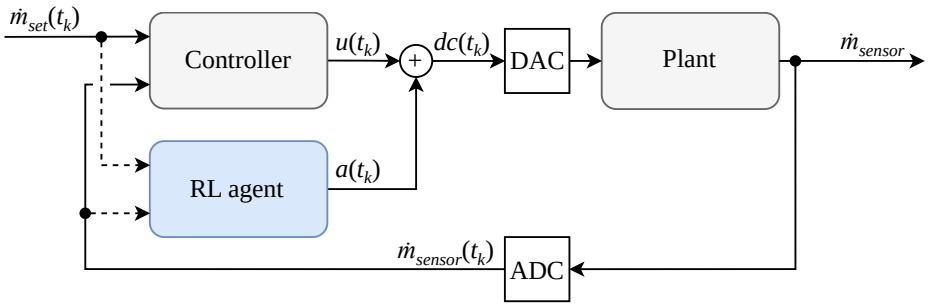


Figure 1: Structure and main components of training model

3. PLANT MODEL

This section deals with the plant model visualized in Figure 1 as counterpart to the controller and RL agent. For reasons of simplicity, the entire fermentation process is reduced to boundary conditions acting at the inlet and outlet of the MFC. Thus, the plant model only includes the physical effects inside the MFC, interacting with the application through (constant) boundary conditions. As shown in **Figure 2**, a MFC consists of 3 main components – an electronic control unit, a proportional valve, and a flow rate sensor. Because the controller is part of the control unit and modeled together with the RL agent in a discrete manner, the physical plant model reduces further. In the end, the modeling task will only include the valve and the sensor.

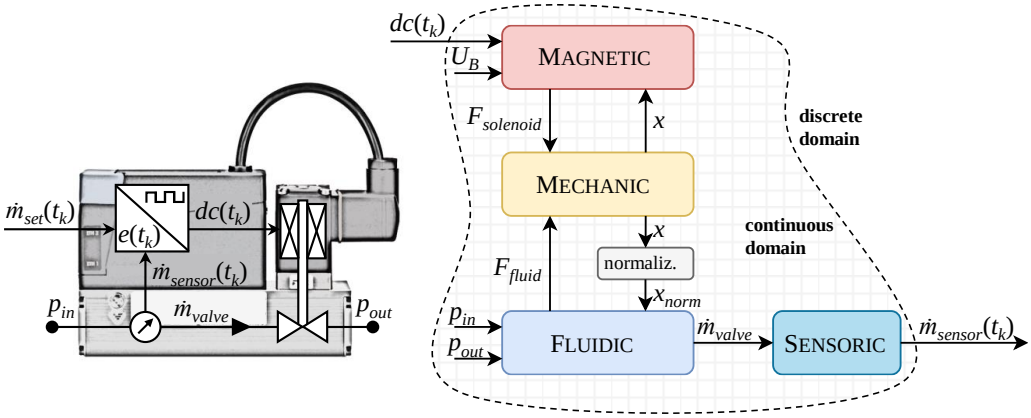


Figure 2: Mass flow controller layout (left), detailed signal flow interface chart for continuous plant model (right)

Taking a closer look at the time-continuous plant model, the governing equations from first principles must describe several physical domains and their interactions as indicated in Figure 2. Hereby, the modeling of the dynamic response of the proportional valve is the largest subtask, which will be discussed in more detail in the sections below.

The electro-magnetic actuator is made from soft-magnetic material excited by a current-driven coil. All relevant data for this solenoid is calculated with the help of finite element simulation yielding the nonlinear current inductance $\partial\psi/\partial i_{coil}|_{x,i}$ and the moving inductance $\partial\psi/\partial x|_{x,i}$ as well as the corresponding reluctance force map $F_{solenoid}$. The underlying equations (1) and (2) describe the current and force behavior for the electro-magnetic actuator.

$$\frac{\partial i_{coil}}{\partial t} = \frac{1}{\frac{\partial \psi}{\partial i_{coil}}} \cdot \left(dc \cdot U_B - R_{coil} \cdot \left(i_{coil} + i_{eddy} \right) - \frac{\partial \psi}{\partial x} \frac{\partial x}{\partial t} \right) \tag{1}$$

$$F_{\text{solenoid}} = f\left(i_{\text{coil}}, x \right) \tag{2}$$

Considering Figure 2 again, there is a nonlinear time delay between the applied duty cycle dc of the PWM excitation and the coil current or solenoid force, respectively. The supply voltage U_B for the PWM-controlled coil acts only as a constant boundary condition and is not changed during the different training episodes. An additional eddy current model fitted with transient finite element results completes the underlying dynamic simulation model of the actuator.

On the fluidic side, the modeling is limited to the description of the mass flow $\dot{m}_{valve}$ as a function of valve stroke x and input or output pressures p_{in}, p_{out} applied. It is known from ISO 6358 that the upstream pressure p_{in} together with the pressure ratio p_{out}/p_{in} is sufficient to calculate the mass flow through a restriction. The dependence from valve opening is implemented through the scaling of the conductance $C(x_{max})$ for the fully opened valve with the normalized valve stroke x_{norm}. Above the critical pressure ratio b, mass flow changes with outlet pressure as equation (3) indicates. Below the critical pressure ratio, the flow is saturated and mass flow depends only on upstream pressure, see equation (4). The missing parameters such as the conductance, the critical pressure ratio or the valve stroke normalization are determined from flow measurements at different valve openings.

$$\dot{m}_{valve} = x_{\text{norm}} \cdot p_{\text{in}} \cdot C(x_{\text{max}}) \cdot \rho_0 \cdot \sqrt{\frac{T_0}{T_{\text{in}}}} \sqrt{1 - \left(\frac{\frac{p_{\text{out}}}{p_{\text{in}}} - b}{1 - b} \right)^2} \tag{3}$$

$$\dot{m}_{valve} = x_{\text{norm}} \cdot p_{\text{in}} \cdot C(x_{\text{max}}) \cdot \rho_0 \cdot \sqrt{\frac{T_0}{T_{\text{in}}}} \tag{4}$$

In addition to these descriptions resulting from the state-of-the-art, the flow forces play an important role. On the one hand, an increase in mass flow results in an additional upstream pressure drop reducing the nominal lift force. On the other hand, the pressure profile interacting with the valve spool is highly velocity-dependent, resulting in an additional deviation form nominal lift force assumptions. Without significantly increasing the measurement effort, the same procedure as for the actuator is used, and the missing force characteristic map F_{fluid} is calculated via computational fluid dynamics.

$$F_{fluid} = f\left(\frac{p_{out}}{p_{in}}, x\right) \tag{5}$$

The translational movement of the valve spool connects the electro-magnetic with the fluidic domain. From balance of forces as shown in equation (6), the position x and velocity $\dot{x}$ are determined. These interface variables complete the calculations in the adjacent domains.

$$\frac{d^2x}{dt^2} = \frac{1}{m} \cdot \left(F_{\text{solenoid}} - F_{\text{fluid}} - F_{\text{friction}} - F_{\text{stop}} \right) \tag{6}$$

Due to the thermal measurement principle for the mass flow, a first order time delay is used as a simplified sensor model, see equation (7).

$$\frac{\partial \dot{m}_{sensor}}{\partial t} = \frac{1}{T} \cdot \left(\dot{m}_{valve} - \dot{m}_{sensor} \right) \tag{7}$$

With all model equations and parameters in hand, a complete model verification is performed for each domain as well as for the entire system. Without going too much into details, a final verification of the fluidic performance dependent on electrical excitation is highlighted in **Figure 3**. The measured mass flow is compared to the calculated mass flow. The two values are in good agreement.

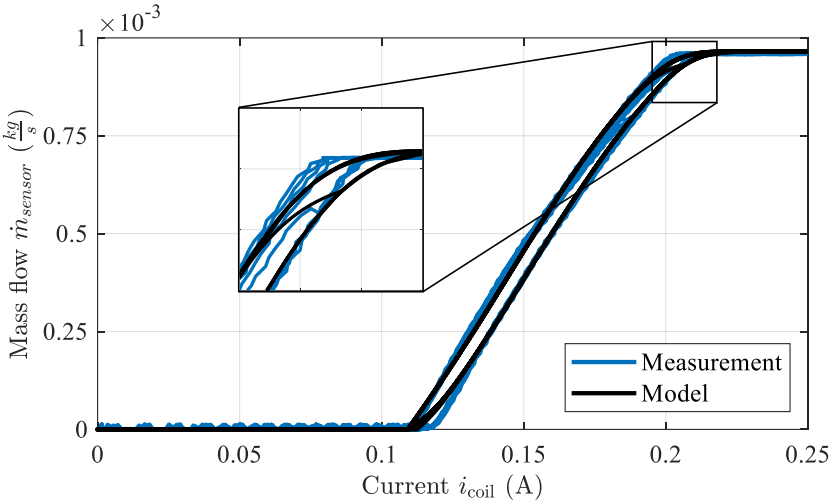


Figure 3: Verification of measured mass flow behavior $\dot{m}_{sensor} = f(i_{coil})$

The verification results suggest that the dynamic simulation model is suitable for the reinforcement learning approach as a counterpart to the real world.

4. REINFORCEMENT LEARNING AGENT AND CONTROLLER

This section describes the setup of the RL agent and the PID controller. As shown in Figure 1, the DDPG adjusts the control signal of the controller. Recalculation of the valve's duty cycle dc takes place at fixed points t_k in time according to equation (8).

$$dc(t_k) = a(t_k) + u(t_k) \tag{8}$$

Accordingly, the parameters of the PI controller remain constant and so does the closed-loop control behavior. The agent interacts with the control-loop at the same point as a feedforward control, but its action is generated in a different manner. In general, the control action of the agent accelerates or delays the opening of the valve depending on the current system status.

4.1. Observation Vector

The state observed by the agent consists of the following elements listed in equation (9). The control error $e(t_k)$ and the actual mass flow $\dot{m}_{sensor}(t_k)$ provide information for the control. The previous control error $e(t_{k-1})$ contains information about the dynamics of the system. Observing the action $a(t_k)$ informs the agent about its own behavior.

$$s(t_k) = \left[e(t_k), e(t_{k-1}), \dot{m}_{sensor,norm}(t_k), a(t_k)\right]^T \tag{9}$$

It is not necessary to include a history of the control error. The current duty cycle can also be omitted. Both approaches were examined in advance and showed no positive effects. Therefore, only the current action remains as part of the state vector. For training purposes, the values for the setpoint $\dot{m}_{set}(t_k)$ and actual value $\dot{m}_{sensor}(t_k)$ are scaled to a range between [0,1]. This ensures that all values of the observation vector are in a similar range of values.

4.2. DDPG Agent Networks

The used configuration of the actor and critic networks of the DDPG agent is shown in **Figure 4**. Both networks use fully connected (FC) layers as hidden layer and a ReLu layer as activation function. The concatenation layer inside the critic network combines the current state $s(t_k)$ with the action $a(t_k)$. The actor network uses a tanh-layer as the last layer to limit the value range of the action $a(t_k)$ between $[-1,1]$.

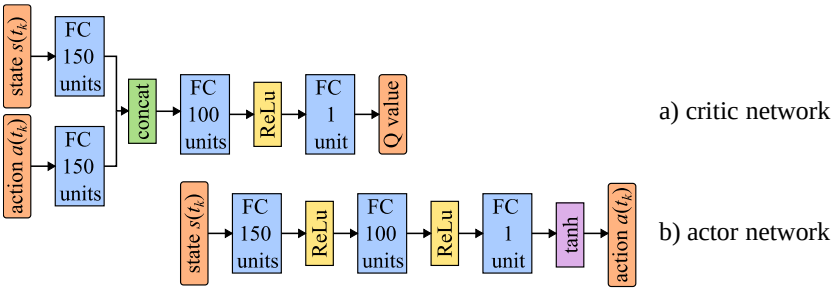


Figure 4: Network configuration of DDPG agent: a) critic network and b) actor network

The action $a(t_k) = 1$ opens the valve completely, irrespective of the PID control action. Conversely, this applies also to an action value $a(t_k) = 0$, the valve is then always closed.

4.3. DDPG Agent Hyperparameters

The hyperparameters of the agent and its training configuration are listed in **Table 1**. The DDPG uses the Ornstein-Uhlenbeck (OU) process as a noise model to solve the problem between exploitation and exploration.

Table 1: Hyperparameter of DDPG agent and training

Hyperparameter	Unit	Setting
Sample time	ms	10
Steps per episode	-	150
Discount factor	-	0.99
Replay buffer size	-	$1 \cdot 10^6$
Batch size	-	64
Optimizer	-	Adam
Smoothing factor	-	$1 \cdot 10^{-3}$
Learning rate - critic	-	$1 \cdot 10^{-3}$
Learning rate - actor	-	$1 \cdot 10^{-4}$
OU standard deviation	-	0.2
OU standard deviation decay rate	-	$5 \cdot 10^{-5}$

4.4. Reward

The reward function used is described in equation (10). The reward function consists of two parts: the first part rewards the control error, and the other part rewards the reduction in the change in the action signal. The first part aims to teach the agent to minimize the control deviation.

$$r(t_k) = -\sqrt{|e(t_k)|} - 0.5 \cdot \sqrt{|a(t_k) - a(t_{k-1})|} \tag{10}$$

First training results only including the control error showed that the agent's action tends to oscillate. For this reason, further investigations were necessary leading to the integration of the second part shown above. The agent should minimize the control error by making as few changes as possible to

the action signal $a(t_k)$. Weighting factors ensure that the agent's focus remains on minimizing the control error, which is why the second part is only weighted by half.

4.5. Episode Management

The scaled mass flow setpoint of each training episode is chosen randomly between $[0.05,1]$. For the dynamic simulation model of the plant, the initial condition of the valve opening is always fully closed meaning that there is no mass flow. Other environmental conditions such as inlet pressure and heating are not initially varied. The boundary conditions for the MFC model are fixed with an inlet pressure of 6 bar and with a solenoid temperature of 20°C. These fixed environmental conditions simplify the complexity of the training.

5. RESULTS

In this section, the performance of the trained DDPG-PI controller is evaluated using the simulation model of the MFC. **Figure 5** shows the control result of the DDPG-PI controller (solid lines) compared to the result of the conventional PI controller (dotted lines) for different setpoints (dashed lines). The PI controller is usually optimized for a specific setpoint using an autotune algorithm. In this case, the PI controller was optimized for the middle of the normalized operating range. Due to the nonlinear plant characteristics (overcoming the spring preload), settling times and overshoots in other operating points should differ from the tuned configuration.

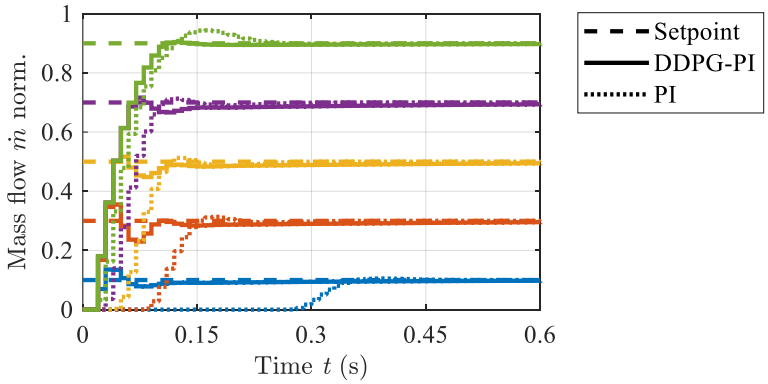
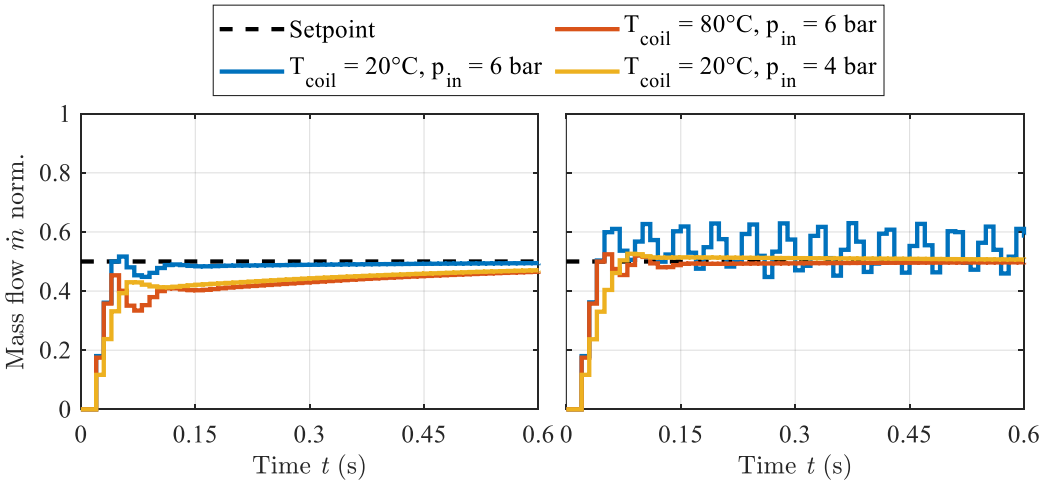


Figure 5: Comparison of the control performance between DDPG-PI and conventional controller

Compared to the conventional PI controller, the control performance of the DDPG-PI controller performance is significantly better in terms of settling times. This is not unexpected, as the agent's interaction with the control loop can be like the behavior of a static feedforward action. With the agent, all settling times are shorter regardless of the setpoint value. Particularly with low set values, the reduced settling time is linked to larger overshoots. However, these are acceptable as they do not damage the system. Because all mass flow transients start in the same way, it can be assumed that the agent has learned the nonlinearity of the plant and has thus implicitly linearized it.

Environmental conditions may vary during the use of MFCs. The capability of the DDPG to handle untrained system states is therefore being investigated. **Figure 6** a) shows the control behavior that results when the DDPG agent has to deal with untrained changes in the environmental conditions. In the fixed environment, the boundary conditions are 6 bar for the inlet pressure and 20°C for the coil temperature. This control result is displayed as blue line. The red line corresponds to the case, where the coil temperature inside the plant model has changed from 20°C to 80°C. The orange line

corresponds to the case in which the inlet pressure of the model has been reduced to 4 bar. With the help of the PI controller, the agent always reaches the setpoint regardless of the changes made. The integral component of the controller is mainly responsible for this because it minimizes the control deviation continuously. However, the settling time for reaching steady state has increased for both variations of the environmental conditions. To achieve the same mass flow at reduced inlet pressure, the valve must be opened further. A rise in coil temperature also delays the valve opening. The integral part requires more time to overcome this.



a) MiL inference with untrained environmental variations

b) MiL inference with trained environmental variations

Figure 6: Control result of DDPG agent with a) untrained and b) trained environmental variations

Different variations of the environmental conditions are part of the further training sequences to improve the control results. The inlet pressure and coil temperature are randomly selected from the following range of values at the start of each training episode: inlet pressure 3…6 bar, coil temperature 20…80°C. Figure 6 b) shows the control results for the same environmental variations as in Figure 6 a), with the difference that the environmental variations were now part of the training. The settling time of the mass flow, especially the last few percent before reaching steady state, has improved significantly. For all variations, the setpoint is reached within 100 ms. However, the robustness of the RL-PI controller has decreased. Although the RL-PI controller can robustly control the mass flow in the first inference at 20°C coil temperature and 6 bar inlet pressure, it is no longer capable in this inference. The mass flow oscillates around the setpoint after it has been reached.

6. CONCLUSION

As part of a feasibility study on RL agents for control engineering problems, this paper demonstrates the possibility of control synthesis based on a virtually trained neural network with RL for mass flow control. The DDPG agent is used as RL algorithm and extends the conventional PI controller by manipulating its control action. The training of the DDPG-PI controller is shifted to a simulation environment while retaining the advantage of high computing power and avoiding the disadvantage of system wear and tear. Therefore, a dynamic simulation model of a MFC is created using MATLAB, Simulink and the RL Toolbox. The DDPG-PI controller is validated using the simulated system. The findings reveal that a DDPG-PI controller improves the control result for different setpoints compared

to a conventional PI controller with fixed parameters. In particular, the settling time is reduced for all setpoint values. Of course, more advanced, e.g. model-based control approaches are also suitable for achieving such improvements.

Tolerable variations in the ambient conditions are also adequately handled by the RL-based controller due to the continuous action of the (fixed) PI controller. Extending the training sequences to include these variations improves the settling time, but also leads the DDPG-based controller losing stability. Further considerations are required to maintain robustness under variable environmental conditions, especially with regard to the signals to be included into the observation vector.

NOMENCLATURE

a	Action of agent	-
e	Control deviation normalized	-
$\dot{m}$	Mass flow	kg/s
p	Pressure (absolute)	Pa
r	Reward	-
s	State of environment	-
T	Temperature	°C
DDPG	Deep Deterministic Policy Gradient	
MFC	Mass flow controller	
PID	Proportional-integral-derivative	
RL	Reinforcement Learning	

REFERENCES

[1] Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M A (2013) Playing Atari with Deep Reinforcement Learning. Computing Research Repository (CoRR): abs/1312.5602

[2] Gheisarnejad M, Khooban M H (2021) An Intelligent Non-Integer PID Controller-Based Deep Reinforcement Learning: Implementation and Experimental Results. IEEE TRANSACTIONS IE 68(4):3609-3618

[3] Siraskar R (2021) Reinforcement Learning for Control of Valves. Machine Learning with Applications 121–134 68(4):3609-3618

[4] Lillicrap T P, Hunt J J, Pritzel A, Heess N, Erez T, Tassa Y, Silver D., Wierstra D (2019) Continuous control with deep reinforcement learning.

A TRAJECTORY-SPECIFIC APPROACH FOR CALCULATING THE REQUIRED HOLDING FORCE FOR SURFACE GRIPPERS

peer reviewed

Tobias Eberhardt[1,2,4,*], Valentin Stegmaier[1,3,4], Walter Schaaf[4], Alexander Verl[2]

[1] *Graduate School of Excellence advanced Manufacturing Engineering (GSaME), University of Stuttgart, Nobelstraße 12, 70569 Stuttgart, Germany*
[2] *Institute for Control Engineering of Machine Tools and Manufacturing Units, University of Stuttgart, Seidenstraße 36, 70174 Stuttgart, Germany*
[3] *Institute of Industrial Automation and Software Engineering, University of Stuttgart, Pfaffenwaldring 47, 70569 Stuttgart, Germany*
[4] *J. Schmalz GmbH, Pre-Development, Johannes-Schmalz-Str. 1, 72293 Glatten, Germany*

* Corresponding author: Tel.: +49 7443 2403-7516; E-mail address: tobias.eberhardt@gsame.uni-stuttgart.de

ABSTRACT

With increasing demands on the productivity and efficiency of manufacturing plants and rising energy costs, manufacturers of components and systems in industrial automation must also ensure that their products are used to their full potential in the intended application. This applies to automated handling with vacuum handling systems too. The energy consumption of such systems is highly dependent on the required holding force. In industrial automation, the required holding force is calculated from a few discrete positions of the surface gripper and the gripping object from the handling process. These positions are called load cases. However, the continuous transition between the kinematic positions of the individual load cases is not observed. The new approach presented in this paper allows an accurate calculation of the holding force required in the transition position when moving along a trajectory. Based on this new approach, the energy consumption of the vacuum handling system can be reduced. A validation with several experiments show good results.

Keywords: Holding force, Surface gripper, Energy efficiency, Vacuum handling system

1. INTRODUCTION

Handling systems are common and important subsystems of modern production plants. A handling system is used to load and unload workpieces to and from various stations in the production plant. They are also used to transport workpieces between processing stations. Vacuum handling systems are often used for such handling tasks [1]. One of their main advantages is the ability to grip objects from only one side using surface grippers in form of suction cups. This allows vacuum handling systems to be easily adapted to different shapes, weights and sizes of objects without the need of modifying the system. These features make vacuum handling systems an attractive solution for future flexible production systems [2].

In general, a vacuum handling system consists of components from six groups, which are vacuum generators, connections, fasteners, switching and system monitoring elements, valves, and suction cups [3]. The three most crucial components in the design of a vacuum handling system are vacuum generators, connections and suction cups [4]. For the purposes of the following considerations, the suction cup will be used as a representative of surface grippers.

The holding force that can be achieved by a vacuum handling system with respect to a gripping object depends largely on the effective suction area and the pressure difference achieved in the suction cup with respect to the environment. The effective suction area is mainly determined by the number of cups used, their size and type, and is therefore highly dependent on the geometric dimensions of the

gripping object. The pressure difference that can be achieved in a leakage-free system depends on the vacuum generator used, the evacuation time and the internal flow resistance of the system and is often referred to as the vacuum level. The vacuum level correlates directly with the energy consumption of the system, with particularly high vacuum levels being very energy intensive [5], [6]. **Figure 1** shows a typical structure of a simple vacuum handling system and an industrial application.

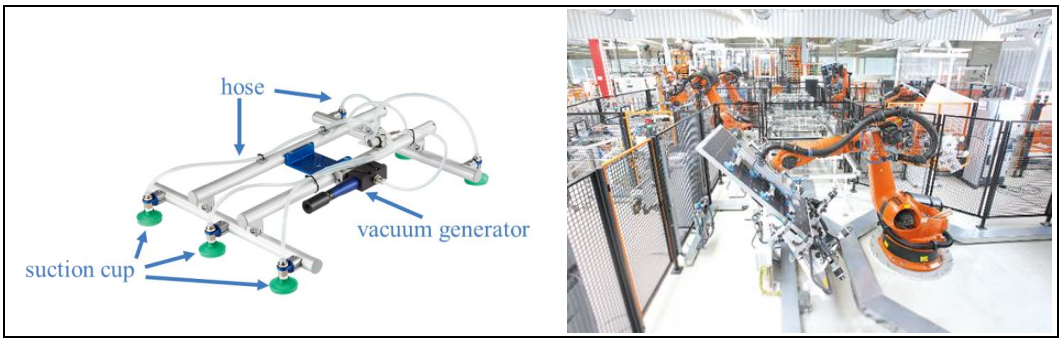


Figure 1: Typical structure of a vacuum handling system and a typical industrial application [7]

In practice, the holding force required for handling is usually calculated using discrete load cases. However, there are many applications in industry, such as rotating the gripping object with a robot, which cannot be adequately represented using these discrete load cases. See **Figure 2**. To compensate for this uncertainty, a high safety factor is added to the calculated holding force [2].
The new approach outlined below aims to change this. For this purpose, the state of the art is presented in section 2. In section 3, the mathematical formulation of the new approach is described. This is followed by experiments for validation in section 4. This section also compares the calculated and measured results. The paper ends with a short conclusion and an outlook.

2. STATE OF THE ART

After describing the calculation of the vacuum holding force of a vacuum handling system, existing approaches to calculate the theoretical holding force are presented.

2.1. Calculation of the suction force

According to Hesse [8], the holding force that can be applied by a vacuum suction cup depends on the effective suction area A, the pressure difference Δp, the safety factor S, the deformation coefficient n_3 to take into account the suction lip deformation due to the pressure drop and the efficiency η (leakage losses of the system). It can be calculated as follows:

$$F_{Suction} = \Delta p \cdot A \cdot \eta \cdot n_3 \cdot \frac{1}{S} = (p_0 - p_U) \cdot A \cdot \eta \cdot n_3 \cdot \frac{1}{S} \quad (1)$$

The holding force $F_{Suction}$ exerted by the suction cup must be greater than the theoretical holding force F_{Th} required by the process and the gripping object for safe handling.

$$F_{Suction} \geq F_{Th} \quad (2)$$

2.2. Existing approaches to calculate the required holding force

An existing approach is to define and mathematically describe some basic load cases in order to calculate the theoretically required holding force [9], [1], [2], [10]. For this purpose, the suction cup and the gripping object are considered either in horizontal alignment or in vertical alignment. **Figure 2** shows the described alignments of the suction cup and the gripping object. For the horizontal

alignment of the suction cup and the gripping object, a further distinction is often made between a vertical and a horizontal movement [9]. This subdivision is often referred to as load cases 1 to 3. Load case 1 is considered the most favourable and load case 3 the least favourable [9].

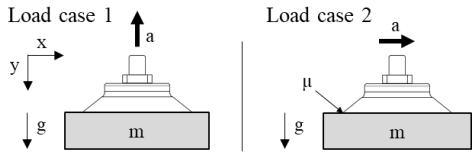
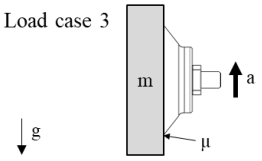
horizontal alignment	vertical alignment
	
$Load\ case\ 1\text{:}\quad F_{Th} = m \cdot (g + a)$ (3)	$Load\ case\ 3\text{:}\quad F_{Th} = m \cdot \frac{(g + a)}{\mu}$ (4)
$Load\ case\ 2\text{:}\quad F_{Th} = m \cdot \left(g + \frac{a}{\mu}\right)$ (5)	

Figure 2: Load case 1 to 3 for handling systems

Hesse [8], on the other hand, combines load cases 1 and 2 described in **Figure 2** by splitting a force acting obliquely on the centre of mass into an *x* and *y* component based on the angle of attack. The individual components are then calculated according to load cases 1 and 2. The results are added together to give the theoretical holding force required. In addition, Hesse extends the load cases 1 and 2 just described by the case of off-centre force application in the case of an eccentric centre of mass of the gripping object in the gripping plane.
Spivak's [11] approach considers an inclined suction cup. However, Spivak focuses on the holding force applied by the suction cup at a given vacuum and vertical movement.

2.3. Conclusions from the introduction and the initial situation

The presented approaches describe only discrete alignments (horizontal/vertical) of the suction cup - gripping object combination and map them to up to three load cases. A trajectory specific calculation is therefore not possible. Furthermore, they neglect the tipping moment due to a spatial expansion of the gripping object. In addition, the holding force applied by the suction cup at a certain vacuum is dealt with. However, not with the theoretically necessary holding force.

3. NEW APPROACHES FOR CALCULATING THE REQUIRED TRAJECTORY SPECIFIC HOLDING FORCE

Two approaches are presented for calculating the required holding force. The first approach neglects the spatial expansion of the gripping object. This extension is included in the second approach. We assume a point mass. Its distance cannot be neglected. The influence of air resistance on the gripping object during handling is neglected in both approaches. The occurring forces and accelerations are divided into *x* and *y* or *N* and *K* components according to the coordinate systems used and are marked with an appropriate subscript. The accelerations in the x and y directions and the angle of rotation α can be expressed as a function of time t to describe the trajectory of the combination of suction cups and gripping object. Accordingly, the theoretical holding force required F_{Th} is also time-dependent.

3.1. Approach 1: Theoretically necessary hooding force without tipping moment

In this approach, the spatial expansion of the gripping object is neglected and instead the gripping object is assumed to be a point mass. The centre of gravity of the gripping object is at the centre of the support surface of the suction cup. **Figure 3** shows the mechanical modelling assumed for the mathematical formulation.

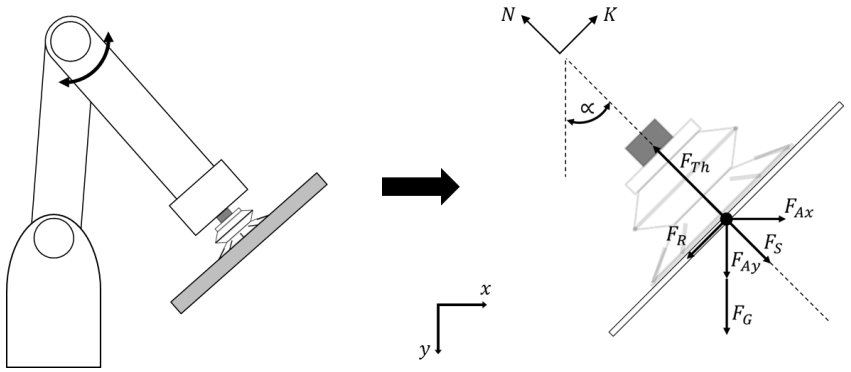


Figure 3: Mechanical modeling for thin gripping objects

To calculate the theoretically required holding force, static equilibrium with respect to the point mass is assumed. To do this, a second coordinate system is used with its principal axis in the direction of F_{Th}. The static equilibrium of inertia force F_A as a result of acceleration, weight force F_G, the suction restoring force F_S and the holding force F_{Th} along the N-axis then results as follows:

$$\sum F_i = 0 = F_{AxN} + F_{AyN} + F_{GN} + F_S - F_{Th} = 0 \tag{6}$$

The forces are transformed into the *N-K* coordinate system according to the angle of rotation α. The inertial forces can be transformed as follows:

$$F_{AxN} = F_{Ax} \cdot sin(\alpha) = m \cdot a_x \cdot sin(\alpha) \tag{7}$$

$$F_{AyN} = F_{Ay} \cdot cos(\alpha) = m \cdot a_y \cdot cos(\alpha) \tag{8}$$

$$F_{GN} = F_G \cdot cos(\alpha) = m \cdot g \cdot cos(\alpha) \tag{9}$$

The restoring force F_S of the suction cup is formulated as a function of the frictional force F_R. It depends on the magnitude of the frictional force and not on its direction. In the static case, it can be described by an equilibrium of the individual force components in the direction of the *K* axis. Converting the state of equilibrium to F_S gives the restoring force as follows:

$$F_S = \frac{|F_R|}{\mu} = \frac{\left|F_{Ax} \cdot cos(\alpha) - \left(F_{Ay} + F_G\right) \cdot sin(\alpha)\right|}{\mu} = m \cdot \frac{\left|a_x \cdot cos(\alpha) - \left(a_y + g\right) \cdot sin(\alpha)\right|}{\mu} \tag{10}$$

Substituting (7), (8), (9) and (10) into equation (6) gives the following formula for calculating the theoretical holding force F_{Th}:

$$F_{Th} = F_{Ax} \cdot sin(\alpha) + \left(F_{Ay} + F_G\right) \cdot cos(\alpha) + \frac{\left|F_{Ax} \cdot cos(\alpha) - \left(F_{Ay} + F_G\right) \cdot sin(\alpha)\right|}{\mu} \tag{11}$$

$$F_{Th} = m \cdot a_x \cdot sin(\alpha) + m \cdot \left(a_y + g\right) \cdot cos(\alpha) + m \cdot \frac{\left|a_x \cdot cos(\alpha) - \left(a_y + g\right) \cdot sin(\alpha)\right|}{\mu} \tag{12}$$

For an angle of rotation with α of 0°, the combination of load cases 1 and 2 results from the state of the art. For an angle of rotation of α = 90°, load case 3 results from the state of the art. **Figure 4** shows the theoretical holding force required along the trajectory for a smooth 90° tilt of the robot from the horizontal position.

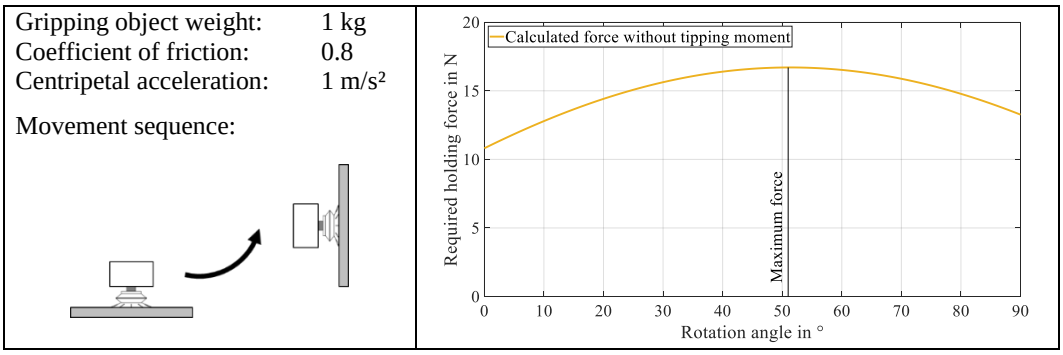


Figure 4: Theoretical required holding force when swivelling thin gripping objects

3.2. Approach 2: Theoretically necessary holding force with tipping moment

In the second approach, the gripping object has a spatial extension. However, the calculation still assumes that the centre of mass of the gripping object is on the same axis as the theoretically required holding force and that there is no relative displacement of the centre of mass during handling. **Figure 5** shows the corresponding modelling.

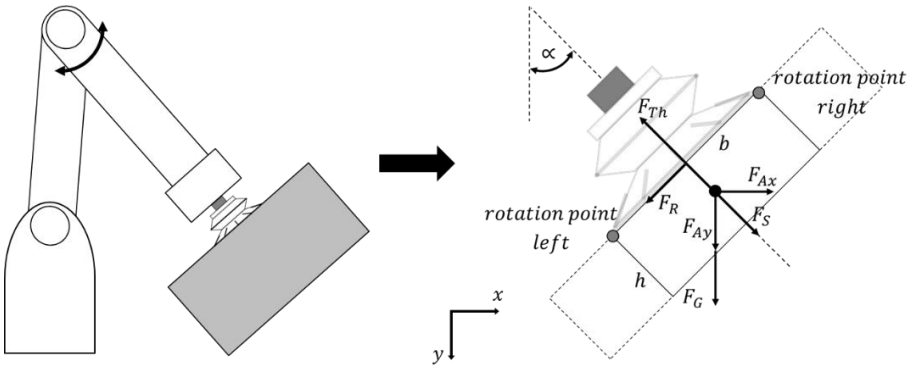


Figure 5: Mechanical modeling for gripping objects with spatial extension (h and b)

To calculate the required holding force, a static moment equilibrium is established around the left or right rotation point. The rotation point depends on the current angle of rotation α and the ratio of the accelerations in the *x* and *y* directions, as well as the geometric dimensions of the gripping object, and is located either at the top or bottom edge of the suction cup. For reasons of clarity, the basic equations for the forces are not used in the following. The following equation for the static moment equilibrium results for the left rotation point:

$$\sum M_i = 0 = M_{Th} \pm M_{Ax} - M_G - M_S - M_{Ay} = 0 \tag{13}$$

For the right pivot point, the moment equilibrium is as follows:

$$\sum M_i = 0 = M_{Th} - M_{Ax} - M_G - M_S \pm M_{Ay} = 0 \tag{14}$$

M_{Ax} at the left rotation point and M_{Ay} at the right rotation point have different effects on the moment equilibrium depending on the current rotation point. This must be taken into account in the geometric modelling of each moment. The modelling of the lever arms for both rotation points is shown in **Figure 6**.

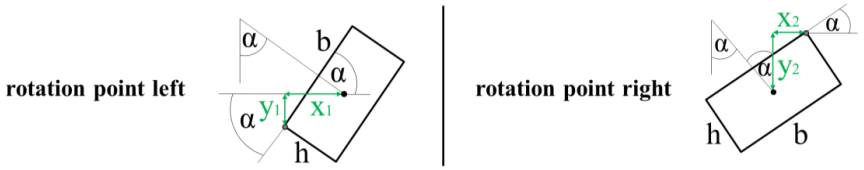


Figure 6: Left and right pivot points for moment equilibrium

According to **Figure 6**, the lever arms for the left rotation point can be described as follows:

$$x_1 = \frac{h}{2 \cdot sin(\alpha)} + cos(\alpha) \cdot \left(\frac{b}{2} - \frac{h}{2 \cdot tan(\alpha)}\right) \tag{15}$$

$$y_1 = \frac{1}{2} \cdot (b \cdot sin(\alpha) - h \cdot cos(\alpha)) \tag{16}$$

The lever arms of the forces with respect to the right rotation point, on the other hand, result from **Figure 6** as follows:

$$x_2 = \frac{cos(\alpha)}{2} \cdot (b - h \cdot tan(\alpha)) \tag{17}$$

$$y_2 = \frac{h}{2 \cdot cos(\alpha)} + \frac{sin(\alpha)}{2} \cdot (b - h \cdot tan(\alpha)) \tag{18}$$

Based on the geometric boundary conditions above, the following equations result for the moments of acceleration (M_{Ax} and M_{Ay}), gravity (M_G) and restoring (M_S) for the left rotation point:

$$M_{Ax} = F_{Ax} \cdot \frac{1}{2} \cdot (b \cdot sin(\alpha) - h \cdot cos(\alpha)) \tag{19}$$

$$M_{Ay} = F_{Ay} \cdot \left(\frac{h}{2 \cdot sin(\alpha)} + cos(\alpha) \cdot \left(\frac{b}{2} - \frac{h}{2 \cdot tan(\alpha)}\right)\right) \tag{20}$$

$$M_G = F_G \cdot \left(\frac{h}{2 \cdot sin(\alpha)} + cos(\alpha) \cdot \left(\frac{b}{2} - \frac{h}{2 \cdot tan(\alpha)}\right)\right) \tag{21}$$

$$M_S = F_S \cdot \frac{b}{2} = \frac{|F_R|}{\mu} \cdot \frac{b}{2} = \frac{\left|F_{Ax} \cdot cos(\alpha) - \left(F_{Ay} + F_G\right) \cdot sin(\alpha)\right|}{\mu} \cdot \frac{b}{2} \tag{22}$$

Merging the equations (19), (20), (21), and (22) into (13) gives the following equation for the theoretically required holding force F_{Th} in relation to the left rotation point:

$$F_{Th} = \frac{F_{Ax}}{b} \cdot (b \cdot sin(\alpha) - h \cdot cos(\alpha)) + \frac{\left(F_{Ay} + F_G\right)}{b} \cdot \left(\frac{h}{sin(\alpha)} + cos(\alpha) \cdot \left(b - \frac{h}{tan(\alpha)}\right)\right) + \frac{\left|F_{Ax} \cdot cos(\alpha) - \left(F_{Ay} + F_G\right) \cdot sin(\alpha)\right|}{\mu} \tag{23}$$

For the right pivot, the equations for the accelerations (M_{Ax} and M_{Ay}), gravity (M_G) and restoring (M_S) are as follows:

$$M_{Ax} = F_{Ax} \cdot \left(\frac{h}{2 \cdot cos(\alpha)} + \frac{sin(\alpha)}{2} \cdot (b - h \cdot tan(\alpha))\right) \tag{24}$$

$$M_{Ay} = F_{Ay} \cdot \left(\frac{cos(\alpha)}{2} \cdot (b - h \cdot tan(\alpha))\right) \tag{25}$$

$$M_G = F_G \cdot \left(\frac{cos(\alpha)}{2} \cdot (b - h \cdot tan(\alpha))\right) \tag{26}$$

$$M_S = F_S \cdot \frac{b}{2} = \frac{|F_R|}{\mu} \cdot \frac{b}{2} = \frac{\left|F_{Ax} \cdot cos(\alpha) - \left(F_{Ay} + F_G\right) \cdot sin(\alpha)\right|}{\mu} \cdot \frac{b}{2} \tag{27}$$

Merging the equations (24), (25), (26), and (27) into (14) gives the following equation for the theoretically required holding force F_{Th} in relation to the right rotation point:

$$F_{Th} = \frac{F_{Ax}}{b} \cdot \left(\frac{h}{cos(\alpha)} + sin(\alpha) \cdot (b - h \cdot tan(\alpha)) \right) + \frac{(F_{Ay} + F_G)}{b} \cdot (cos(\alpha) \cdot (b - h \cdot tan(\alpha))) + \frac{|F_{Ax} \cdot cos(\alpha) - (F_{Ay} + F_G) \cdot sin(\alpha)|}{\mu} \quad (28)$$

Figure 7 shows an extension of the calculation of the theoretically required holding force for the swivel operation from **Figure 4** by a spatial extension of the gripping object.

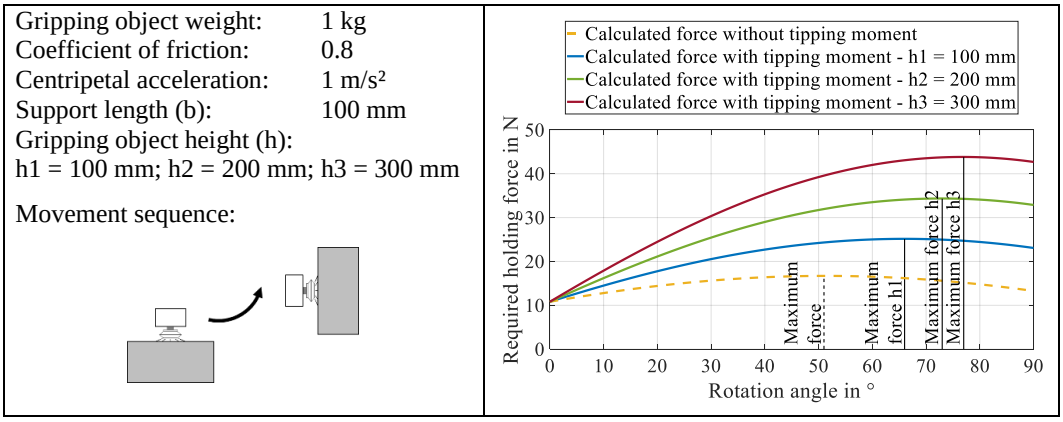


Figure 7: Theoretically required holding force when swivelling gripping objects with relevant thickness

3.3. Summary of the new approaches

The two approaches presented in sections 3.1 and 3.2 can be used to calculate the required holding force of the suction cup along any trajectory in the plane. **Figure 4** and **Figure 7** both show that the required force F_{Th} does not always have its maximum in one of the discrete load cases 1 to 3 according to the state of the art shown in **Figure 2**. Furthermore, the position of the maximum force is strongly dependent on the friction value and the ratio of the accelerations. Apart from that, the position of the centre of mass as a result of the spatial expansion of the gripping object also has a significant influence on the position of the maximum force as well as its magnitude. For this reason, designing using discrete load cases based on the worst-case scenario carries the risk of overestimating the maximum holding force. In addition to the swivel process itself, an intelligent choice of gripping object orientation can reduce the theoretical holding forces required for horizontal movement, for example. This allows a lower vacuum level and therefore lower energy absorption. **Figure 8** shows how, for an example application, a 6° adjustment results in a 17.7 % reduction in the holding force required.

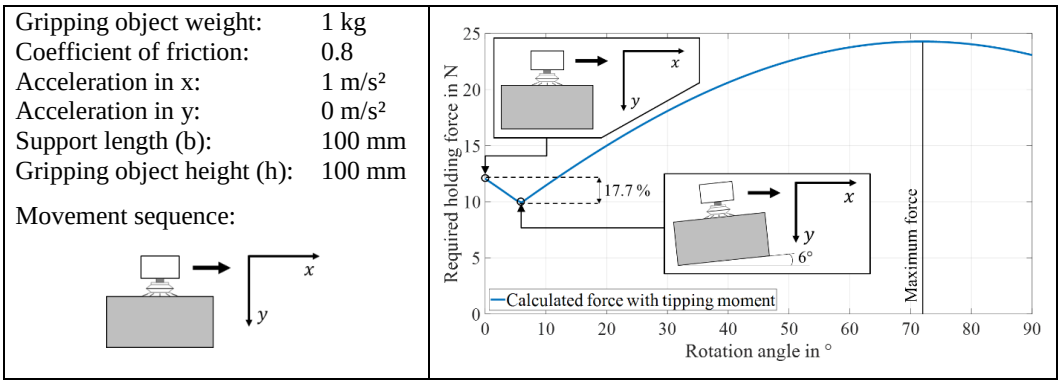
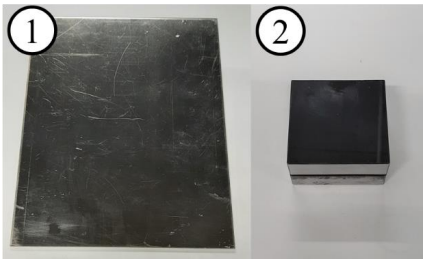


Figure 8: Reduce required holding force through clever alignment of the gripping object

4. VALIDATION

Several tests were carried out to validate the two new calculation methods described in section 3. For the calculation without overturning moment, a sheet metal of supposingly negligible thickness was chosen. On the other hand, for the calculation with tilting moment, a composite of two steel blocks was chosen. **Figure 9** shows the two parts and their main characteristics.



gripping object 1:	*sheet metal*
material:	*aluminium*
weight:	*0.880 kg*
dimensions:	*318 mm x 258 mm x 4 mm*
gripping object 2:	*steel block*
material:	*steel*
weight:	*0.887 kg*
dimensions:	*60 mm x 60 mm x 32 mm*

Figure 9: Test gripping objects used and their characteristics

For the suction cup, a seamless 3D printed sealing ring made of resin-acrylic-50 was chosen. This is inserted into a printed holder depending on the specific experimental setup. The combination of sealing ring and holder then forms the suction cup. The sealing rings and suction cups used in the respective test set-ups can be seen in **Figure 10**.

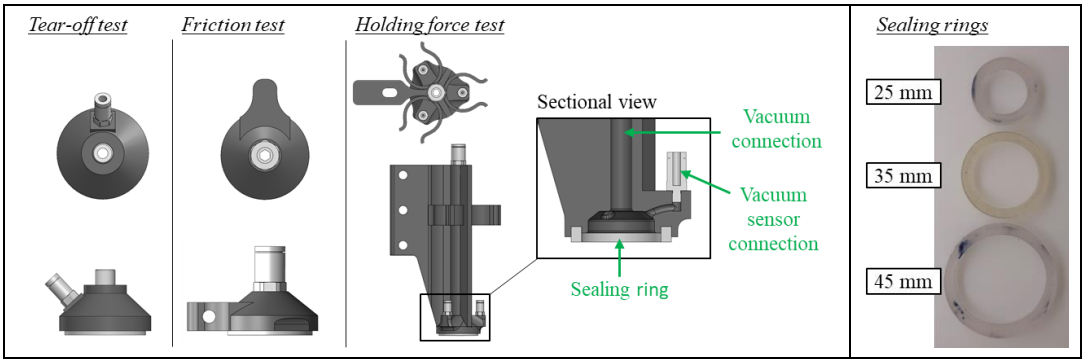


Figure 10: Suction cups and sealing rings used in the test set-ups

4.1. Determination of the suction cup diameter

In a first test, several tear-off tests were carried out to confirm the actual diameter d of the sealing ring. To do this, the seal was placed in another holder and pulled vertically off the test piece using a tensile testing machine. This was repeated at least three times for different vacuum levels. The vacuum was generated by an ejector (= vacuum generator) and manually adjusted to the appropriate level using a ball valve and a vacuum sensor. For evaluation, the maximum force in the force-displacement curve was determined representing the last point before, the cup lifts and is vented. At this point the handling system fails. The test set-up described and the test results obtained are shown in **Figure 11**. A 35 mm sealing ring was used for the tests.

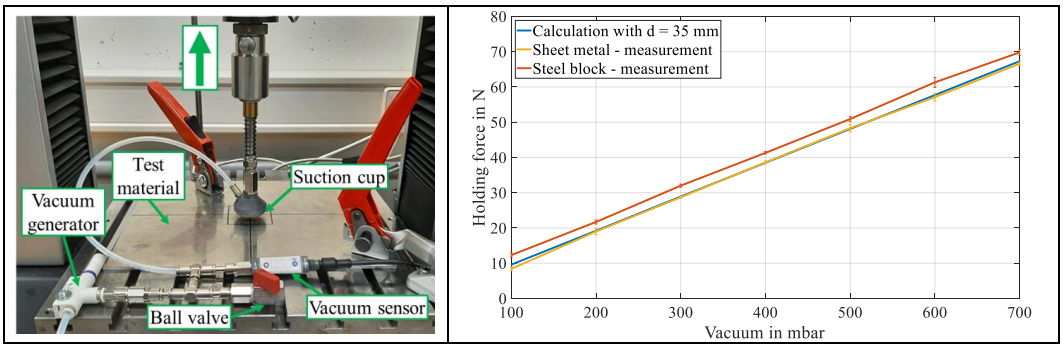


Figure 11: Experimental setup and results of the tear-off test

The test results show a good match between the calculated and measured values for the normal force to be applied. The effective diameter of the gasket on the sheet metal is very close to the nominal diameter. However, there is a slight deviation in the steel block test.

4.2. Determination of friction coefficients

In a second test, the friction value between the two test objects and the sealing ring was measured at different vacuum levels to compare the tests with the calculation. To do this, the sealing ring was placed in a holder and pressed against the test piece with a defined vacuum. The required vacuum was created using an ejector and adjusted to the desired level using a vacuum sensor and ball valve. The suction cup was then pulled upwards using a tensile testing machine. The force application point of the tensile machine on the suction cup is as close as possible to the surface of the test piece so that the lowest possible torque occurs during removal. The highest force value is then used to evaluate the force-displacement characteristic. This corresponds to static friction. Once the static friction limit has been exceeded, the suction cup will slip, which would correspond to a failure of the system in a real application. The pull-off test was performed at least three times for each vacuum level tested.
Based on the results of section 4.1, the applied normal force F_{Nf} can be determined and then, using the static friction force F_{Sf} recorded in the test, the corresponding static friction coefficient μ can be determined as follows [12]:

$$\mu = F_{Sf}/F_{Nf} \tag{29}$$

The test setup described and the results of the test with a 35 mm sealing ring are shown in **Figure 12**.

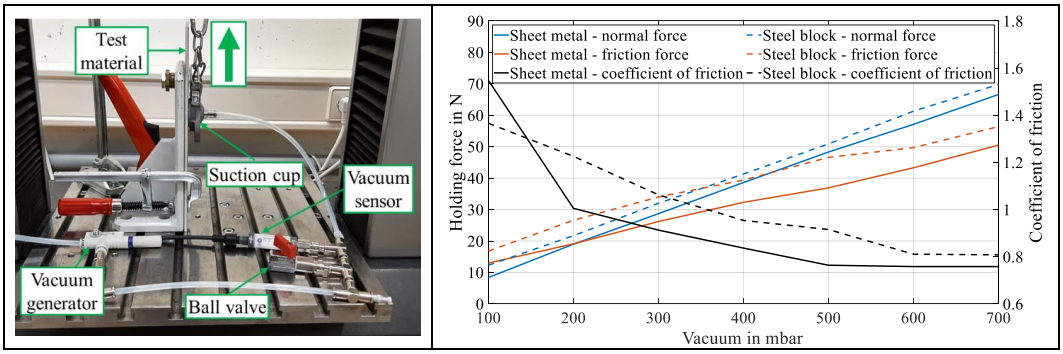


Figure 12: Experimental setup and results of the friction test

The results in **Figure 12** show that the coefficient of friction μ decreases as the normal force F_{Nf} increases.

4.3. Determination of the holding force as a function of the angle

The following test is carried out to validate the calculation approaches from section 3.1 and 3.2. A frame is used to set and fix the suction cup at a specific angle. The test piece is then sucked in the centre. The necessary vacuum is created by an ejector. A proportional valve then automatically releases air into the system so that the vacuum level in the system is continuously reduced. To slow down and stabilise the venting process, a tank is installed between the ejector and the valve. This also smooths any pressure fluctuations in the system. For the test, the vacuum in the system is reduced until the test part detaches from the suction cup and falls off. The vacuum curve is continuously recorded by three vacuum sensors mounted directly behind the sealing ring using a programmable logical controller (plc) with a cycle time of 1 ms. The arithmetically averaged vacuum curve can be used to determine the vacuum value at which the test piece releases from the suction cup. Using the vacuum value and the seal diameter, equation (1) can be used to determine the theoretical holding force F_{Th} required during disengagement. The holding force determined in this way is the same as the theoretical holding force determined by the methods described in sections 3.1 and 3.2. The test is carried out for 3 different sealing ring diameters. In order to record a complete curve, the test is repeated in 10° steps and at least five times per step. The test is also carried out for both test objects. The corresponding test setups are shown in **Figure 13**.

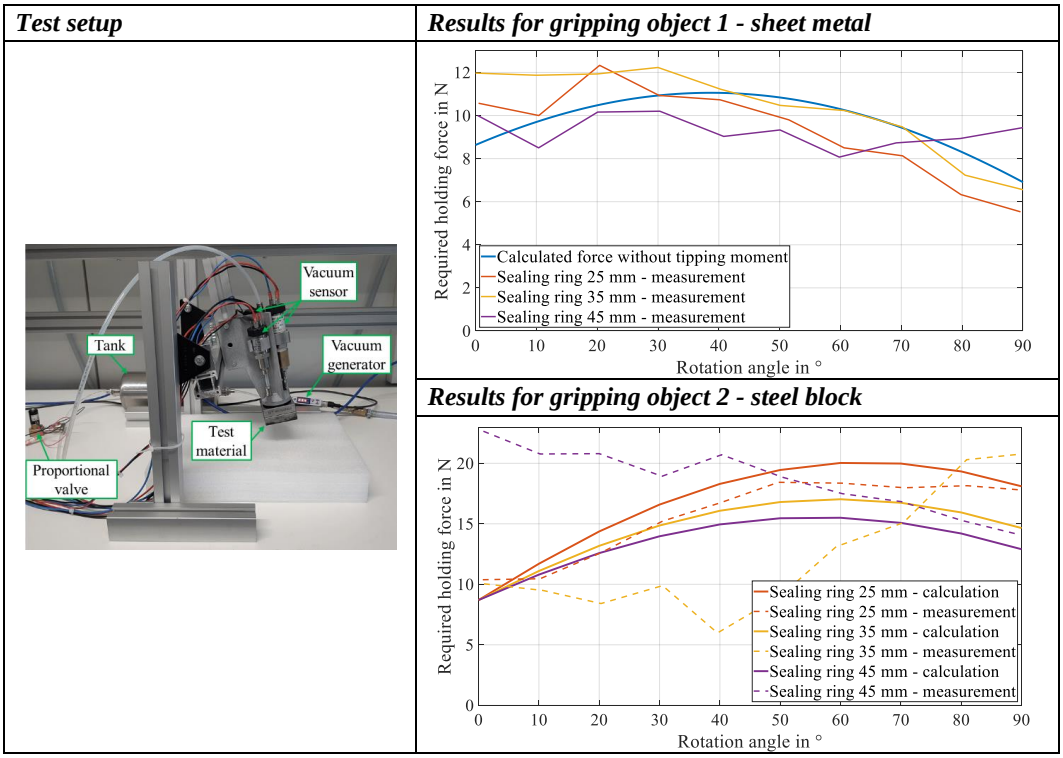


Figure 13: Test set-up and results of the holding force test

The process of gripping failure and thus the falling of objects due to the venting of the system is a highly dynamic process. Thus, it is difficult to determine the exact point of detachment. Several types of sensors (triangulation, hall and acceleration sensors) were tested to determine the detachment point. However, these were much slower than the vacuum sensors and only detected a detachment when the

vacuum sensors had already detected ambient pressure again. Therefore, for the evaluation, the detachment point was placed at the beginning of the abrupt pressure change.

As can be seen in **Figure 14**, the deviation of the release points over the series of measurements was significantly smaller for the small seal. However, the deviations for the larger seals were significantly higher.

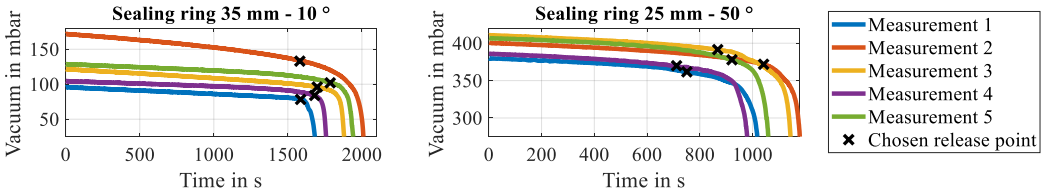


Figure 14: Pressure curve and the chosen release point

In addition, a constant static friction value from **Figure 12** was assumed for the calculation, which is approximately at the pressure level of the separation process.
The comparison of the test results with the calculated values shows that there is a good overlap for all sealing ring diameters for the "sheet metal" gripping object. The comparison for the second gripping object "steel block" shows that for the small seal diameter there is a very good match between calculation and test. For the larger ring diameters there are larger deviations in the range between 20 and 70°. However, the exact cause could not be determined with certainty due to the large number of possible influences. However, with the larger ring diameters the pressure in the system increases more slowly prior to separation than with the smaller ring diameters, making it more difficult to determine a specific detachment point.

5. CONCLUSION AND OUTLOOK

This paper addresses the challenge of efficient and safe design of vacuum handling systems. A review of the state of the art shows that, apart from a few discrete positions of the gripping object and the surface gripper, no other intermediate positions are considered when calculating the theoretically required holding force for handling gripping objects with surface grippers. Furthermore, the tipping moment due to spatial expansion of the gripping object is neglected. These aspects are taken into account in two new calculation approaches. The first approach extends the current state of the art to include additional intermediate positions for the gripping object and the gripper. The load cases known from the state of the art represent discrete individual positions. The second approach extends the first approach and takes the spatial expansion of the gripping object into account too. To validate the two calculation approaches, different static gripper gripping object positions were experimentally investigated. It was found that the experimental results for a small sealing ring were in very good agreement with the calculated values. Larger deviations occurred only for larger ring diameters and when the tipping moment was taken into account. The cause of these deviations could not be clearly identified due to the large number of influencing factors. Future work will focus on further detailing the calculation approach and extending it to the application of spatial and eccentric forces. It would also be interesting to quantify the reduction in energy consumption in various industrial application scenarios. In addition, most gripper systems use more than one suction cup. This therefore represents another research gap for future work.

6. ACKNOWLEDGEMENTS

This work was supported by the Landesministerium für Wissenschaft, Forschung und Kunst Baden Württemberg (Ministry of Science, Research and the Arts of the State of Baden-Wurttemberg) within

the Nachhaltigkeitsförderung (sustainability support) of the projects of the Exzellenzinitiative.

NOMENCLATURE

a	Acceleration	m/s²	M_S	Restoring force torque	Nm
b	Gripping object support width	m	M_{Th}	Theoretical holding force torque	Nm
d	Diameter	m	m	Gripping object weight	kg
F_A	Inertia force	N	n_3	Deformation coefficient	-
F_G	Weight force	N	Δp	Differential pressure	Pa
F_R	Frictional force	N	p_0	System pressure	Pa
F_S	Restoring force	N	p_U	Ambient pressure	Pa
$F_{Suction}$	Applicable suction force	N	S	Safety factor	-
F_{Th}	Theoretical holding force	N	t	Time	s
g	Acceleration due to gravity	m/s²	α	Rotation angle	°
h	Gripping object height	m	η	Efficiency due to leakage losses	-
M_A	Inertia force torque	Nm	μ	Coefficient of static friction	-
M_G	Weight force torque	Nm			

REFERENCES

1. Gabriel F, Bobka P, Dröder K (2020) Model-Based Design of Energy-Efficient Vacuum-Based Handling Processes. Procedia CIRP 93:538–543. https://doi.org/10.1016/j.procir.2020.03.006
2. Straub D (2020) Methode zur technischen Auslegung von Vakuumgreifsystemen mit einer Mindesthaltedauer auf Basis fluidischer Untersuchungen. Dissertation
3. Stegmaier V, Schaaf W, Jazdi N et al. (2023) Simulation Model for Digital Twins of Pneumatic Vacuum Ejectors. Chem Eng & Technol 46:71–79. https://doi.org/10.1002/ceat.202200358
4. Straub D, Huber K (2018) Potentials of vacuum gripping systems in human-robot-collaboration applications. VDE Verlag, Berlin, Offenbach
5. Fritz F, von Grabe C, Kuolt H et al. Benchmark of existing energy conversion efficiency definitions for pneumatic vacuum generators. https://doi.org/10.1007/978-981-4451-48-2_22
6. Stegmaier V, Schaaf W, Jazdi N et al. (2022) Anwendungsfälle und Ansatz zur Erstellung des Digitalen Zwillings aus Sicht eines Komponentenherstellers. In: Automation 2022: 23. Leitkongress der Mess- und Automatisierungstechnik : automation creates sustainability : Baden-Baden, 28. und 29. Juni 2022. VDI Verlag, Düsseldorf, pp 5–18
7. Stegmaier V, Eberhardt T, Schaaf W et al. (2023) Automated Configuration of Optimized Customer Specific Mechatronic Systems Using Behavior Models. CIRP Conference on Intelligent Computation in Manufacturing Engineering. https://doi.org/10.13140/RG.2.2.21121.22889
8. Hesse S (1997) Greiferanwendung: Blue Digest on Automation
9. Festo Didactic GmbH & Co. KG Grundlagen der Vakuumtechnik: Kurzübersicht. https://www.festo.com/net/supportportal/files/9916/grundlagen_vakuumtechnik.pdf. Accessed 10 Oct 2023
10. J. Schmalz GmbH Theoretische Haltekraft eines Sauggreifers: Typische Lastfälle. https://www.schmalz.com/de-de/vakuum-wissen/vakuumsystem-und-seine-bauteile/systemauslegung-berechnungsbeispiel/theoretische-haltekraft-eines-sauggreifers/. Accessed 10 Oct 2023
11. Spivak D (2021) The Science of Vacuum-Cup Forces. https://fluidpowerjournal.com/the-science-of-vacuum-cup-forces/. Accessed 10 Oct 2023
12. Kern P Elastomerreibung und Kraftübertragung beim Abscheren von aktiv betriebenen Vakuumgreifern auf rauen Oberflächen. Dissertation, Karlsruher Institut für Technologie (KIT); Karlsruher Institut für Technologie

Chapter 15

Control

ONLINE LEARNING OF CYLINDER VELOCITY CONTROLLERS FOR EXCAVATOR ASSISTANCE FUNCTIONS USING LOCAL MODEL NETWORKS

peer reviewed

Dr. Ozan Demir[1]*, Dr. Benjamin Hartmann[1], Naresh Mandipalli[1],
Dr. Frank Bender[2], Benjamin Ehlers[2], Michael Mehren[2]

[1] *Robert Bosch GmbH, Stuttgart, Germany*
[2] *Bosch Rexroth AG, Lohr am Main, Stuttgart, Germany*

** Corresponding author: E-mail address: ozan.demir@de.bosch.com*

ABSTRACT

This contribution introduces a data-based modeling approach using Local Model Networks for the online learning of cylinder velocity controllers that are applied for the realization of excavator assistance functions like reference tracking of the tool center point (TCP).

Even without any individual machine data, just using available data from a similar machine or expert knowledge, we can design an initial controller that is adapted during operation to improve the controller performance and to allow for automatic controller calibration. This allows for a significant reduction of manual machine commissioning efforts while ensuring the required accuracy of the assistance functions. In general, changes in the system behavior over machine lifetime could be compensated with our approach.

To show the effectiveness of the proposed strategy, we have applied the proposed machine learning method to a hydraulic excavator. The data-based controllers are adapted online using a rapid-prototyping system and are sufficiently fast to be implemented on a standard control unit. The control performance is comparable to traditional approaches while drastically reducing the time and effort for calibration.

Keywords: Excavator assistance functions, learning-based control, local model networks, online learning

1. INTRODUCTION

Recently, there has been an increasing demand on assistance functions for working machines, which is also triggered with the recent advances in the field of electro-hydraulics and applied robotics. Such functions can simplify machine usage, increase the overall productivity, and improve machine safety during operation [1].

One of the core technological requirements for assistance functions is the ability of an accurate path tracking for the tool center point (TCP) which is located at the tip of the bucket. Since an experienced operator can follow a desired path with a high accuracy, tracking accuracy requirements for assistance functions are also similarly high. Simple empirical controllers fail to provide the required accuracy since the kinematics and the hydraulic system of the excavator is complex and highly non-linear. Deriving first-principles models for the model-based controller design requires lots of time and effort. Hydraulic excavators are typically high-mix / low-volume products, which makes the development of a model-based controller for each excavator economically not feasible. In the course of this paper,

we will apply a learning-based control approach for this control problem. A presentation of the target assistance function for leveling and sloping operations can be also found in the following video: LINK (https://youtu.be/6xE5ZYCYxtw?si=DyCSdjKmNeBXOdOY).

For some time, speed control of hydraulic cylinders for autonomous and semi-autonomous working machines has been in the focus of the control and robotics community. While Bender et. al. [2] and Kalmari et. al. [3] introduce model-based controller structures, recently, pure data-based and hybrid controller structures have been presented. In [4], feedforward Neural Networks (NNs) were used as data-based models, while a Gaussian Process (GP) regression framework has been deployed for the same control problem in [5]. In [6], the authors present a reinforcement learning framework for control, which is trained offline, i.e., without interactions with the real machine using a data-based model. A similar control problem has been solved using local GP models for smaller robots in [7]. In all mentioned papers, the main focus has been the controller design rather than the efficiency of the data-collection process. An efficient data collection method is developed in [8] for the same application using active learning. The importance of an efficient design of experiments (DoE) grows if environmental factors and aging of the system behavior cannot be neglected. Another interesting challenge for data-based controllers is the handling of production tolerances, which can make it necessary to repeat the data collection process for each individual machine.

The main contribution of this work is an efficient online adaptation of data-based controllers for excavator assistance functions, running embedded on the excavator control unit. We apply the concept of local model networks to learn the inverse system behavior so that the obtained model can be used as an adaptive feed-forward controller. Using measurements on a real excavator, we will show that the controller can ensure the necessary accuracy during typical leveling and sloping movements. Moreover, in the case of a deviation from the optimal performance, for example caused by a lack of a rich data sets for the training of the initial model or aging effects, the initial controller can be adapted during regular operation under limited computational and storage resources.

This work is structured as follows. In the next section, we will present the test vehicle and the control problem which will be solved using adaptive data-based controllers. Section 3 introduces the fundamentals of local model networks for non-linear system identification. In Section 4, our proposed online adaptation concept is presented in detail. Vehicle measurements in Section 5 are used to evaluate the effectiveness of the introduced learning-based control strategy. Concluding remarks are given in Section 6.

2. PROBLEM SETUP

2.1. Test Vehicle

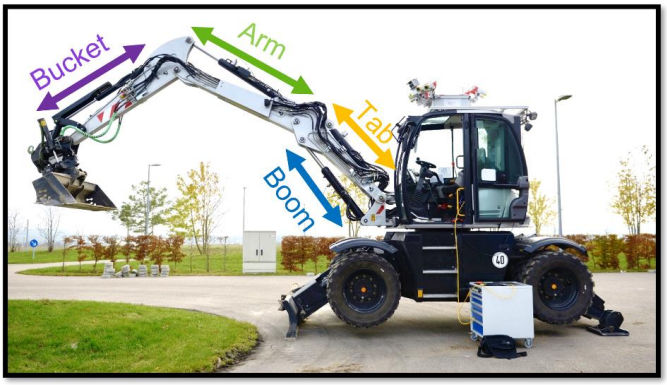


Figure 1: Test Vehicle: JCB Hydradig hydraulic excavator

The test vehicle as shown in Fig. 1 is a JCB Hydradig mobile hydraulic excavator which is used to validate the effectiveness of the proposed control strategy. The working arm of the excavator consists of four links actuated with four cylinders, called boom, tab, arm and bucket cylinder. Three of the cylinders are controlled using our proposed controller while the tab cylinder remains at a constant position during the excavator movement. The control input to the test vehicle is the modified output of the joystick signals. The operational software of the working machine is not modified.

The test vehicle is equipped with a *dSpace MicroAutoBox II* rapid prototyping system. The proposed functionality is implemented in Matlab & Simulink which allows for efficient code generation for the dSpace system.

2.2. Control Problem

The structure of the overall system for the path tracking problem is presented in Fig. 2. As stated in the introduction, focus of this paper will be the learning control for hydraulic cylinders. It is important to notice that the controller structure has also been used in our recent papers [4, 5].

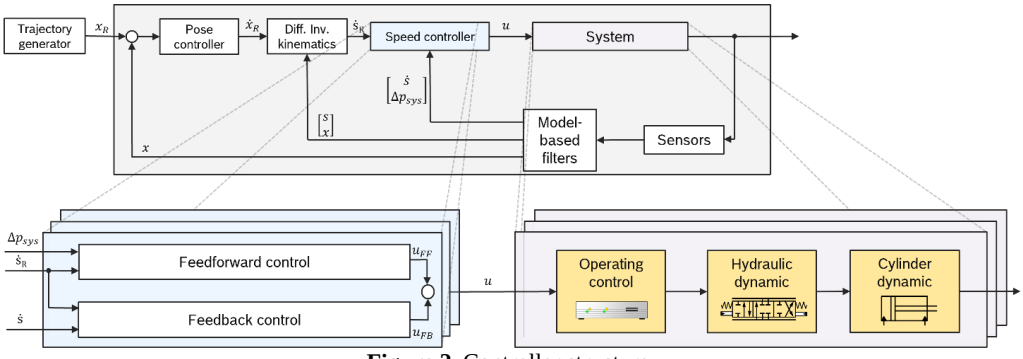


Figure 2: Controller structure

It is assumed that the kinematics of the excavator arm are exactly known so that differential inverse kinematics can be calculated and have not to be learned from data. Details on the optimization-based solution of the inverse kinematics problem, as it has been applied here, can be found in Bender et. al. [2]. Hence, the tracking problem can be transformed from TCP coordinates to joint coordinates. The feedforward controller is learned from data and reflects the inverse system behavior. The control variables are the three joystick signals.

3. LEARNING-BASED FEEDFORWARD CONTROL USING LOCAL MODEL NETWORKS

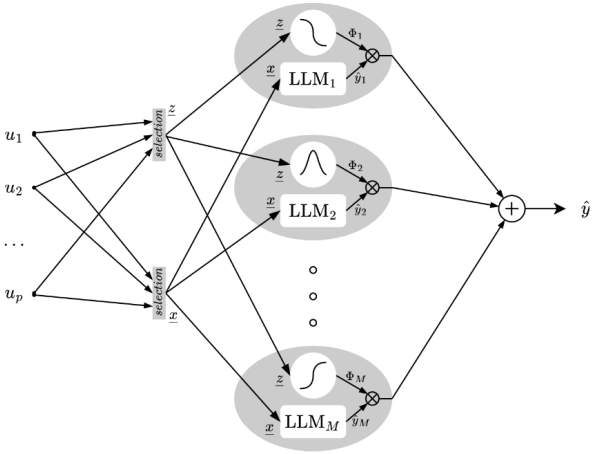


Figure 3: Illustration of a Local Model Network, see [10].

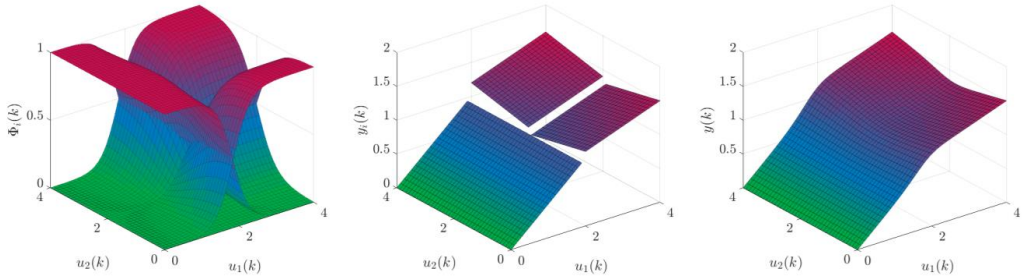


Figure 4: Local model network with two inputs. The model output (right) is calculated as the weighted interpolation of local linear models (middle) using the corresponding validity functions (left).

3.1. Local Model Networks (LoMoNet)

Local Model Networks (Fig. 3) are a special type of neural networks which were developed in the context of nonlinear system identification [10]. The main idea is to calculate the model output $\hat{y}$ as a weighted sum of local models $\hat{y}_i$. Here, the weightings are so-called *validity functions* Φ_i, and the local models are usually chosen to be of *linear* type:

$$\hat{y} = \sum_{i=1}^{M} \hat{y}_i(\underline{x})\Phi_i(\underline{z}) = \sum_{i=1}^{M} (w_{i0} + w_{i1}x_1 + w_{i2}x_2 + \cdots + w_{i,nx}x_{nx})\Phi_i(\underline{z})\,. \tag{1}$$

Using this model construction, it is possible to describe a nonlinear process behavior while having simple linear equations locally, see Fig. 4.

An important aspect of LoMoNet models is that the inputs to calculate the local models as well as the inputs to calculate the validities can be treated individually. Hence, it can be distinguished between the validity function inputs $\underline{z}$ that are influencing the model in a nonlinear way and the local model inputs $\underline{x}$ that affect the model output linearly. This is especially important, e.g., when dealing with dynamic processes where delayed versions of physical inputs have high correlation and should only be incorporated in $\underline{x}$, but are omitted in $\underline{z}$.

As shown in (1), the *i*-th local linear model (LLM) $\hat{y}_\mathrm{i}$ is described with:

$$\hat{y}_i(\underline{x}) = w_{i0} + w_{i1}x_1 + w_{i2}x_2 + \cdots + w_{i,nx}x_{nx} = \underline{\tilde{x}}^T \underline{w_i}\,, \tag{2}$$

where $\underline{w_i} = [w_{i0}\, w_{i1} \ldots w_{i,nx}]^T$ are the local model parameters, i.e., the offset and slopes in each input dimension and $\underline{\tilde{x}} = [1\, x_1 \ldots x_{nx}]^T$ the input vector for the local model. The regressor "1" in $\underline{\tilde{x}}$ extends the input vector $\underline{x}$ for the estimation of the offsets w_{i0}. The linear formulation has the benefit that standard (weighted) Least Squares methods can be applied for estimation. Furthermore, especially in the context of the underlying application, this is an essential feature, because it allows to use *recursive* algorithms for updating the parameters online.

The validity functions determine in which input region a local model is valid. The determination of the position and shape of these validity functions or the so-called partitioning, respectively, is a complex nonlinear optimization problem and, therefore, often is solved heuristically in practice.

The underlying work uses the Local Linear Model Tree (LoLiMoT) algorithm which is described in detail in [10]. Here, the validity functions are calculated as normalized Gaussian membership functions. The partitioning is achieved with orthogonal splits which lead to Gaussians, described with diagonal covariance matrices. Hence, the calculation of the membership functions simplifies to

$$\mu_i(\underline{z}) = \exp\left(-\frac{1}{2}\left(\frac{(z_1 - c_{i1})^2}{\sigma_{i1}^2} + \cdots + \frac{(z_{nz} - c_{i,nz})^2}{\sigma_{i,nz}^2}\right)\right), \tag{3}$$

where σ_{ij} are the standard deviations and c_{ij} the centers of each local model. The membership functions are then normalized to achieve a partition of unity. Hence, for each input sample $\underline{z}$ all validity functions sum up to one:

$$\Phi_i(\underline{z}) = \frac{\mu_i(\underline{z})}{\sum_{j=1}^{M} \mu_j(\underline{z})},\ \sum_{i=1}^{M} \Phi_i(\underline{z}) = 1\,. \tag{4}$$

Please, refer to [10] for more details on the functioning and characteristics of the LoLiMoT algorithm. However, important to note is that the LoLiMoT algorithm has two main hyperparameters: the maximum number of local models M and the smoothness value which proportionally affects the calculation of the standard deviations σ_{ij} in (3). Hence, the interpolation smoothness between the

local models is influenced. The higher the smoothness value is chosen the smoother the interpolation between local models becomes. Note that with higher interpolation smoothness the number of *effective* model parameters is reduced due to the local estimation of the LLM parameters. This leads to an implicit regularization effect, see [10] for details.

For the underlying application, LoMoNet showed similar performance compared to, e.g., Gaussian Process models or neural networks (MLP). However, the architecture of LoMoNet models is especially useful for our use case, because the model meets the memory and computation requirements of the embedded controller and, furthermore, can be adapted in real-time on the target device.

3.2. Online Learning using LoMoNet

As shown in Fig. 2, the speed controller contains a feed-forward part that captures the inverse behavior of a hydraulic actuation and determines a corresponding joystick position based on the desired trajectories, like described in chapter 2.2. In the following, the inverse plant behavior is modeled with the LoMoNet approach.

The learning of LoMoNet consists of two parts:

1. Learning structural parameters (σ_{ij}, c_{ij}) that capture nonlinear characteristics by partitioning the input space.
2. Learning the parameters of local linear model parameters $\underline{w_i}$.

The first part of learning is challenging due to the nonlinear formulation of the partitioning. In our work, the incremental, axis-orthogonal tree construction algorithm LoLiMoT was used to train the partitioning parameters from offline data. We assume that for similar excavators the partitioning does not change significantly, because the nonlinear behavior is comparable. Hence, the partitioning is trained offline based on data of a "golden sample" excavator and is then transferred to a new, similar excavator. Next, using the pre-trained partitioning, only the local model parameters are adapted to the new system. Therefore, it is possible to separate the complex offline training of the validity functions from the online learning of the local model weights. Note that it is even possible to manually define the input space partitioning based on physical system knowledge when the first commissioning of the machine is done with only a low number of inputs and initially no golden sample data is available.

Once the partitioning is available, the local model weights are adapted online, embedded in the control unit software. The initial values of the local model parameters are taken from the offline training as well, hence, are determined by the data of a similar excavator. For (online) adaptation of local model parameters, a state-of-the-art recursive least squares algorithm can be used.

4. EXPERIMENTAL RESULTS

Databased models only ensure a good model accuracy if the working space is covered with data in a dense way. In this work, we have been using an already available, rich dataset which has been also introduced in [4]. The dataset has been generated using quasi-random (multi-sine and amplitude modulated pseudo random binary signal (APRBS)) excitation of single and multiple cylinders. With this dataset, three different LoMoNet models were trained for the three cylinders with the following input features and model settings, see Table 1. Note that in all experiments the inputs for $\underline{x}$ and $\underline{z}$ are chosen equally.

Table 1: Input features and hyperparameters of LoMoNet models for different cylinders (#LMs = number of local models).

Model	Input features for $\underline{x}$ and $\underline{z}$	#LMs	smoothness
Boom Cylinder	speed, acceleration, Δp_{sys}	110	1.5
Arm Cylinder	speed, acceleration	90	1.2
Bucket Cylinder	speed, Δp_{sys}	51	1.2

4.1. Offline Validation of Online Learning using Vehicle Measurements

First, the obtained models have been validated using test data which wasn't included in the training data. While training data has been generated during random APBRS and multi-sine movements, test data has been collected using typical leveling and sloping operations. Levelling and sloping movements are repeatable excavator tasks with high accuracy requirements and low external load. As expected, all the controllers show a good modeling accuracy as shown in Table 2.

Next, the robustness of the obtained controller against changes in the system behavior has been analyzed. Such deviations in the behavior can arise due to aging or other environmental effects. Another possible reason for data-to-behavior deviations would be production tolerances, if collection of individual training data for each vehicle is omitted and initial controllers have been trained using the same nominal data set. In our experiments, in order to analyze the robustness of the databased controller against changes in the system, we switched at our test machine to a different operation mode (power mode); when not only the speed of the combustion engine and consequently the maximum available power for the hydraulic system but also the characteristics between the joystick position and volume flow demand changes. Please note that the vehicle behavior in power mode stimulates the behavior of a similar but not identical vehicle. Consequently, we observed a clear drop of the model accuracy which is reflected in the model errors shown in Table 2. Please notice that the training data has been collected in normal mode and we didn't repeat the data collection in power mode.

The bad performance obtained with the initial controller in power mode (i.e. at a different working machine) is not really surprising and the typical measure against this situation would be to repeat the data collection stage at the "new vehicle" i.e. at power mode. Consequently, new models can be trained using individual training at the cost of a higher data generation effort. However, even if the system behavior in power mode deviates from the behavior in normal mode, it is still similar so that it is reasonable to expect that the databased controller can achieve again a good performance after a slight adaptation of local model parameters. An offline analysis of vehicle measurements also validates our expectation; using online learning, we obtained a very significant improvement of the system accuracy, see Table 2 vs. Table 3. Table 3 shows that even at the normal mode, after online adaptation, the controller performance can be improved since in this way the model parameters can be optimized more for the leveling operation at the cost of the model generalization.

Table 2: Path tracking performance for different cylinders at normal mode and power mode.

Model	Normal mode		Power mode	
	RMSE [%]	R^2 [-]	RMSE [%]	R^2 [-]
Boom Cylinder	4.03	0.974	6.52	0.861
Arm Cylinder	4.83	0.983	5.84	0.952

Bucket Cylinder	4.29	0.973	4.62	0.948

Table 3: Path tracking performance for different cylinders at normal mode and power mode (online adaptation is enabled).

Model	Normal mode		Power mode	
	RMSE [%]	R^2 [-]	RMSE [%]	R^2 [-]
Boom Cylinder	2.14	0.992	1.99	0.987
Arm Cylinder	1.87	0.997	2.08	0.994
Bucket Cylinder	1.36	0.997	3.21	0.975

4.2. Online Validation at the Working Machine

After the offline analysis, the online adaptation of feed-forward databased controllers has been also validated at the test vehicle. Please note that during validation measurements, the feedback part of the cylinder speed controller is deactivated so that different feed-forward controllers can be directly compared.

In the measurements, first, online adaptation is disabled and the path following accuracy during typical levelling movements at normal mode has been evaluated. Figure 5 shows the tracking accuracy at different hydraulic cylinders.

As described above, we wanted to check the robustness of the databased controller against changes of the system behavior. Figure 6 shows the controller performance at power mode which represents a similar but not identical machine. As expected, we observed a significant drop of the controller performance. Especially, strong oscillations arise at boom cylinder which is not acceptable.

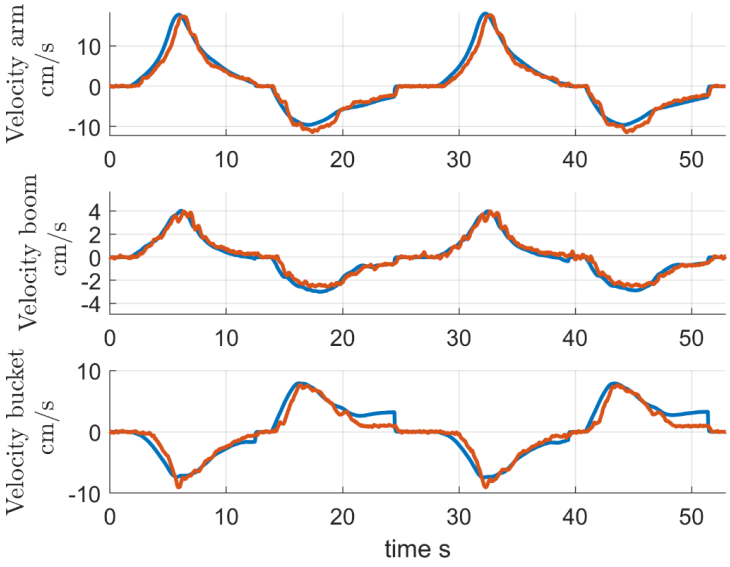


Figure 5: Trajectory tracking performance of databased controllers during leveling at normal mode (nominal machine). The desired cylinder speeds are depicted in blue and the measured speeds in red.

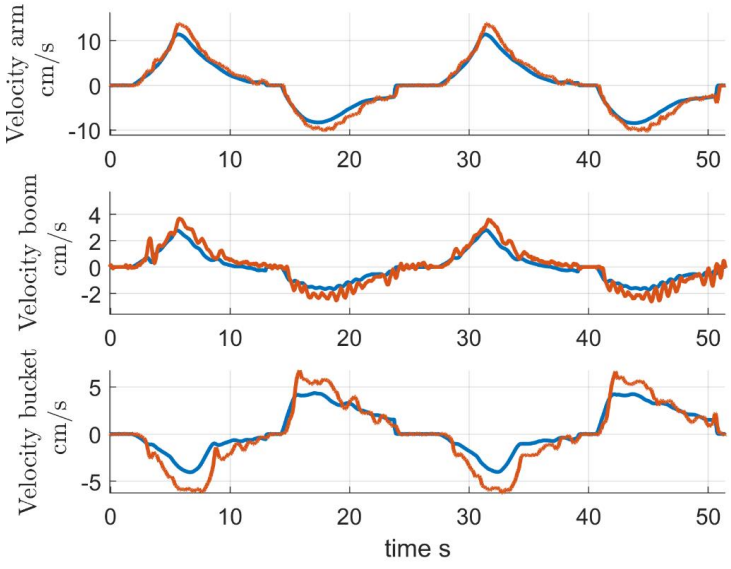


Figure 6: Trajectory tracking performance of databased controllers at power mode (i.e at a similar machine) during leveling. The desired cylinder speeds are depicted in blue and the measured speeds in red.

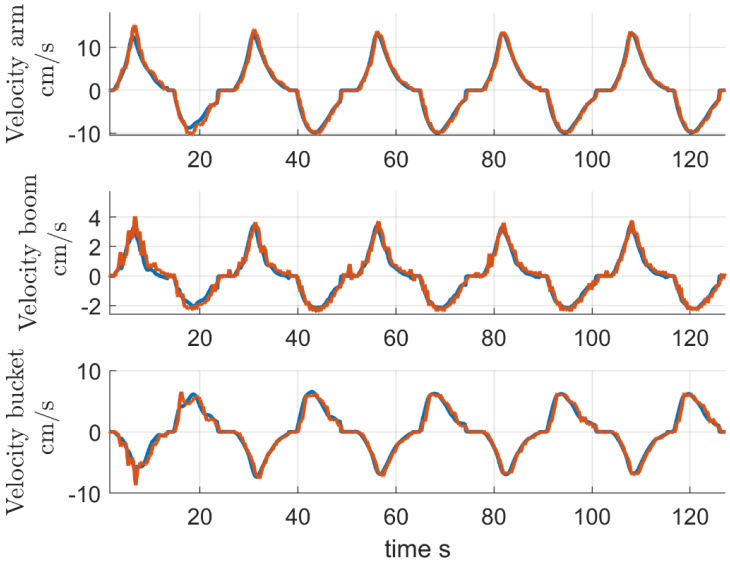


Figure 7: Trajectory tracking performance of databased controllers at power mode (i.e at a similar machine) during leveling. The desired cylinder speeds are depicted in blue and the measured speeds in red. Online adaptation of local model parameters is enabled, which ensured a low tracking error compared with the previous measurement presented in Figure 6.

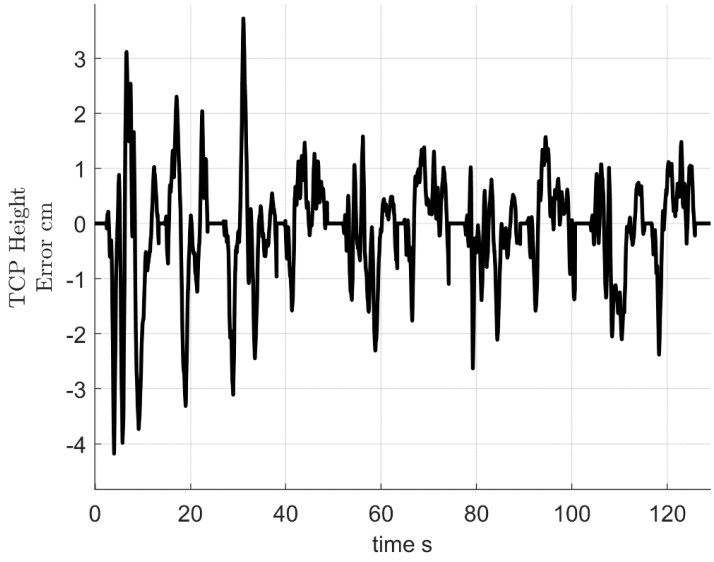


Figure 8: Leveling error (deviation from desired height) during online learning.

To be able to compensate for the change of the system behavior, online adaptation is enabled, and the local model parameters are adapted dependent on the obtained prediction error. Figure 7 clearly shows

that the controller performance can be improved in a significant way after a very short time. With the help of online adaptation, also the TCP-tracking accuracy can be improved, which is presented in Figure 8. Please notice that online adaptation ensures that the deviation from the desired height remains under 3cm which is a very impressive performance for the size of the excavator.

5. CONCLUSION AND OUTLOOK

In this work, we presented a learning-based control concept for the speed control of hydraulic cylinders using local model networks. The proposed framework models the complex hydraulic behavior with good accuracy. Furthermore, model parameters can be adapted in an efficient way, so that it is even possible to run it on the embedded control unit software. As a result, the robustness of the learning-based controller is significantly improved against changes of the system behavior. This is a very important achievement since such deviations are very typical during regular operation due to environmental effects, aging or part replacements.

In our current research, we are working on the robustness of the online learning algorithms against "bad" data, so that learning can be activated all the time and functionality shall decide which data can be used to adapt databased models. Another interesting research point is to learn databased models of the forward path of the system which can be then used for controller design.

NOMENCLATURE

TCP	Tool center point	
x	Pose of tool center point in horizontal, vertical, and relative bucket angle	m, m, rad
$\underline{x}_R$	Desired pose of tool center point in horizontal, vertical, and relative bucket angle	m, m, rad
s	Measured cylinder position for boom, arm, and bucket cylinder	m
s_R	Desired cylinder position for boom, arm, and bucket cylinder	m
u	Controller output (modified joystick output)	%
Δp_{sys}	System pressure i.e. pressure difference between pump pressure and load sensing pressure	bar
LLM	Local linear model	
w_i	Parameters of the i-th local linear model	
σ_{ij}	Standard deviations of i-th local model inputs	
c_{ij}	Center of i-th local model	
RMSE	Root mean square error	
R^2	Coefficient of determination	

REFERENCES

[1] Bender, F., Kaiser, N., Jayakumar, B., *Industry Spotlight: Automated operations for construction machines*, In: New Horizons – Building the Recovery. Hillhead Digital Conference, March 30-31, 2021

[2] Bender, F., Sonntag, M., Sawodny, O., *Nonlinear model predictive control of a hydraulic excavator using Hammerstein models*, In: Proceedings of the 6th International Conference on Automation,Robotics and Applications, pp. 557–562, 2015.

[3] Kalmari, J., Backman, J., Visala, A., *Nonlinear model predictive control of hydraulic forestry crane with automatic sway damping*, In: Computers and Electronics in Agriculture, vol. 109, pp. 36–45, 2014

[4] Weigand, J., Raible, J., Zantopp, N., Demir, O., Trachte, A., Wagner, A., Ruskowski, M., *Hybrid*

data-driven modelling for inverse control of hydraulic excavators, In: Proceedings of the International Conference on Intelligent Robots and Systems, 2021, pp. 2127–2134

[5] Rabenstein, G., Demir, O., Trachte, A., Graichen, K., *Data-driven feed-forward control of hydraulic cylinders using Gaussian process regression for excavator assistance functions*, In: Proc. of CCTA. Trieste, Italy, 2022, pp. 962–969.

[6] Egli, P., Hutter, M., *Towards RL-based hydraulic excavator automation*, In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020

[7] Nguyen D., Seeger M., Peters J., *Real-time local GP model learning*, In: From Motor Learning to Interaction Learning in Robots, Springer, 2010, pp. 193–207.

[8] Dio M., Demir O., Trachte A., Graichen K., *Safe Active Learning and Probabilistic Design of Experiment for Autonomous Hydraulic Excavators*, In: Proceedings of the International Conference on Intelligent Robots and Systems, 2023 (accepted).

[9] Siciliano, B., Sciavicco, L.,Villani, L., Oriolo, G., *Robotics: modelling, planning and control*. Springer, London, 2010.

[10] Nelles, O., *Nonlinear System Identification*. Springer, 2020.

VALIDATION OF A HYDRAULIC PULSE CONTROLLER ON AN OFF-HIGHWAY MACHINE

peer reviewed

Marvin Schell[1]*, Paul Foschum[2], Andreas Plöckinger[2], Thomas Zehetbauer[2], Mikko Huova[3], Matti Linjama[3], Jyrki Tammisto[3]

[1] *Andreas Lupold Hydrotechnik GmbH, Eythstraße 11, 72189 Vöhringen, Germany*
[2] *Linz Center of Mechatronics GmbH, Altenbergerstrasse 69, 4040 Linz, Austria*
[3] *Tampere University, Korkeakoulunkatu 6, FI-33720, Tampere, Finland*

* Corresponding author: Tel.: +49 7454 944-148; E-mail address: marvin.schell@lupold.de

ABSTRACT

There are different ways of controlling hydraulic variable displacement pumps, mostly the choice of control depends on the application and the respective requirements. With a focus on open circuit off-highway applications, the classic control structures such as pressure cut-off, load sensing, positive or negative flow control or power control are standard. The design of these controllers is usually the responsibility of the pump manufacturer and is often solved purely hydraulically. With the current trends, such as connectivity and increased efficiency of systems, a development into electrically actuated pumps becoming apparent. The preferred technology for this is usually a modified or improved version of the mechanical controller. Using proportional technology for oil flow control of the control system of a pump is one possibility but does not use the entire potential of an intelligent and efficient pump. With the approach of system optimization, a new controller based on two independent digital valves was already presented at the IFK in 2022. This hydraulic pulse controller (HPC) has since been developed further and a system comparison with different pumps has been tested on a modified test rig. In addition, its suitability for controlling an application was tested as part of a proof of concept on a real working machine. Furthermore, the development is now focused on the definition of this as a new smart subsystem and optimal system integration into existing architectures.

Keywords: digital hydraulic, Hydraulic Pulse Control (HPC), digital valve, variable displacement pump, efficiency, pump control, working maschine, flexibility, smart subsystem

1. INTRODUCTION

According to the current state of the art, the adjustment of variable displacement pumps is mostly managed by proportional valves. The in- or outflow of the control chamber is influenced by spool valves. In most cases, the system is controlled purely hydro-mechanically (passive), based on the area ratio, mechanical components like springs and the respective pressure level. Electro-hydraulic control valves can also be found in such applications; these are based on identical spool technology but are operated and positioned by a solenoid and controlled according to electrical signals.

In difference to the state of the art, the Hydraulic Pulse Controller (HPC) uses two independent fast-switching digital valves, which are controlled in a digital hydraulic manner (**Figure 1**). The valves have a hydraulic switching time of 2ms and are triggered with a PWM signal, which defines the

opening duration and the level of opening. Rapid opening and closing results in a corresponding cross-section, which can be changed quasi arbitrarily by adjusting the PWM dutycycle. This makes it possible to achieve control accuracy analogous to proportional valves.

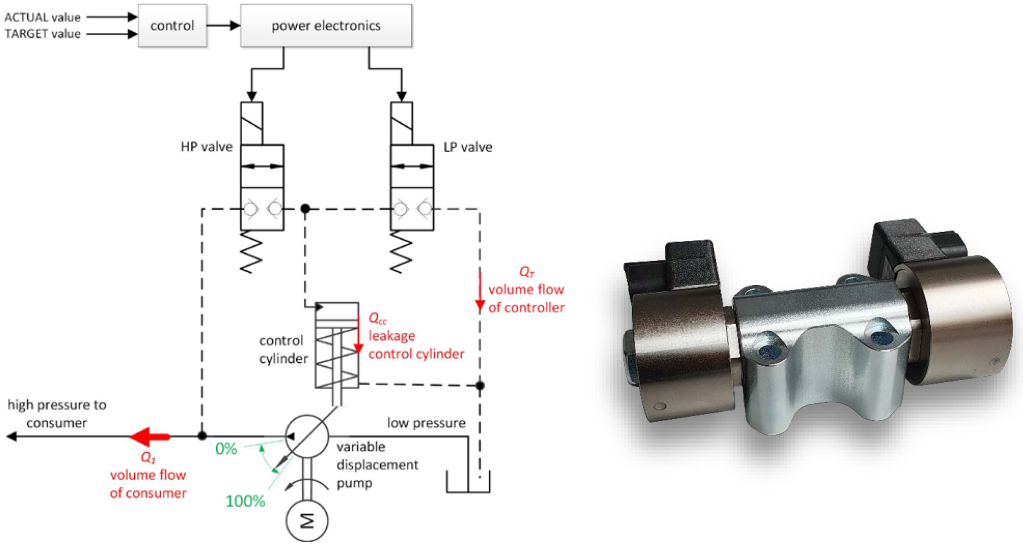


Figure 1: left: HPC (Hydraulic Pressure Control), digital pressure control schematic; right: HPC realized with 2 PULSE valves

This new way of acting on the control system results in significant advantages in terms of performance and energy savings. For example, the leakage or power loss of a previous spool valve is almost completely eliminated. Also, the two independent digital valves no longer represent a compromise between inflow and outflow of the oil flow, as it is on a spool with multiple control edges. This system is also less susceptible to contamination and offers enormous potential for energy savings by changing the operating mode and reducing the control pressure in the overall system. This controller was published for the first time on IFK2022 [1] and its functionality and advantages were proven through testing on various test benches. What is new in the meantime, however, is the extended control architecture, further conclusive tests with different pumps on a modified test rig in the laboratory and a test carried out on a real working machine in cooperation with the Innovative Hydraulics and Automation (IHA) of Tampere University, Linz Center of Mechatronics GmbH (LCM) and Andreas Lupold Hydrotechnik GmbH (ALH). Before the HPC was installed, comparative measurements were performed on an existing system with electro-hydraulic control (Rexroth). With a definition of certain test cycles, it was possible to evaluate the performance achieved and the general suitability of the system on a forwarder.

2. LABORATORY TESTS

For the established test rig [1] a hydraulic drive was used for powering the test pump (UUT, Unit Under Test) equipped with different controller types including the digital HPC controller. At the beginning the standard A10 DFR-Controller made by Lupold were applied and tested as some sort of reference and subsequent the idea of the digital control alternative was realized with on the market available 2/2-way seat type valves. They showed up to be too small for the aimed performance of the pump control and are therefore not suitable for such applications. Bigger valves were either not available or too slow in reaction time and not reaching the desired dynamic performance, so the digital valve PULSE was developed and optimized [2] upon which the current HPC (Hydraulic Pulse

Control) is based.

Figure 2 shows a picture of the PULSE valve and its duty cycle flow-rate behaviour at different pressure levels. It can be seen, that nominal down to very low flow rates can be achieved by choosing the appropriate duty cycle (50Hz excitation frequency), even at high pressure drops. Due to the fact, that the valve is not pressure equalized, the moment of opening respective minimal reasonable duty cycle to achieve useful flow rates varies depending on the differential pressure. Higher pressure drops lead to lower necessary minimum valve excitation times.

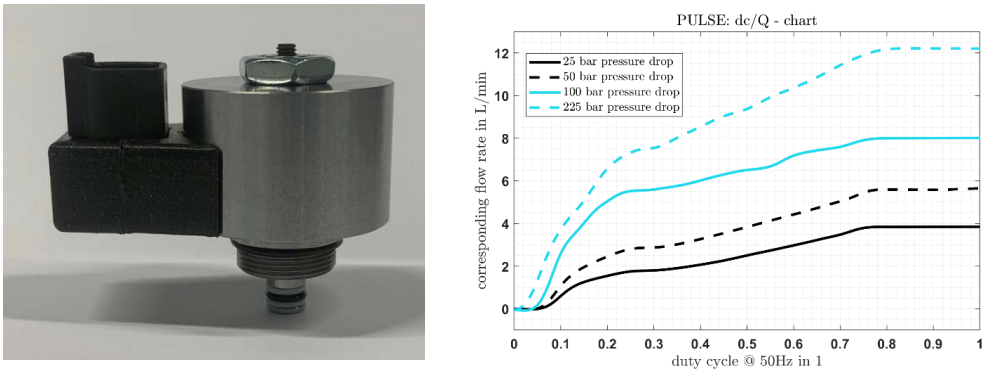


Figure 2: left: picture of the PULSE valve, right: characteristic diagram of flow rate at different duty cycles and pressure drops (based on 50Hz PWM-frequency)

2.1. The new measurement setup:

During completion of the former measuring objectives some issues and new requests showed up: The different controller types naturally lead to different results in performance and efficiency factor. Clearly determining the efficiency factor of the UUT in every working point would have been desirable. On the primarily side, in fact, only the pressure for the hydraulic drive was measured and the oil flow just estimated by the hydraulic source's swivel angle sensor. Due to the unknown characteristic of the hydraulic drive itself, it was impossible to clearly determine the mechanical power fed to the UUT. Comparing different controllers was of course possible but not the overall change of the efficiency factor of the UUT. Another issue was the motor-pump-load constellation, combining several hydraulic controllers, UUT and hydraulic load including the pressure control of primary oil source and hoses, which led to oscillations at some operating points and so fragmentary efficiency factor charts.

To get 'cleaner' results and additional information the test rig was advanced (**Figure 1Figure 3**). To reduce the hydraulic complexity the hydraulic drive was replaced with an 22kW electric asynchronous motor with frequency converter and to measure the mechanical power, driving the UUT, and a torque transducer with integrated measurement for the rotation speed was applied. To provide tank pressure oil for the UUT it was necessary to actively pump oil from the oil power unit below the laboratory upstairs. To keep the pressure at a level of about 3 bar a check valve was integrated in the back flow hose. This minimum pressure is also checked to prevent the UUT to be rotated without proper tank oil supply.

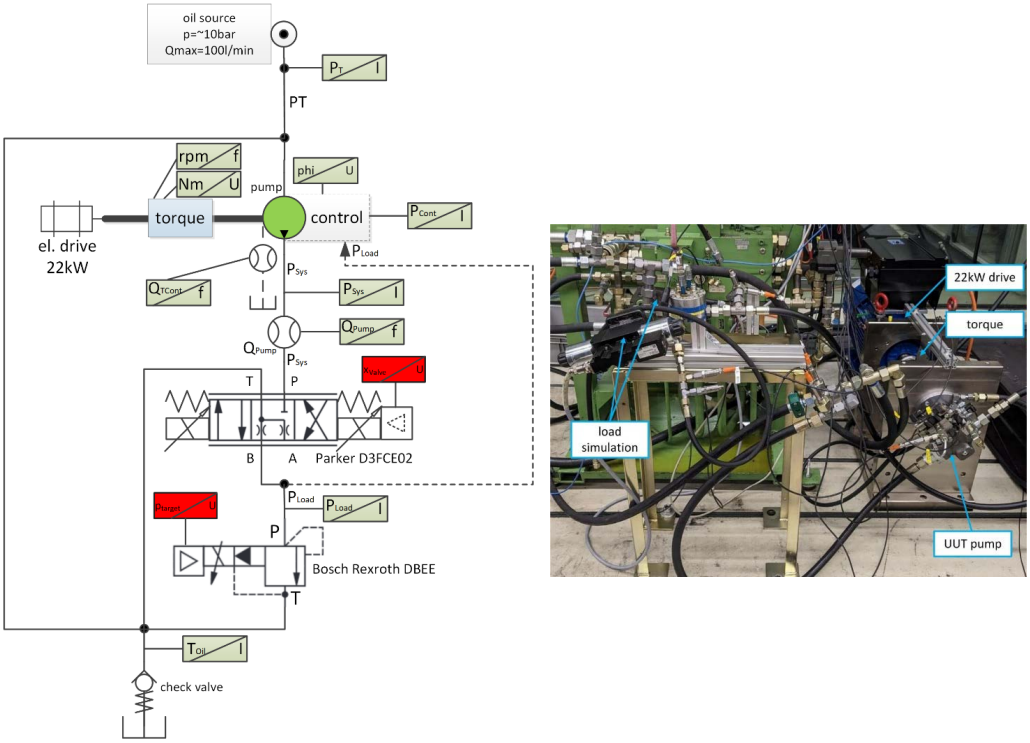


Figure 3: left: test rig setup schematic; right: picture of test rig at laboratory

Together with the volume flow sensor measuring Q_T (**Figure 1**) of the controller it is not only possible to distinguish between the mechanical power needed by the UUT to generate the required hydraulic power but also to identify the losses of the controller itself. The leakage Q_{CC} of the control cylinder is not considered as a loss of the controller but only the leakage of the low pressure valve. With a leakage free control cylinder at a stationary operation point the oil coming (leakage) from the controller's high pressure side would raise the pressure. To keep the operation point the oil would have to be dumped whether by the low pressure valve's leakage undesignedly or actively by piloting the low pressure valve. Either way, the amount of oil dumped by the low side valve comes from the high side valve at the pressure of p_S and so the power loss of the controller is estimated by:

$$P_{cont} = p_S \cdot Q_T \tag{1}$$

When it comes to comparing different controllers mounted on the same UUT the leakage of the control cylinder would be the same at the same operation points due to the leakage of the control cylinder. This loss is included in the overall efficiency factor as well as the hydraulic power needed for the controller.

$$P_{mech} = M \cdot \omega \tag{2}$$

$$P_{hydr} = p_S \cdot Q_1 \tag{3}$$

$$\eta_{tot} = P_{hydr}/P_{mech} \tag{4}$$

With the present setup the overall efficiency of the UUT and the characteristic of the controller can be determined (e.g. **Figure 4**, **Figure 5**). Especially the profiles in the right plots show the high

benefit in the overall efficiency using the HPC controller in comparison to the established hydro-mechanical controller, reaching an increase of up to 10% at lower loads. Some investigations with proportional valves can be found in [5]. Our measurements with similar valves showed comparable behaviour and the high leakage like the hydro-mechanical controller (DFR), so the digital seat valve based HPC noticeable exploits its strength vs. both technologies. This energy saving will also help to reduce losses at the upcoming electrification of mobile machines [6].

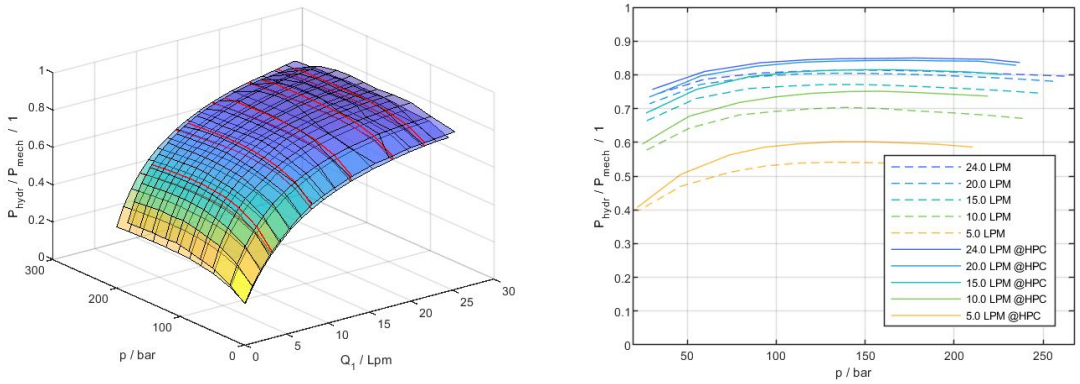


Figure 4: efficiency of 28ccm Bosch Rexroth A10 pump; hydro-mechanical vs. HPC controller

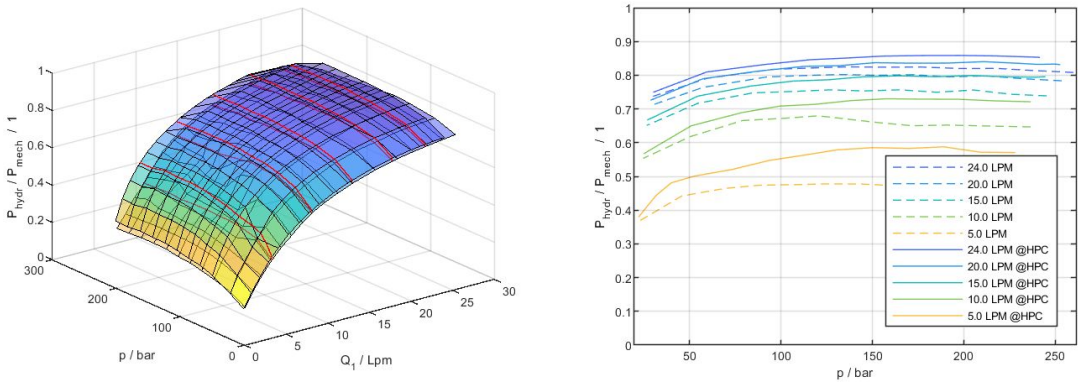


Figure 5: efficiency of 28ccm Liebherr pump; hydro-mechanical vs. HPC controller

2.2. Dynamic Behaviour

Besides the efficiency also the dynamic behavior of the pump-controller combination could be performed as it was formerly done [1] but in a more exact way (**Figure 6**, **Figure 7**, **Figure 8**).

The test cycle consists of several system excitations to determine the controller and UUT dynamics but also the accuracy. Starting with constant flow rates at 2 different states using constant xValve positions and the controller keeping the pressure drop at 16 bar the system pressures are ramped up and down at different pressure rates. In the second part of the test cycle (time > 120s) the system pressures should be kept at constant pressure levels while the flow rates are stepwise changed. Due to these impacts it is possible to show whether the controller works fast enough and if there are any oscillations at some operation point. Another benefit of such a challenging test cycle is the possibility to compare the different controller settings and pump reactions at extreme stress.

The overall goal of the comparison with the existing hydro-mechanical controller was not only to be more efficient but also to reduce oscillations (as shown in **Figure 6** between the time period of 80s

to 145s) without being less dynamical (rising or reducing pressure rates). Of course, another great advantage of an electrically actuated system is the possibility to change the controller's characteristic just in software without any mechanical modification even in already operating conditions.

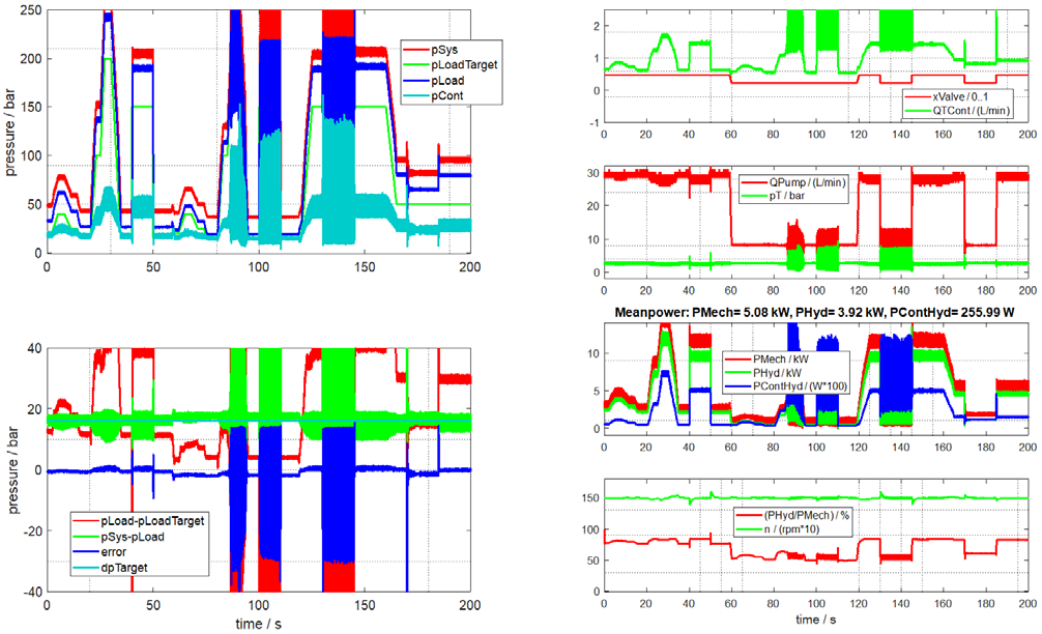


Figure 6: test cycle with Liebherr 28ccm and mechanical DFR controller

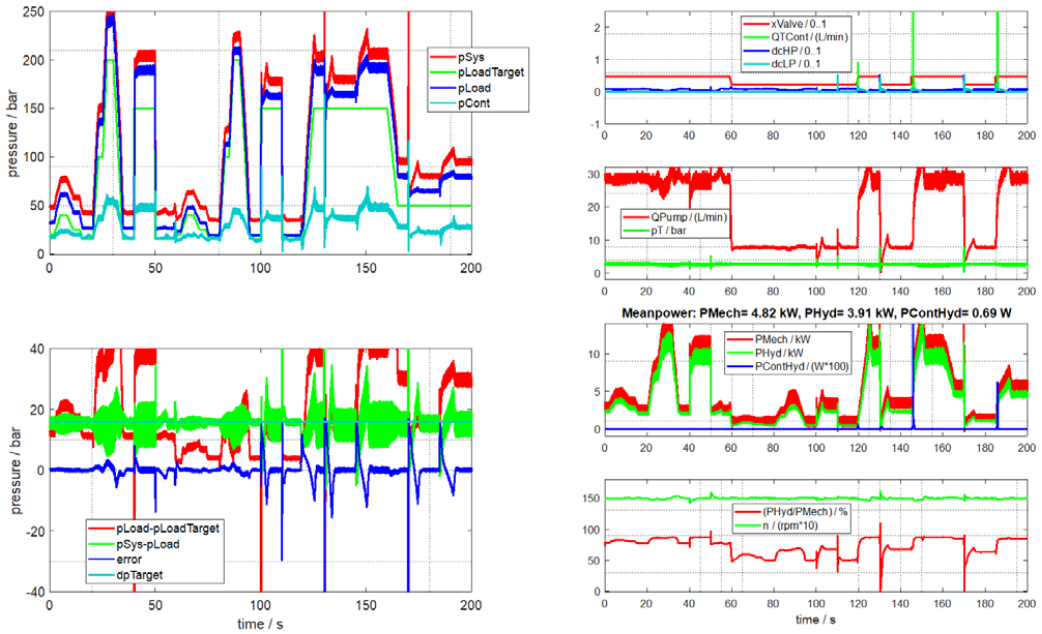


Figure 7: test cycle with Liebherr 28ccm and HPC controller

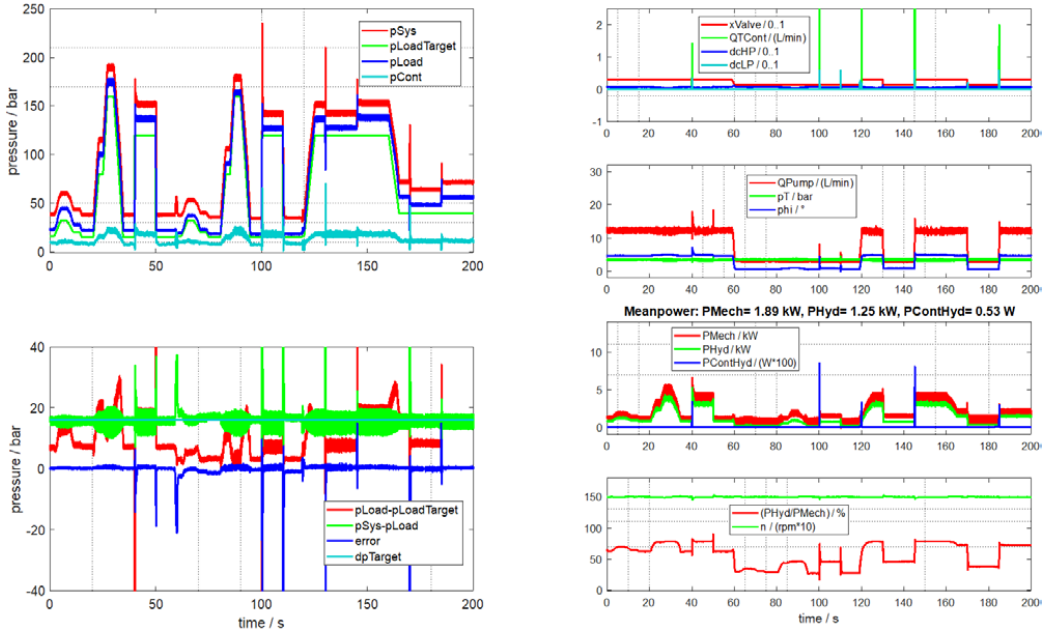


Figure 8: test cycle with Bosch Rexroth A10 28ccm and HPC controller with optimized parameters

The tests were performed with several variable displacement pumps from 18ccm to 32ccm of different suppliers which showed to be controllable out of the box without severe changing of the controller parameters. The performance could be improved with small adjustments during the measurements (e.g. **Figure 7** with standard; **Figure 8** with optimized parameters resulting in a shorter settling time and less noise). These promising results led to the next step, the proof of concept on a real working machine without the fear of serious problems.

2.3. Comparison of power consumption

Table 1 shows the power distribution of a Bosch Rexroth A10 (18ccm) with different controller types. The mean powers over one test cycle are listed and the corresponding power consumed by the controller split in electric and hydraulic power. Of course, the DFR controller doesn't consume any electrical power. It can be seen that the HPC controller is roughly 13% better than the DFR and about 9% better than using a standard proportional valve, in terms of energy efficiency.

Table 1: comparison of mean powers over a test cycle using Bosch Rexroth A10, 18ccm

	HPC with PULSE valves	Standard proportional valve	hydro-mechanical DFR
□□□□$_h$ / (kW)	2.01	2.32	2.37
□$_h$□□□ / (kW)	1.5	1.54	1.47
□$_{\text{electric,controller}}$ / (W)	0.76	6.1	-
□$_{\text{hydraulic,controller}}$ / (W)	1	252	385
η_{tot} / (1)	75%	66%	62%

3. PROOF OF CONCEPT ON A REAL WORKING MACHINE

Even though majority of development and testing of a pump controller could be carried out in laboratory, proof of concept testing on real applications is necessary. LCM, IHA and ALH decided to begin real application testing at IHA mobile machine test area located in Tampere Finland. Forwarder was selected as a test platform for the pump controller studied. The Ponsse Caribou S10 forwarder (**Figure 9**) is a machine used in cut-to-length (CTL) timber harvesting. The CTL method utilizes two machine types: 1) the harvester felling, delimbing and cutting, 2) the forwarder picking and transporting the piles of logs to roadside [3].



Figure 9: Ponsse Caribou forwarder as a test platform for the pump controller

The work functions of the forwarder include: slew, lift, luffing, extension, grapple rotator and grapple. These functions are realized using hydraulic actuators each powered by a common open circuit variable displacement pump (Bosch Rexroth 71 ccm A10 series) in load sensing configuration. The actuators are controlled by Parker K170LS mobile proportional directional control valve as presented in the simplified hydraulic circuit diagram in **Figure 10**.

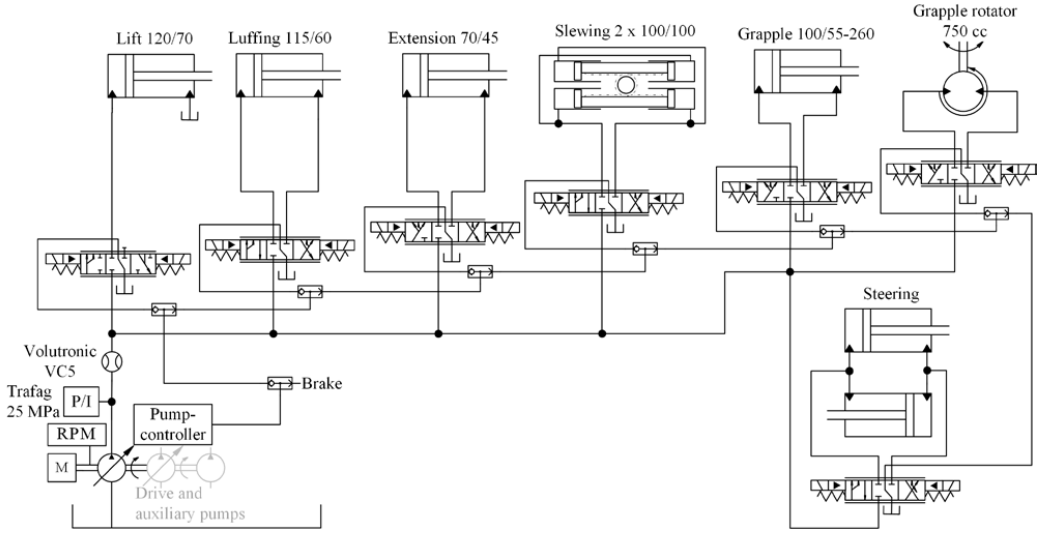


Figure 10: Simplified hydraulic circuit diagram of the forwarder

The data acquisition and control system is presented in **Figure 11**. The 4/3 control valves of the hydraulic actuators are driven by IFM current controller which receives the coil current references

from the Epec 5050 main controller. The main controller measures joystick signals, boom angles, cylinder positions and part of the actuator pressures. The rest of the pressures are measured using CBX-CAN converter. Crane slewing angle is measured with resolver which sends angle information to main controller via CAN bus.

The second Epec 5050 unit (Pump controller) measures diesel engine rotational speed, LS-pressure and pump angle. All the measured data is transmitted via CAN buses to Kvaser Memorator which is used here for data acquisition. Measurement results are post-processed and presented in MATLAB.

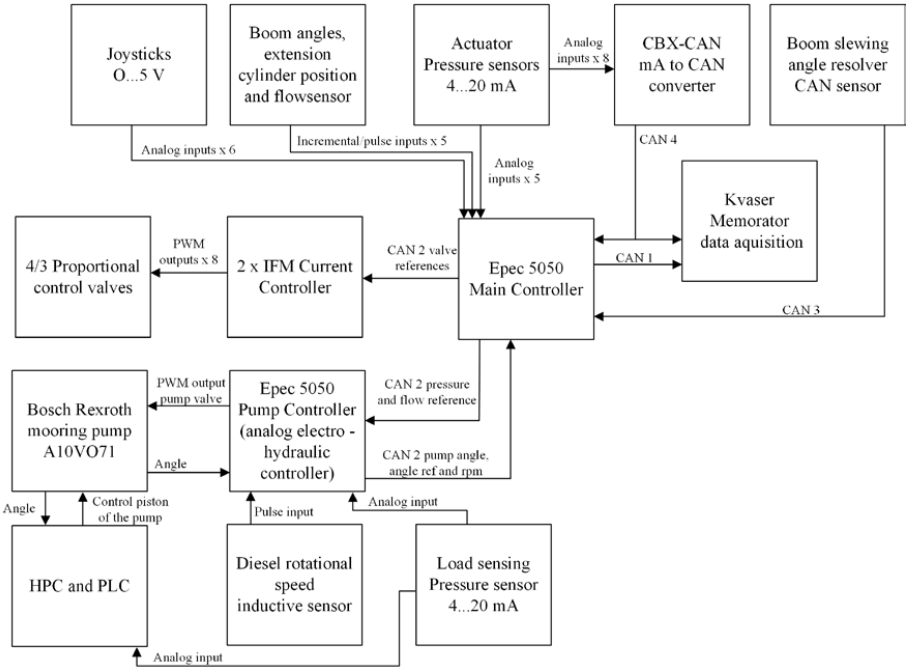


Figure 11: Data acquisition and control system used in the proof-of-concept testing

Figure 11 show both the HPC controller and the EPEC 5050 controller unit for the analog electro-hydraulic pump controller (EHC) using proportional valve from Bosch Rexroth. Naturally, only single pump controller is active at a time during the tests. The controller unit for the EHC uses the LS-pressure signal measured from the control valve LS-port to generate the pump pressure reference. The pressure reference is shaped using non-symmetrical rate limitation which allows fast increase of pressure, but slows down the decrease of pressure significantly. This results in relatively fast reaction to increased load pressure, while maintaining sufficient damping characteristics of the actuators and good operator feel when operating the crane. The pump pressure controller is of cascaded structure, where the inner loop controls the pump displacement. The reference value for the pump displacement is a sum of a flow feedforward and the output of the outer control loop, which controls the pump pressure. The flow feedforward is generated based on the coil currents of the actuator control valves. More information of the data acquisition and control system is given in [4].

The measurements performed with the machine include step responses of the lift actuator. It is worth to note that the actuator command is rate limited thus resulting in fast ramp-wise command. The purpose of these is to enable repeatable tests with less operator influences. The tests are carried out with luffing and extension actuators at minimum position and without load in the grapple. Initial position of the lift actuator is set such that grapple touches the ground.

To prove the applicability of the proposed pump controller in real mobile machine operation, also loading of logs from ground level into the load space of the forwarder is tested. These work cycles include simultaneous operation of more than one actuator and represent typical use of the machine.

Figure 12 and **Figure 13** present the step responses of the lift actuator with the baseline analog electro-hydraulic pump controller EHC and the proposed HPC controller. First column of diagrams presents piston position and velocity of the lift actuator and its valve command. The recorded valve command curve shows the compensation of the dead zone and the ramp-wise rate limitation. The second column of diagrams presents piston side pressure of the lift actuator, pump pressure, LS-pressure, pump flow rate, pump displacement, mechanical output power of the lift actuator and the output power of the pump.

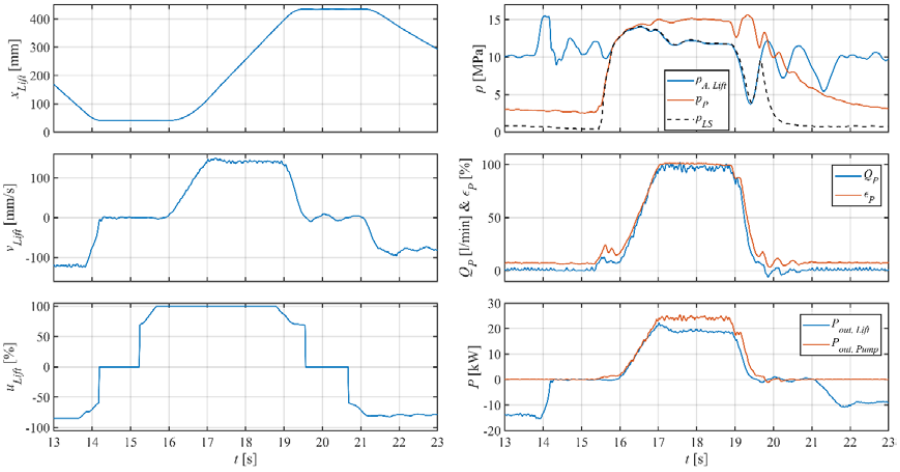


Figure 12: Lift actuator response with the EHC

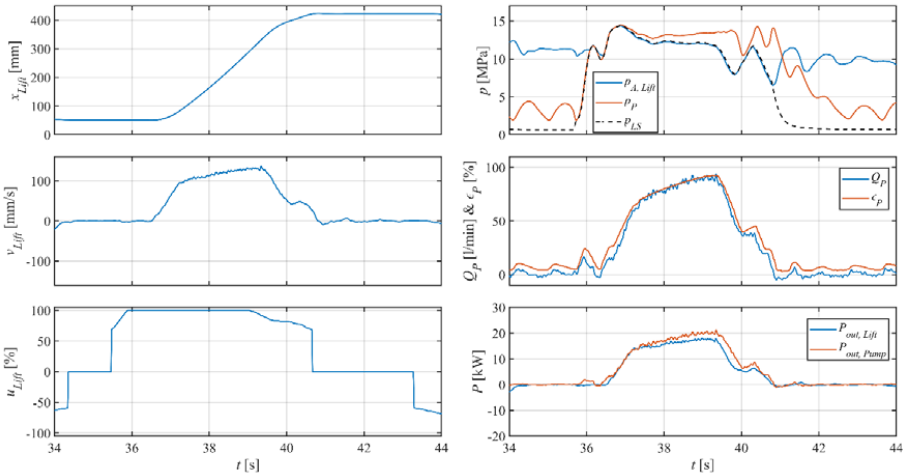


Figure 13: Lift actuator response with the HPC

Figure 14 presents an excerpt of the measured loading cycle. Piston positions and velocities of the four main actuators are presented as well as the pump pressure, LS-pressure, pump flow rate and

displacement. The results show that the HPC controller is capable of controlling the load sensing pump pressure and the crane can be used for real loading work. At certain operating points, however, a some oscillation of the pump pressure can be observed, which will be further optimised. Despite slight oscillation of pump pressure in these first proof-of-concept tests, the overall operator feel of the machine equipped with the HPC and pressure compensated control valves is good.

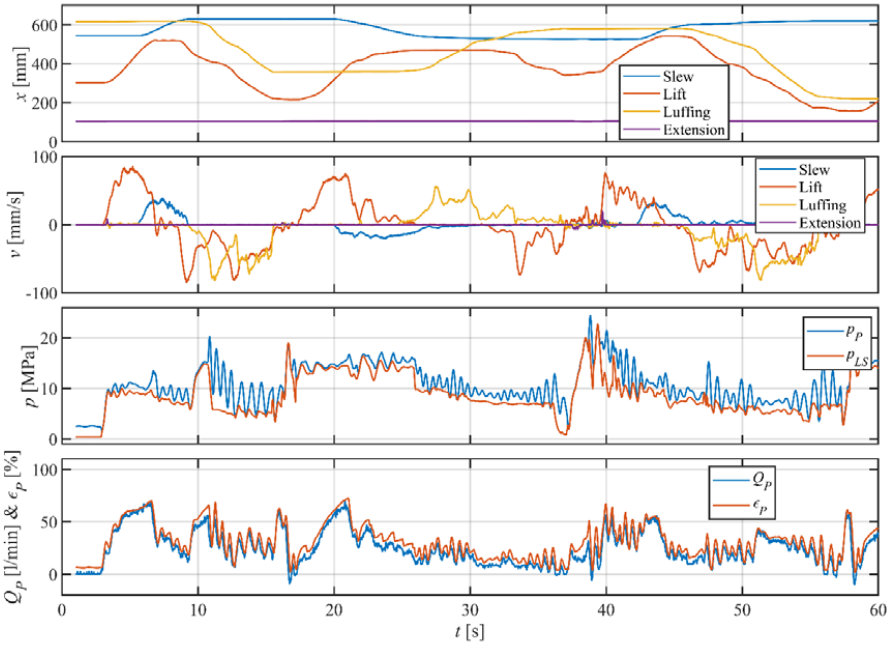


Figure 14: Excerpt of the measured loading cycle with HPC controller

4. CONCLUSION & OUTLOOK

The results of the tests in the laboratory have already shown that the Hydraulic Pulse Controller can be optimally combined with various pumps from different manufacturers. The structure of the controller is designed in such a way that tuning to a different pump is quickly possible. By modifying the test rig for better and measurable comparison of different controllers, it was also demonstrated that the efficiency increase was achieved. In the context of the tests in the laboratory the influencing variables from outside were less, but still these tests provide essential knowledge for the further development and the preparation of tests on complete machines.

The complexity to control a pump in a real working machine is a big challenge in many aspects. For example, the pump must supply oil to a number of actuators at the same time and react optimally to influences such as pressure changes, temperature or oscillations in order to ensure the operability of the machine. For this reason, it is also necessary to prove the suitability of the HPC in this environment. The pre-setup of the control parameters was originated from the lab tests and not specifically adjusted to the pump (71ccm A10 Bosch Rexroth) of the machine. From the first attempt, the HPC was able to control the pump according to the operator's requirements. The test period with the HPC installed on the machine was limited to a few hours, during which, with the help of some analysis, the controller parameters were further tuned until the performance was at a comparable level to the previously installed system. The full potential of the control quality was not yet fully exploited in this short time frame. However, the commissioning phase has already shown that the HPC can significantly reduce the time required here. In addition, adaptation to the respective application or a

specific work cycle is much better possible. This is mainly realized by moving essential functions into software. Now it is possible to influence a characteristic or the behavior of the pump without adapting the mechanics. Another advantage is the connectivity between the machine and the pump, so that important information can be exchanged, such as information of the pump status.

In interaction with the machine, the HPC takes up the role of enabling the intelligence of the pump. The importance of software-based functions is growing, as this is the only way to react flexibly to new requirements according to [7]. In this paper, there was mostly only 'HPC' mentioned, but this has developed increasingly into a smart subsystem. An intelligent node in the bus-communication closes the gap in the network between the machine and the pump. This control system can be smoothly integrated into existing architectures. With this subsystem, further potentials are now available, such as optimized interaction of the entire hydraulic system to reduce the standby pressure. To validate these further advantages of the HPC, further tests will be carried out on the test bench and on mobile machines. The controller architecture will also be further developed and additional functions added. These include, for example model-based control including temperature dependencies and pump design parameters, or AI-based data evaluation including condition monitoring for early fault detection such as increasing leakage or aberrant behavior.

NOMENCLATURE

Variable	*Description*	*Unit*
Q_T	Volume flow of the controller measured at low pressure side	[L/min]
Q_1	Volume flow of pump to consumer (load simulation) at high pressure side	[L/min]
p_S, p_A, p_T, p_V	System pressure, controller p., lowside p., pressure used by consumer	[bar]
E_{hydr}	Hydraulic energy	[Ws]
P_{hydr}	Hydraulic power	[W]
P_{mech}	Mechanical power of the drive shaft	[W]
P_{cont}	Hydraulic power of controller	[W]
Q	Volume flow	[L/min]
M	Mechanical torque of the motor shaft	[Nm]
ω	Angular velocity	[rad/s]
p	Hydraulic pressure	[bar]
phi	Swivel angle of the variable displacement pump	[°]
dc	Duty cycle	[%]
η_{tot}	Overall efficiency factor	[1]

ACKNOWLEDGEMENT

The presented research work has partly been supported by Andreas Lupold Hydrotechnik GmbH (ALH), Tampere University and the Linz Center of Mechatronics GmbH (LCM), which is part of the COMET/K2program of the Federal Ministry of Transport, Innovation and Technology and the Federal Ministry of Economics and Labor of Austria. The authors would like to thank the Austrian and Upper Austrian Government for their support.

REFERENCES

[1] Foschum P, Plöckinger A, Winkler B, Schell M, Lohmiller D (2022) Development of a hybrid pump control based on a digital hydraulic controller. 13th International Fluid Power Conference. March 21-23, 2022, Aachen, Germany.

[2] Zehetbauer T, Plöckinger A, Foschum P, Schell M, Lohmiller D (2022) A journey of a compact digital valve development; from scratch to series production. 11th Workshop on Digital Fluid Power. September 19-20, 2022, Edinburgh, Scotland.

[3] The cut-to-length method. Accessed 24th Oct 2023. https://www.ponsse.com/cut-to-length#/

[4] Huova M, Tammisto J, Linjama M (2020) Open-loop independent metering control of a multi-DOF forwarder boom. International Journal of Fluid Power, vol. 21, pp. 147–168.

[5] Mohn U G, Tetik S, Hefner J (2023) Electronic Open Circuit (eOC) – Elektronische Regelung von Axialkolbenpumpen im offenen Kreislauf; Karlsruher Schriftenreihe Fahrzeugtechnik, 9. Fachtagung, Hybride und energieeffiziente Antriebe für mobile Arbeitsmaschinen. February 28, 2023, Karlsruhe, Germany.

[6] Fassbender D, Brach C, Minav T (2022) Development of a hybrid pump control based on a digital hydraulic controller. 13th International Fluid Power Conference. March 21-23, 2022, Aachen, Germany.

[7] Dr.-Ing. Martin Petzold (2023) MOBILE Off-Highway Solutions. Dresdner Fluidtechnisches Kolloquium, Professur für Fluid-Mechatronische Systemtechnik. October 13, 2023, Dresden, Germany.

MODEL PREDICTIVE CONTROL OF ELECTRO-HYDRAULIC SYSTEMS WITH MULTIPLE DEGREES OF FREEDOM

Thomas Sendelbach[1]*, Tobias Leutbecher[1]

[1]*Bosch Rexroth AG, Partensteiner Straße 23, 97816 Lohr, Germany*

* Corresponding author: Tel.: +49 9352 18-4620; E-mail address: thomas.sendelbach@boschrexroth.de

ABSTRACT

Modern hydraulic drives have an ever-increasing power density and robustness, however they become more and more complex in their design and control. In many systems, there is the possibility of using an overdetermination of system control inputs to optimize the operating strategy for energy efficiency or tracking error. Many applications are still largely based on empirical and hard coded rules. More advanced methods are using offline optimization algorithms [1] to calculate an optimized trajectory for more than one manipulated variable for a given command trajectory. However, this approach leads to a lack of robustness and flexibility if model equations are not exact enough, real time control is required or the operating point changes. Traditional algorithms lack in standardisation and scalability which is also crucial for success in the industry. To overcome the disadvantages of rule based or offline optimization methods this paper presents the fundamentals and the application of Model Predictive Control (MPC) with respect to electro-hydraulic drives. Furthermore, possibilities are described to make advanced algorithms economically transferable into series production.

Keywords: model predictive control, digitalization, sustainability, optimization

1. INTRODUCTION

1.1. Motivation

The system under investigation comprises two variable displacement hydraulic axial piston pumps attached to one single speed variable motor shaft. The pumps drive respective flows for two chambers of a differential hydraulic cylinder. Leveraging the flexibility of this setup, MPC is used to exploit the available input degrees of freedom (DOF) to match and weight different performance goals and constraints [2]. The algorithm demonstrates the effectiveness of optimization-based control in managing the system for optimal energy strategies and control under state constraints.

2. OPTIMAL CONTROL OF AN ELECTROHYDRAULIC SYSTEM WITH MULTIPLE INPUTS

2.1. Problem Description

Model predictive control is a broad family of methods which solve an optimal control problem iteratively. If one stays in matured space of MPC a well-established set of standardized implementations is at hand. Here, the application is a complex electro-hydraulic system with three-degree-of-freedom (3DOF) in its control inputs. The system consists of a permanent magnet synchronous motor, two displacement variable axial piston pumps on a single speed variable shaft and a differential hydraulic cylinder. The basic idea is to actively use the shaft speed and the two

pump displacements to control the cylinder position while optimizing a cost function and complying with the defined boundary conditions.

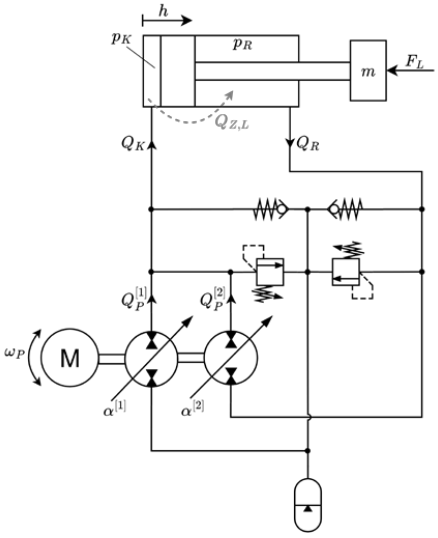


Figure 1: Scheme for the electrohydraulic drive system – summating transformer

Table 1: Parameters of the system

Cylinder parameter	Value
Area A_K	85.33×10^{-3} [m²]
Area A_R	42.66×10^{-3} [m²]
Mass m	3.0×10^{3} [kg]
Volume pump 1,2	4.0×10^{-5} [m³]
Piston stroke	0.4 [m]

2.2. Modelling of the drive system

The continuous-time model of the hydraulic cylinder drive system, with two pumps on a speed variable shaft, can be expressed in state space as in figure 2. The control inputs are $\omega_{P,cmd}$, $\alpha_{cmd}^{[1]}$ and $\alpha_{cmd}^{[2]}$. From top to bottom $\dot{p_K}$ and $\dot{p_R}$ describe the pressure dynamics of the respective chamber. The piston velocity $\dot{h}$ and piston acceleration $\ddot{h}$ describe the mechanical equation of the cylinder. The swivel rate $\dot{\alpha}^{[1,2]}$ and swivel acceleration $\ddot{\alpha}^{[1,2]}$ describes the swivel angle dynamics of both pumps, simplified to a second order system. The rotational speed ω_p and torque generating current I_q finishes the set of equations. Several simplifications were done by neglecting the suction circuit and pressure relief valve in the prediction model. Both assumptions are permissible because the operating strategy will be chosen to avoid the opening of the suction valves by penalising the controller to drop below defined threshold pressure.

$$\dot{x} = \begin{pmatrix} \dot{p}_K \\ \dot{p}_R \\ \dot{h} \\ \ddot{h} \\ \dot{\alpha}^{[1]} \\ \ddot{\alpha}^{[1]} \\ \dot{\alpha}^{[2]} \\ \ddot{\alpha}^{[2]} \\ \dot{\omega}_P \\ \dot{I}_q \end{pmatrix} = \begin{pmatrix} \frac{E'(p_K)}{V_K(h)} \cdot \left(-G_{Z,L}(p_R - p_K) + Q_P^{[1]}(\eta_v) + Q_P^{[2]}(\eta_v) - \frac{\rho(p_K)}{\rho_0^{[f]}} A_K \dot{h} \right) \\ \frac{E'(p_R)}{V_R(h)} \cdot \left(G_{Z,L}(p_R - p_K) - Q_P^{[2]}(\eta_v) + \frac{\rho(p_R)}{\rho_0^{[f]}} A_R \dot{h} \right) \\ \frac{\mathrm{d}}{\mathrm{d}t} h \\ \frac{1}{m} \cdot (p_K A_K - p_R A_R - F_R - F_L) \\ \frac{\mathrm{d}}{\mathrm{d}t} \alpha^{[1]} \\ \frac{1}{\left(T^{[1]}\right)^2} \alpha_{cmd}^{[1]} - \frac{2D^{[1]}}{T^{[1]}} \dot{\alpha}^{[1]} - \frac{1}{\left(T^{[1]}\right)^2} \alpha^{[1]} \\ \frac{\mathrm{d}}{\mathrm{d}t} \alpha^{[2]} \\ \frac{1}{\left(T^{[2]}\right)^2} \alpha_{cmd}^{[2]} - \frac{2D^{[2]}}{T^{[2]}} \dot{\alpha}^{[2]} - \frac{1}{\left(T^{[2]}\right)^2} \alpha^{[2]} \\ \frac{1}{J} \cdot \left(k_M I_q - M_{M,R} - M_L^{[1]}(\eta_m) - M_L^{[2]}(\eta_m) \right) \\ \frac{1}{T_q} \cdot \left(k_n \left(\omega_{P,cmd} - \omega_P \right) + \frac{k_n}{T_n} \int \left(\omega_{P,cmd} - \omega_P \right) \, \mathrm{d}t \right) - \frac{1}{T_q} I_q \end{pmatrix}$$

Figure 2: Detailed dynamical model for the electro-hydraulic system

M_L represents the load torque, including mechanical efficiency model of the axial piston pumps. The function is described with the help of mechanical efficiency maps (3).

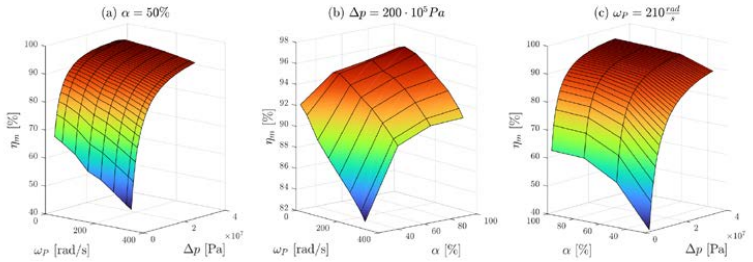


Figure 3: Mechanical efficiency maps regarding the axial piston pumps M_L

Q_p represents the pump flow, including the volumetric efficiency model of the axial piston pumps, which are described by volumetric efficiency maps (4).

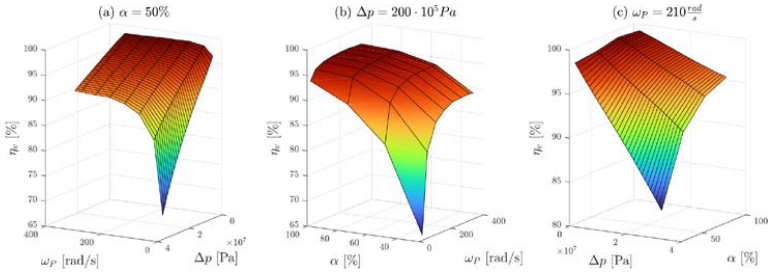


Figure 4: Volumetric efficiency maps regarding the axial piston pumps Q_P

The electric power losses are comprised by simplified formula for copper (1) and iron losses (2).

$$P_{Mot,Cu,eff} = \frac{3}{2} \cdot \mathrm{R_S} \cdot \left(\mathrm{I_{q,eff}}\right)^2 \tag{1}$$

Where R_S is the electrical resistance of the copper windings and $I_{q,eff}$ is the effective torque generating current.

$$P_{Mot,Fe} = \mathrm{k_{Hys}} \cdot \omega_M \cdot sign(\omega_M) + \mathrm{k_w} \cdot (\omega_M)^2 \tag{2}$$

ω_M is the rotational motor speed and $\mathrm{k_{Hys}}, \mathrm{k_w}$ are fitted motor factors. All these loss-models build the foundation for an operating strategy which allows the efficient operation of the system [1]. The aggregate loss model (3) consists of the volumetric and mechanical losses for both pumps, the iron and copper losses of the motor, the inverter losses and mechanical cylinder losses. For further details, we recommend studying [2].

$$\begin{aligned} P_{loss} = {} & P_{P1,Vol} + P_{P2,Vol} + P_{P1,Mech} + P_{P2,Me} + P_{Mot,Fe} \\ & + P_{Mot,Cu} + P_{Mot,Me} + P_{Cv,f} + P_{Cyl,Me} \end{aligned} \tag{3}$$

3. CONTROL DESIGN

3.1. MPC and adaptive linear MPC

Model Predictive Control (MPC) is an advanced control strategy that computes control inputs by solving an optimization problem at each time step. It involves the use of a model of the system to predict its future behaviour over a finite horizon. The MPC algorithm computes the control inputs in a way that minimizes a cost function, typically by finding a compromise between the deviation from a desired command trajectory and the control effort. Once the optimization problem is solved, only the first control input of the optimized sequence is applied to the system. Then the whole process is repeated at the next time step. This 'receding horizon' strategy allows MPC to handle multivariable systems with constraints on inputs, system states, and outputs.

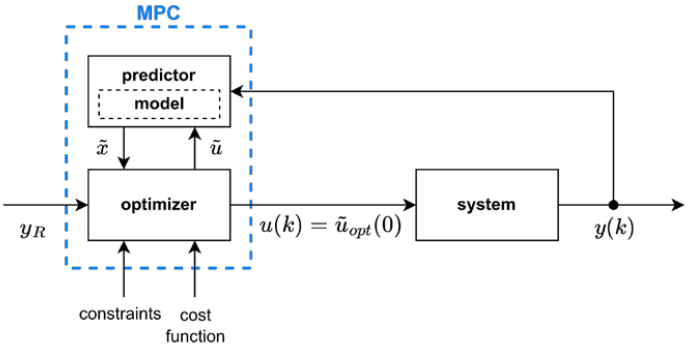


Figure 5: Traditional MPC scheme

Adaptive Linear MPC (ALMPC) on the other hand, is an extension of the traditional MPC approach, that is tailored for systems with time-varying linear dynamics. It involves an online update of the linear model used in the MPC optimization as new data becomes available. This enables the controller to adjust to changes and nonlinearities in the system dynamics.

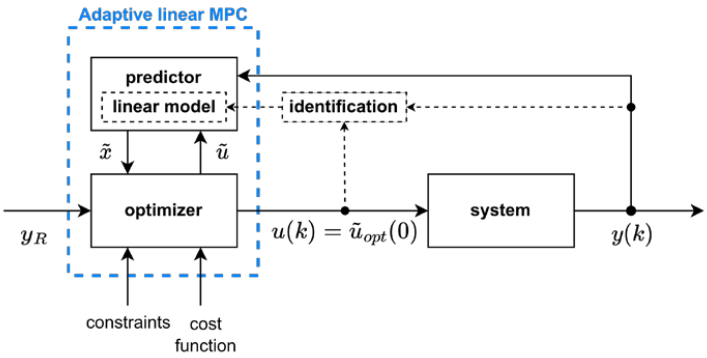


Figure 6: Adaptive linear MPC scheme

A general disadvantage of MPC is the high computational demand. It scales significantly with number of states N, number of constraints m and number of inputs n. For example, for the interior point method the complexity is $\mathcal{O}(\mathrm{N}^3 \cdot (n+m)^3)$ [4].

Additionally, the prediction horizon n_P, the control horizon n_c and the controller sample time T_s are also critical parameters in MPC frameworks. n_P limits the time frame for forecasting the system's trajectory, while n_c defines the duration over which control actions are optimized. A higher n_P leads to a more forward-looking but also increases the computational burden exponentially. Similarly, extending n_c generally results in more optimized control sequences but at the cost of higher computational demand. A smaller T_s leads to a higher computational burden but also to more accurate control behavior.

Therefore, the target is to use the least number of states, constraints, and inputs as well as the biggest control sample-time, the shortest prediction and control horizon as possible, while preserving the advantages of the optimal control strategy. To reach both, accuracy and computational viability for industrial control systems, a detailed operating strategy (reference) and a reduced operating strategy are implemented and validated in simulation.

3.2. Full order model operating strategy

Optimization problem formulation

The detailed variant is benchmarked against a state-of-the-art P-controlled system [5] (with acceleration feedback). The controller is only utilising the speed input of the pump and the swivel angle is set to maximum for both pumps.

For the strategy itself there are two important aspects, the cost function (4), and the state and input constraints (5) of the problem. The cost function consists of the following parts: control deviation

$e_{Control}$, rate of control inputs du and the total loss energy use $E_{Sys,loss}$.

$$[\mathrm{u, x, s}]^{\mathrm{T}} = \mathbf{argmin}\, \mathrm{f}_{cost}\,(\mathrm{e}_{Control}\,, \mathrm{du}, \mathrm{E}_{\mathrm{Sys,loss}}) \tag{4}$$

The constraints are set by the system boundaries. It is distinct between hard and soft constraints, whereas hard constraints (5) are mandatory to be hold. In this case minimal and maximal pump speed and minimal and maximal swivel angle are considered.

$$\begin{aligned} \omega_P - \omega_P^{[max]} &\leq 0 \\ -\omega_P - \omega_P^{[max]} &\leq 0 \\ \pm\alpha_{cmd}^{[1,2]} - 1 &\leq 0 \\ \pm\omega_{P,cmd} - \omega_P^{[max]} &\leq 0 \\ -|\omega_P| - \omega_P^{[min]} &\leq 0 \end{aligned} \tag{5}$$

Soft constraints (6) are on the pressure A and B port. Soft constraints can be violated if no feasible solution can be found. For this reason, a slack variable is introduced to weight such violations negatively in the cost function. However, the slack variables have been omitted for reasons of clarity.

$$\begin{aligned} p_K - p^{[max]} &\leq 0 \\ p_R - p^{[max]} &\leq 0 \\ -p_K + p^{[min]} &\leq 0 \\ -p_R + p^{[min]} &\leq 0 \end{aligned} \tag{6}$$

The total loss energy of the system (7) is calculated by the integral over the total loss power equation (3).

$$E_{Sys,loss} = \int P_{loss}\, dt \tag{7}$$

Results

The design and performance evaluation of the MPC algorithm is based on a reference cycle (7). The use of a reference cycle helps to evaluate the advantage of the proposed control scheme for a specific or generic task.

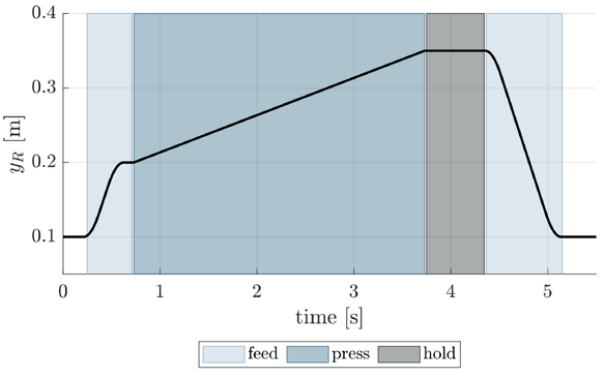


Figure 7: Reference cycle

Figure (8) shows the comparison of three different approaches: A state-of-the art p-controller (with acceleration state feedback), the ALMPC method without energy efficiency optimization and ALMPC method with energy efficiency optimization.

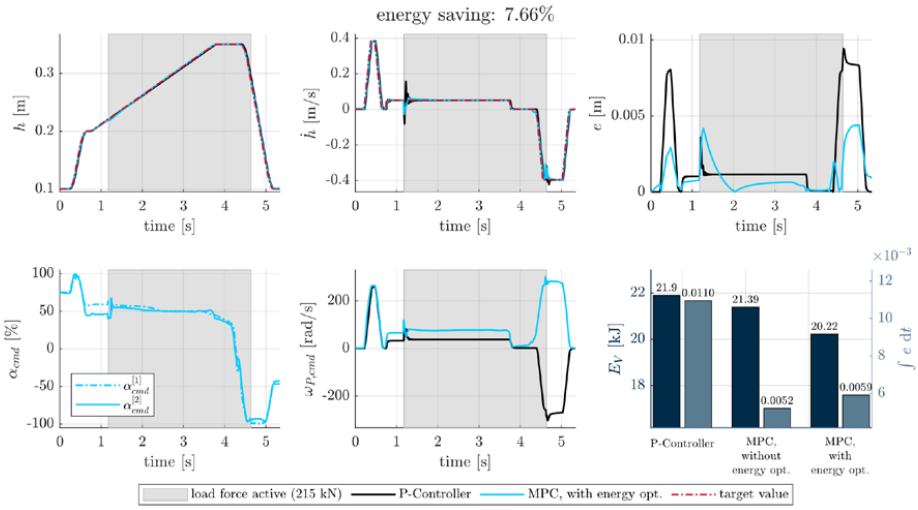


Figure 8: Result for reference cycle n_P = 40, n_C = 12 and T_S = 10ms

After a hyper parameter tuning (parameter set n_P = 40, n_C = 12 and T_S = 10 ms) the results for the benchmark cycle show that both conflicting performance indicators could be improved at the same time. For MPC without energy optimization the control accuracy is slightly better than MPC with

energy optimization. Both approaches yield a significant improvement over the P-controller benchmark for both energy usage and tracking performance.

3.3. Reduced Order Model Operating Strategy

Optimization problem formulation

To bring the MPC approach closer to the real-time applicability, computational burden must be reduced. To speed up the embedded optimization problem a reduced order prediction model, fewer constraints and a new reduced cost function is formulated. These simplifications are done while the goal is to preserve most of the results of the high order model.

$$\dot{x} = \begin{pmatrix} \dot{p}_K \\ \dot{p}_R \\ \dot{h} \\ \ddot{h} \end{pmatrix} = \begin{pmatrix} \frac{E'(p_K)}{V_K(h)} \cdot \left(-G_{Z,L}(p_R - p_K) + Q_P^{[1]}(\eta_v) + Q_P^{[2]}(\eta_v) - A_K \dot{h} \right) \\ \frac{E'(p_R)}{V_R(h)} \cdot \left(G_{Z,L}(p_R - p_K) - Q_P^{[2]}(\eta_v) + A_R \dot{h} \right) \\ \frac{\mathrm{d}}{\mathrm{d}t} h \\ \frac{1}{m} \cdot (p_K A_K - p_R A_R - F_R - F_L) \end{pmatrix}$$

Figure 9: Reduced dynamical model for the electro-hydraulic system

First, the required effort to tune the weights of the cost function is reduced considerable by moving the control deviation $e_{Control}$ out of the function and transform it into a soft constraint. Also, the weight tuning is simplified by having a reduced size weight matrices to trade off the opposing optimization objectives [3].

$$[\mathrm{u}, \mathrm{x}, \mathrm{s}]^{\mathrm{T}} = \mathbf{argmin}\, \mathrm{f}_{cost}\,(\mathrm{du}, \mathrm{E}_{\mathrm{Sys}}) \tag{8}$$

As the swash plate dynamics is removed from the model, no constraints are set on the minimum and maximum swivel angle and swivel rate. The control deviation $e_{Control}$ is now formulated as a soft constraint, which penalizes the exceeding of a limit e_{tol}. As in the first optimization problem the slack variables have been omitted for reasons of clarity.

$$\begin{aligned} e_{control} - e_{tol} &\leq 0 \\ -e_{control} - e_{tol} &\leq 0 \\ p_K - p^{[max]} &\leq 0 \\ p_R - p^{[max]} &\leq 0 \\ -p_K + p^{[min]} &\leq 0 \\ -p_R + p^{[min]} &\leq 0 \end{aligned} \tag{9}$$

While searching for controller hyperparameters, of the reduced order problem, we observed better

results for lower prediction and control horizons than for long prediction and control horizons. This can be explained by the fact that ALMPC uses linearization. Therefore, the model uncertainty of the prediction model increases with a rising distance to the operating point such that modelling errors and linearizing effects accumulate along the prediction horizon. However, this fact comes in very handy when we try to have the controller run in an embedded environment. The shorter horizons result in less computing effort.

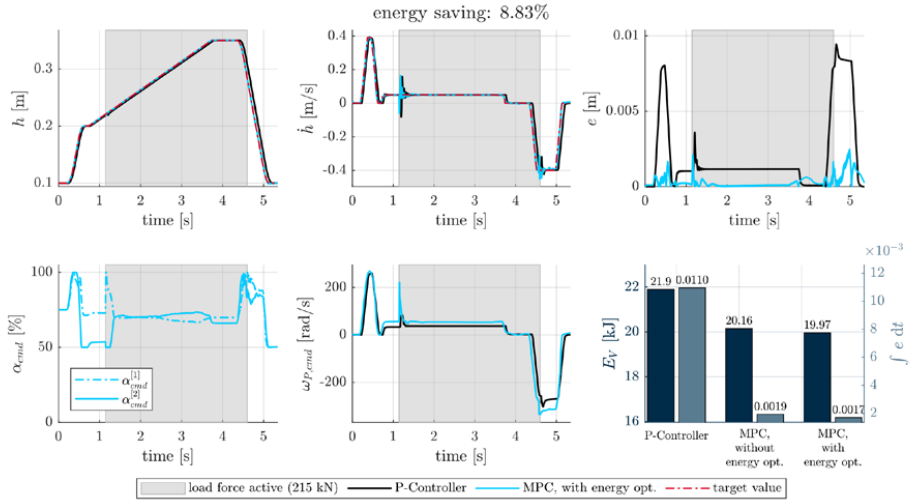


Figure 10: Result for reference cycle n_P = 12, n_C = 6 and T_S = 5ms

Noteworthy is that by reducing the problem complexity the amount of compute dropped significantly and allowed for a wider and more granular hyperparameter search. This ultimately led to hyperparameters (parameter set n_P = 12, n_C = 6 and T_S = 5 ms) which gave similar results on the benchmark cycle versus the full order model. Additionally, the reduced amount of compute brought the software into the realm of real time operation on an industrial computer.

4. CONCLUSION AND OUTLOOK

In summary, our work has demonstrated the scalability and potential of Model Predictive Control (MPC) for control systems with multiple degrees of freedom in its inputs. The optimization-based approach makes it a scalable and versatile tool for achieving different performance goals under constraints and safety margins.

Looking forward, the future of MPC in hydraulic systems is promising. Combining control algorithms with AI and machine learning will further enhance capabilities, and the demand for energy-efficient solutions underscores the importance of MPC in optimizing hydraulic operations. The competencies of hydraulic control engineers need to evolve towards dynamical modelling and optimization-based control to harness the power of discussed approaches. Collaborative efforts between academia and industry will facilitate the transition of these advancements into practical applications, benefiting numerous sectors that rely on hydraulic systems. In conclusion, MPC offers substantial benefits, and its future applications promise improved control performance and energy

efficiency in various hydraulic industries, driving innovation in engineering and automation.

REFERENCES

[1] Willkomm J (2016) Modellprädiktive Optimierung drehzahlvariabler Verstellpumpen, Fluidmechatronische Systeme, Dresdner Verein zur Förderung der Fluidtechnik e.V.

[2] Leutbecher T (2023) Optimization based control of hydraulic systems with multiple degrees of freedom, Bosch Rexroth, Universität Ulm, Technische Hochschule Ulm, Masterthesis

[3] Cecchin L (2023) Nonlinear Model Predictive Control for Efficient Control of Variable Speed Variable Displacement Pumps, Research of Robert Bosch GmbH

[4] Richter S, Jones C, Morari M (2012) Computational Complexity Certification for Real-Time MPC With Input Constraints Based on the Fast Gradient Method Research, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 57

[5] Gebhardt N, Weber J (2020) Hydraulik Fluid-Mechatronik Grundlagen, Komponenten, Systeme, Messtechnik und virtuelles Engineering

DATA-DRIVEN VIBRATION CONTROL STRATEGY FOR HYPERGRAVITY CENTRIFUGAL SHAKING TABLE

peer reviewed

Yang Zhu[1], Xie Haibo[2*]

[1] *State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, 310058, Hangzhou, Zhejiang, China*
[2] *State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, 310058, Hangzhou, Zhejiang, China*

* Corresponding author: Tel.: +86 13116788322; E-mail address: hbxie@zju.edu.cn

ABSTRACT

The hypergravity centrifugal shaking table is widely used in the field of civil engineering, which has the scaling effect and is the most effective means of studying the disaster effect of geotechnical earthquake. The compound control of hypergravity centrifugal shaking table is mainly composed of two parts: position servo control and acceleration vibration control. The position servo control is used to ensure that the shaking table works in a safe working area, and the acceleration vibration control is helpful to further improve the seismic wave reproduction ability of the shaking table. Due to the complexity and diversity of the working conditions of the shaking table, it is proved that under the premise of robust position loop control, the vibration control strategy of the acceleration ring of the hypergravity centrifugal shaking table based on the data-driven idea can effectively improve the vibration control accuracy of the hypergravity centrifugal shaking table. At the same time, the vibration control strategy is only based on the current working state of the shaking table. Therefore, It has stronger adaptive control ability.

Keywords: Shaking table, Vibration control, Data-driven, Hypergravity

1. GENERAL INSTRUCTIONS

Earthquake is one of the major natural disasters faced by mankind. It is of great significance to study the earthquake disaster effects such as building damage and sand liquefaction caused by earthquake. Shaking table is a necessary means to study the earthquake disaster effect. The seismic wave generated naturally during the earthquake disaster is a low-frequency elastic wave with a frequency below 10 Hz.

The hypergravity centrifugal shaking table [1, 2] is mainly composed of two parts: a centrifuge that generates an additional heavy force field and an airborne shaking table that is used to reproduce high-frequency seismic waveforms. By using a large centrifuge to attach the heavy force field of the airborne shaking table, the stress equivalence between the scale experiment under hypergravity and the original physical model experiment under normal gravity is realized. Under the action of n g high gravity, the size and vibration duration of the experimental model are reduced by n times, but the vibration frequency and amplitude are also increased by n times.

High-precision seismic signal reproduction is the key technology of airborne shaking table, and it is also the prerequisite for reproducing the real earthquake effect on the corresponding civil structure. The control forms of airborne shaking table are mainly divided into single servo control, vibration control or the combined control strategy of servo control and vibration control. The composite control takes into account the characteristics of good stability of servo control and high precision of vibration

control waveform reproduction, and is an effective means to improve the ability of shaking table waveform reproduction. For servo control, the common control strategy is mainly three-state control [3, 4]. Three-state feedback adjusts the zero position of the system, in which displacement feedback ensures the stability of the system, acceleration feedback improves the damping of the system, and velocity feedback improves the bandwidth of the system. Meanwhile, three-state feedforward is used to eliminate the stable zero near the virtual axis of the system and further improve the bandwidth characteristics of the system. However, different from the executive cylinder of the normal gravity shaking table, which has a long shape path and low natural bandwidth characteristics of the system, the executive cylinder of the high gravity airborne shaking table has a short and thick structure, and its natural bandwidth is high, generally up to about 200 Hz. The system bandwidth is mainly limited by the bandwidth characteristics of the system servo valve. Therefore, the three-state control of the hypergravity airborne shaking table has certain limitations. In terms of vibration control, the relevant adaptive control vibration control strategy [5] and iterative control strategy [6] are more actively studied, but their research objects are mainly aimed at the normal gravity shaking table condition, and the vibration control strategy of the airborne shaking table with hypergravity needs to be further studied and verified.

This paper mainly includes the following parts, the second part is the schematic description and the dynamic modeling of airborne shaking table, the servo control and vibration control strategy of the hypergravity airborne shaking table as the third part, the fourth part is the simulation and experimental results, and the fifth part is conclusion.

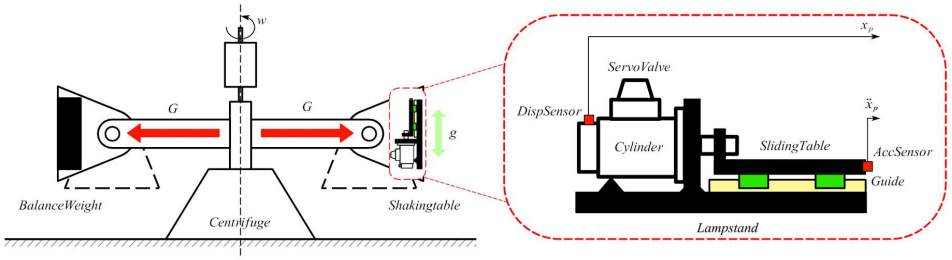


Figure 1: Hypergravity unidirectional centrifugal shaking table structure

2. SYSTEM ANALYSIS

2.1. Physical system

The hypergravity centrifugal shaking table system is shown in Figure 1, which is mainly composed of two parts: centrifuge system and unidirectional airborne shaking table.

Centrifuge rotates at a high speed under the action of the driving motor to generate centrifugal force G with a specific gravity value. The unidirectional airborne shaking table is installed in the basket on one side of the centrifuge, and the basket on the other side is the centrifuge counterweight. The direction of the vibration acceleration g of the unidirectional airborne shaking table is perpendicular to the centrifugal force G and is naturally decoupage. Under this premise, in order to facilitate the study of the hypergravity centrifugal shaking table, the unidirectional shaking table studied here is in the normal gravity environment. The study of the high-frequency vibration characteristics of the airborne shaking table in the normal gravity environment is equivalent to the study of the airborne shaking table in the actual heavy force field.

2.2. System analysis and modelling

The schematic of the airborne shaking table system is shown in Figure 2, which is a typical single-cylinder double-valve system. The purpose of the dual valve parallel [7] is to consider that the main factor affecting the bandwidth of the unidirectional airborne shaking table system is the bandwidth characteristics of the servo valve. The higher bandwidth characteristics of the servo valve often correspond to the working condition of the servo valve small opening, that is, the working condition of the servo valve small flow rate. When the double valves are in parallel, the demand flow rate of the single valve is reduced by half theoretically, so the double valves in parallel can greatly improve the bandwidth characteristics of the system.

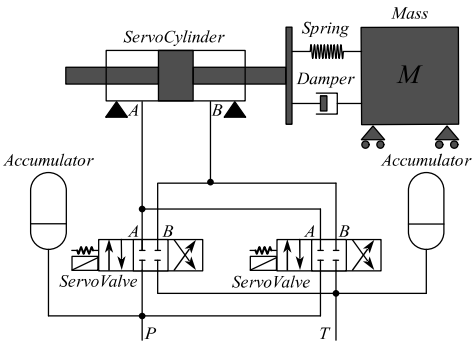


Figure 2: Schematic of unidirectional airborne shaking table

The load force balance equation for the single-cylinder double-valve system is established as shown in equation (1):

$$A_p P_L = m\frac{d^2 x_p}{dt^2} + B_p\frac{dx_p}{dt} + F_L \tag{1}$$

where, A_p is the actuating area of the cylinder, m is the mass of the piston and load converted to the load rod, x_p is displacement, B_p represents the viscous damping coefficient, F_L is the accidental load force, P_L is the load pressure. The continuity equation of the flow is established as follows (2):

$$Q_L = A_p\frac{dx_p}{dt} + \frac{V_t}{4\beta_e}\frac{dP_L}{dt} + C_{tp}P_L \tag{2}$$

where, Q_L represents the system load flow rate, V_t represents the total volume of the cylinder and the pipeline between the cylinder and the servo valve, β_e represents the effective bulk elastic modulus, the linear equation of the servo valve flow rate is shown in equation (3):

$$Q_L = K_q x_v + K_c P_L \tag{3}$$

K_q is the flow coefficient, K_c is the flow pressure coefficient of the servo valve, and x_v is the spool displacement of the servo valve. Let $K_{cy} = \frac{K_q}{A_p}$ indicate the system open-loop gain, $K_{ce} = K_c + C_{tp}$ indicate the total flow pressure coefficient, and $C_{tp} = C_{ip} + \frac{C_{ep}}{2}$ indicate the total leakage coefficient, $w_h = \sqrt{\frac{4\beta_e A_p^2}{V_t m}}$ and $\xi_h = \frac{K_{ce}}{A_p}\sqrt{\frac{\beta_e m}{V_t}}$ respectively represent system natural frequency and the hydraulic

damping ratio. Considering that the load viscous damping coefficient B_p is small, meet $\frac{K_{ce}B_p}{A_p^2} \ll 1$. Connected equations (1-3), the transfer function of single cylinder double valve can be obtained as follows (4):

$$X_p = \frac{K_{cy}}{s(\frac{s^2}{w_h^2} + \frac{2\xi_h}{w_h}s + 1)} X_v \tag{4}$$

The two-valve parallel dynamic model can be approximately simplified into a second-order oscillation system as shown in equation (5):

$$X_v = \frac{K_{sv}}{\frac{s^2}{w_{sv}^2} + \frac{2\xi_{sv}}{w_{sv}}s + 1} U \tag{5}$$

where, K_{sv} is the gain coefficient of dual valves in parallel, w_{sv} is the servo bandwidth and ξ_{sv} is the damping coefficient. The dynamic model of shaking table can be obtained by connecting equations (4-5), as shown in equation (6):

$$X_p = \frac{K_s U}{s\left(\frac{s^2}{w_h^2} + \frac{2\xi_h}{w_h}s + 1\right)(\frac{s^2}{w_{sv}^2} + \frac{2\xi_{sv}}{w_{sv}}s + 1)} \tag{6}$$

where, $K_s = K_{cy}K_{sv}$ represents the total open-loop gain of the system.

3. ALGORITHM DESIGN

3.1. Servo control strategy

Due to its inherent limitations, the single degree of freedom controller cannot meet the control requirements of tracking and disturbance rejection at the same time, and the single degree of freedom controller must make a compromise between these two performance objectives, while the two degrees of freedom controller fundamentally overcomes the inherent defects of the single degree of freedom controller, and has two degrees of freedom to handle disturbance rejection and tracking tasks separately 0. Based on the displacement feedback signal, a controller with two degrees of freedom is established as the servo controller of the airborne shaking table. While ensuring the sufficient robustness of the shaking table displacement ring, the displacement control accuracy of the shaking table is improved as much as possible, thus laying the foundation for the high-precision control of the acceleration of the shaking table. The servo control frame of the airborne shaking table is shown in Figure 3:

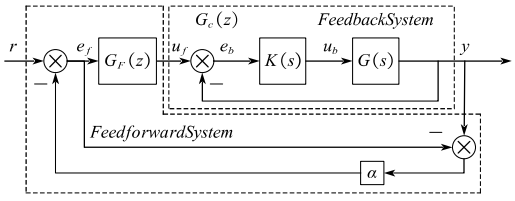


Figure 3: Servo control frame of unidirectional airborne shaking table

where, the system transfer function model G(s) and the feedback controller K(s) designed based on H_∞ method 0 constitute airborne shaking table displacement closed loop. Further, the nominal model $G_c(z^{-1})$ of the system position closed loop is determined through identification experiments. The feedforward controller $G_F(z) = G_c^{-1}(z)F(z)$ in the system is constructed by using the improved internal model control strategy 0, where $G_c^{-1}(z)$ and F(z) represent the nominal inverse model of the system 0 and the low-pass filter to prevent the system's high-frequency gain from being too large, respectively. $0 < \alpha < 1$ is the gain factor in the improved internal model controller.

3.2. Vibration control strategy

Under a single servo control link, the accuracy of acceleration seismic reproduction is often difficult to meet the application requirements. The main reasons are that firstly, the actual acceleration signal is deviated from the theoretical displacement corresponding to the corresponding acceleration, which is obtained by integrating the input servo link twice. Secondly, the vibration components of the shaking table are not completely rigid. Elastic deformation will occur and the installation and connection of each component will inevitably leave a gap. The existence of the above factors makes the acceleration signal in the high-frequency vibration of the shaking table subject to obvious harmonic interference. Therefore, only a single servo control strategy can not meet the high-precision reproduction of the acceleration waveform of the shaking table. The data-driven vibration control frame of the airborne shaking table is shown in Figure 4:

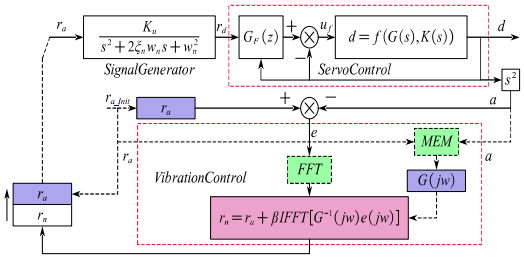


Figure 4: Vibration control frame of unidirectional airborne shaking table

where, the acceleration reference signal r_a generates the corresponding displacement signal r_d after passing through the second-order signal generator. Under the action of the servo control link, the shaking table completes the excitation experiment and collects the acceleration signal generated by the current excitation experiment at the same time. The vibration controller uses the corresponding input and output signals and uses the maximum entropy method 0, to obtain and save the frequency response of the system acceleration ring, and updates the new control input signals according to equation (7, 8):

$$r_n = r_a + \beta IFFT[G^{-1}(jw)\mathrm{e(jw)}] \tag{7}$$

$$\mathrm{e(jw)} = \mathrm{FFT}(r_a - a) \tag{8}$$

where $0 < \beta < 1$. The acceleration signal of the control input is iteratively optimized by data-driven method, and then the accuracy of the random acceleration signal of the shaking table is improved. For the airborne shaking table system, set the frequency response as G, random process input as x(n), output as y(n), and sequence length as N, then $\mathrm{G}(f) = \sqrt{\frac{P_{yy}(f)}{P_{xx}(f)}}$ in the frequency domain. $P_{yy}(f)$ and $P_{xx}(f)$ represent the self-power spectrum of the output signal and the input signal respectively. In the traditional spectral estimation method, it is assumed that when $\mathrm{n} \geq \mathrm{N}$, the autocorrelation of the signal is zero. In many cases, the above assumption is not true, so the resolution and accuracy of the

estimated spectrum will be significantly reduced by this window method. The maximum entropy method is a nonlinear spectral estimation method which can obtain high resolution by extrapolating the autocorrelation function at the known finite delay point according to the maximum entropy criterion.

It is especially suitable for spectral estimation of short data series. Based on this, the maximum entropy method is used to estimate the frequency response of the airborne shaking table acceleration ring during the implementation of vibration control strategy.

4. SIMULATION AND EXPERIMENT

4.1. System identification and simulation

The structure of the unidirectional airborne shaking table is shown in Figure 5. Under the joint action of the relief valve and the accumulator, the system works in a stable pressure environment as far as possible. The double valve parallel system can ensure that the flow supply of single valve is reduced under the same flow demand of the system, and then effectively improve the bandwidth characteristics of the airborne shaking table system. At the same time, the inertial load with a mass of 82 kg is installed on the surface of the airborne shaking table.

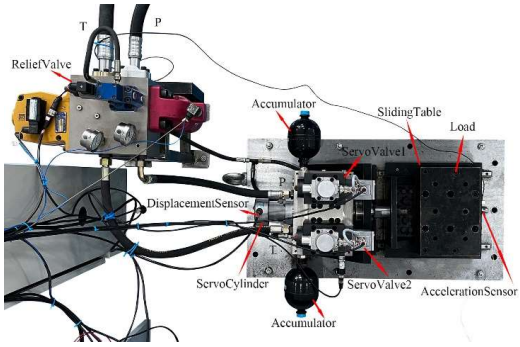


Figure 5: Mechanical structure of unidirectional airborne shaking table

Specific identification experiments are carried out on unidirectional airborne shaking table, and the transfer model of unidirectional airborne shaking table can be obtained as follows (9):

$$\hat{G}(s) = \frac{2.002e13}{s^5 + 1188s^4 + 513400s^3 + 2.362e08s^2 + 3.084e10s} \tag{9}$$

Analyze and compare the step response of the system model and the actual step response of the system, and then verify the effectiveness of the identification model. The specific comparison results are shown in Figure 6.

As can be seen from Figure 6, the 5th-order transfer function model of the system obtained by the identification experiment can better fit the step response of the actual system, and has sufficient system identification accuracy.

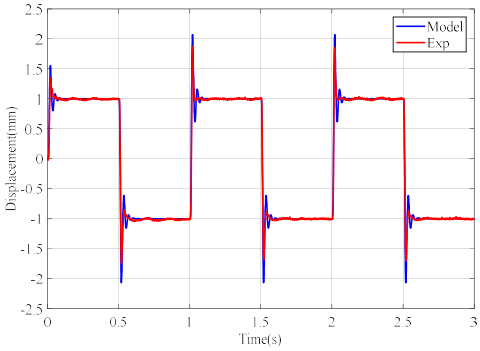


Figure 6: The open loop identification result of displacement of airborne shaking table

4.2. Experimental results

According to the displacement open-loop model, the servo controller of the airborne shaking table is designed. The displacement reference signal is obtained by integrating the corresponding random seismic acceleration signal twice. The specific expressions of the servo controller are shown in equations (10-12):

$$K(s) = \frac{\begin{matrix}806.1s^4 + 4.403\text{e}05s^3 + \\ 2.451\text{e}08s^2 + 6.718\text{e}10s \\ +6.323e9\end{matrix}}{\begin{matrix}s^5 + 1265s^4 + 1.18\text{e}06s^3 + \\ 3.952\text{e}8s^2 + 2.12\text{e}11s \\ +1.8e11\end{matrix}} \tag{10}$$

$$G_c^{-1}(z) = \frac{\begin{matrix}(z-0.9776)(z-0.7507) \\ (z^2-1.585z+0.7567) \\ (z^2+1.126z+0.751) \\ (z^2-0.5812z+0.7312) \\ (z^2+0.2856z+0.7664)\end{matrix}}{\begin{matrix}0.011534z(z-0.9807) \\ (1.494z^2+2.339z+1) \\ (1.122z^2-1.159z+1) \\ (1.299z^2+1.362z+1) \\ (1.122z^2-0.2889z+1)\end{matrix}} \tag{11}$$

$$F(z) = \frac{0.5887z + 0.2574}{z^2 - 0.2924z + 0.1085} \tag{12}$$

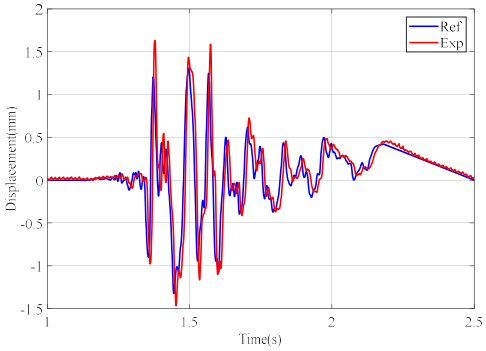


Figure 7: Input and output waveform of random seismic displacement signal on unidirectional airborne shaking table

Figure 7 shows that under the action of the servo controller, where $\alpha = 0.2$. the displacement ring of the shaking table can have a high displacement reproduction accuracy while ensuring the stability of the system. Under the premise of closed-loop displacement of the system, the input and output signals of random acceleration signal are collected, and the maximum entropy method (MEM) is used to identify the system in the frequency domain, and the corresponding system frequency response is obtained as shown in Figure 8.

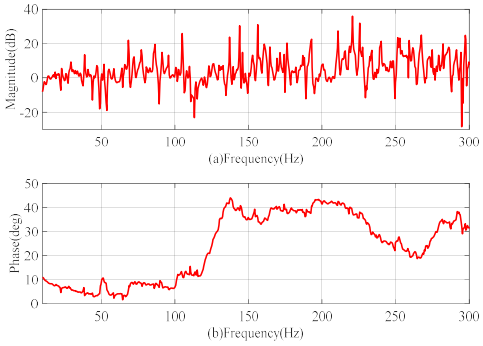


Figure 8: Frequency response of unidirectional airborne shaking table acceleration ring

In order to verify the effectiveness of the vibration strategy proposed here, the same random seismic signal is used for experiments under the two conditions of only servo control and the combination of servo and vibration control. The experimental results are shown in Figure 9.

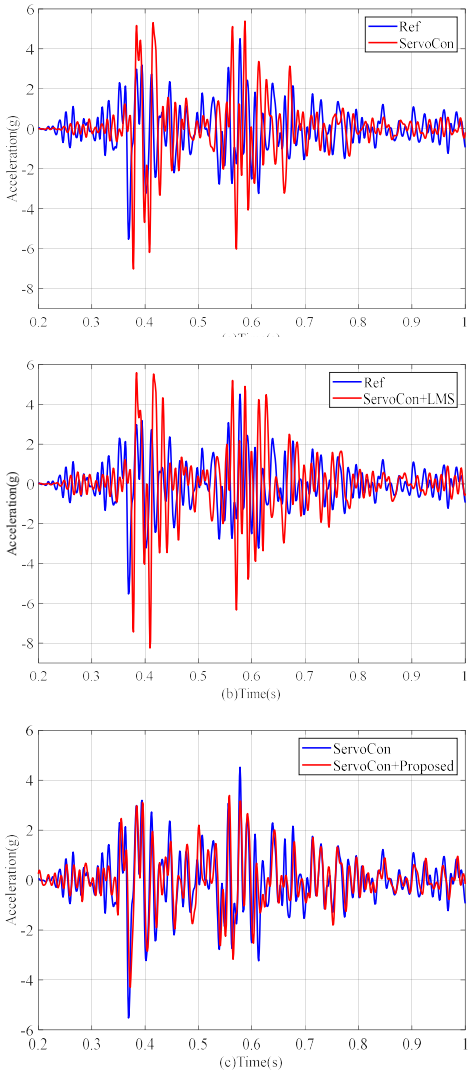


Figure 9: Input and output waveform of random seismic signal airborne shaking table

Figure 9 respectively shows the comparison between the output acceleration waveform and the reference acceleration waveform of the airborne shaking table under the same acceleration input in three cases. Figure (a): Only two degrees of freedom servo control strategy is used; Figure (b): On the basis of the servo control strategy, a vibration control strategy based on LMS 0 is further introduced, where $\mu = 0.001$; Figure (c): A data-driven vibration control strategy is introduced based on servo control measurements, where $\beta = 0.5$. Compared with the above three figures, it is found that the vibration control strategy based on data drive can obviously improve the reproduction accuracy of the acceleration waveform of the airborne shaking table under the same servo control.

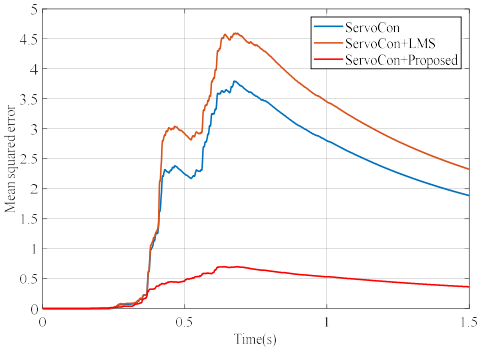


Figure 10: MSE comparison of input and output waveforms of random seismic signal airborne shaking table

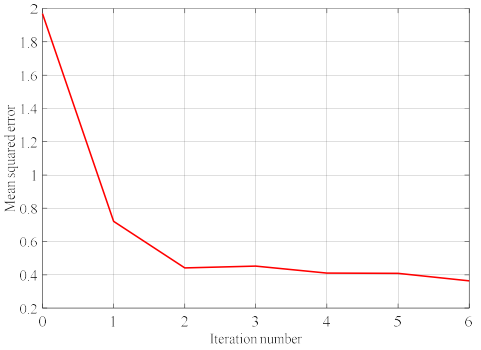


Figure 11: Global MSE iteration trend chart

The following conclusions can be drawn from the analysis of Figure 10 and Figure 11:(1) Compared with a single servo control strategy, only the combination of appropriate vibration control strategy and servo control strategy can further improve the accuracy of acceleration waveform reproduction of the shaking table; (2) Under the same servo control link, the proposed shaking table vibration control strategy, after a certain number of iterative optimization, can effectively improve the acceleration waveform reproduction accuracy of the shaking table, and it is iteratively converging.

5. CONCLUSION

In this paper, taking the hypergravity centrifugal shaking table as the research object, the problem of the poor reproduction accuracy of the acceleration waveform of the airborne shaking table under the action of a single servo controller is studied. When an airborne shaking table operates under a single servo controller, the reproduction accuracy of the shaking table acceleration waveform is often poor due to the deviation of the displacement signal obtained by the double integration of theoretical acceleration, the finite stiffness of the shaker table and the existence of installation gap. Based on this, a data-driven vibration control strategy for airborne shaking table is proposed, and the corresponding vibration control strategy is introduced into the airborne shaking table control system on the basis of servo control. The final experimental results show that the proposed vibration control strategy can effectively improve the acceleration waveform reproduction accuracy of airborne shaking table. At the same time, the proposed vibration control strategy is compared with the existing vibration optimization control strategy based on LMS, which further verifies the advanced nature of the proposed vibration control strategy.

REFERENCES

[1] Zhou, Y. G., Meng, D., Ma, Q., Ling, D. S. (2020). Frequency response function and shaking control of the ZJU-400 geotechnical centrifuge shaker. *International Journal of Physical Modelling in Geotechnics*, *20*(2), 97-117.

[2] Hung, W. Y., Liang, Y. C., Huang, J. X., & Pham, T. N. P. Effect of soil particle size on seismic response of gentle slope by centrifuge shaking table test.

[3] Qianli, L., Zhangwei, C., & Jinrong, X. (2014). Three-variable control technique for a seismic analog shaking table. *Journal of Vibration and Shock*, *33*(8), 54-60.

[4] Luan, Q. L., Chen, Z. W., Xu, J. R., & He, H. N. (2014). Three-variable control parameter tuning technology on seismic simulation shaking tables. *Journal of Vibration Engineering*, *27*(03), 416-425.

[5] Shen, G., Li, X., Zhu, Z., Tang, Y., Zhu, W., & Liu, S. (2017). Acceleration tracking control combining adaptive control and off-line compensators for six-degree-of-freedom electro-hydraulic shaking tables. *Isa Transactions*, *70*, 322-337.

[6] Tang, Y., Shen, G., Zhu, Z. C. (2014). Time waveform replication for electro-hydraulic shaking table incorporating off-line iterative learning control and modified internal model control. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, *228*(9), 722-733.

[7] Bai, Y., & Quan, L. (2016). Improving electro-hydraulic system performance by double-valve actuation. *Transactions of the Canadian Society for Mechanical Engineering*, *40*(3), 289-301.

[8] Pal, M., Sarkar, G., Barai, R. K. (2019). Two-Degree-of-Freedom Control of Non-minimum Phase Mechanical System. In *Modelling and Simulation in Science, Technology and Engineering Mathematics: Proceedings of the International Conference on Modelling and Simulation (MS-17)* (pp. 365-378). Springer International Publishing.

[9] Shiralkar, A., Kurode, S., Gore, R., & Tamhane, B. (2019). Robust output feedback control of electro-hydraulic system. *International Journal of Dynamics and Control*, *7*, 295-307.

[10] Huang, L., Pei, H., & Cheng, Z. (2023). System identification and improved internal model control for yaw of unmanned helicopter. *Asian Journal of Control*, *25*(2), 1619-1638.

[11] Rigney, B. P., Pao, L. Y., & Lawrence, D. A. (2009). Nonminimum phase dynamic inversion for settle time applications. IEEE Transactions on Control Systems Technology, 17(5), 989-1005.

[12] Tanc, A. K. (2010). Iterative maximum entropy power spectrum estimation for multirate systems. *AEU-International Journal of Electronics and Communications*, *64*(2), 93-98.

[13] Ferrante, A., Pavon, M., & Zorzi, M. (2011). A maximum entropy enhancement for a family of high-resolution spectral estimators. *IEEE Transactions on Automatic Control*, *57*(2), 318-329.

Chapter 16

System Design and Architecture-I

EFFICIENT MODEL-BASED THERMAL SIMULATION METHOD DEMONSTRATED ON A 24-TON WHEEL LOADER

Eric Pohl[1]*, Jürgen Weber[1]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*

* Corresponding author: Tel.: +49 351 463-39081; E-mail address: eric.pohl2@tu-dresden.de

ABSTRACT

With the ongoing emission reduction and electrification of mobile machinery the efficiency of the hydraulics is becoming increasingly important. Here, the thermal behavior of the entire hydraulic drive system, including all cooling circuits, plays a decisive role. For a comprehensive quantification of the thermal behavior, a model-based analysis method is the most cost- and time-efficient option. In the recent past, thermo-hydraulic network models with lumped parameters have become established for this purpose. Due to their approach of solving coupled time-dependent differential equations, very small time steps and thus very long simulation times can occur for complex systems such as mobile machines. Thus, an efficient design process including parameter studies and optimization is highly uneconomical. In this paper, an efficient computational approach for the determination of local temperatures in thermal equilibrium is presented and used for the simulation of the thermal behavior of a 24 t wheel loader. The results are simulatively validated by a comparison with the classical transient calculation approach with exactly the same parameterization in order to prove the accuracy and the reduction of simulation time. The comparison shows that the presented simulation approach is able to reduce the simulation time by more than 4000 times with almost the same accuracy. This demonstrates the great potential of the simulation method for the design process including economic optimization and parameter studies.

Keywords: mobile machines, thermal, simulation, system design, lumped-parameter model, thermal equilibrium

1. INTRODUCTION

The thermal behavior of mobile machines has a substancial influence on their design, efficiency, operability and reliability. Efficiency in particular will become increasingly important in the near future, as the trend is towards low-emission and energy-efficient systems. For example, the efficiency of electrified battery-powered machines has a direct impact on machine costs due to the high battery costs. As of yet the thermal behavior is only taken into account to a limited extent during the development and design of the different machine or drive subsystems. Design adjustments are often carried out iteratively via costly prototype measurements or the systems are greatly oversized from a thermal point of view. In order to reduce development time and costs, an accurate prediction of the thermal behavior of the entire system is required as part of the design process. Due to the complex thermal interactions within and between different domains in hydraulic drive systems, efficient simulation methods are required to calculate the local operating temperatures within the thermal equilibrium. For stationary applications thermal lumped-parameter models have been successfully established as the most efficient simulation method in recent years [1] – [3]. A vivid example of a thermo-hydraulic network model is shown in **Figure 1**.

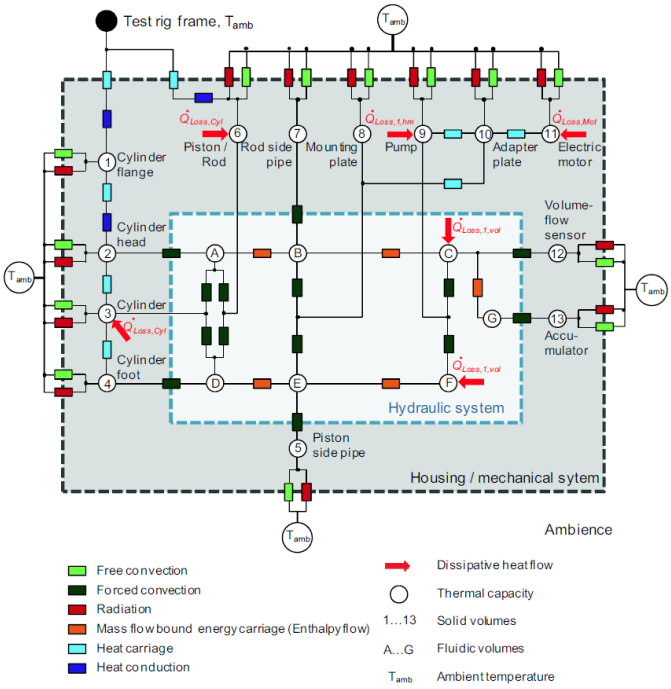


Figure 1: Exemplary thermal resistance network model of a compact drive [3]

However, in the field of mobile machines the highly application- and operation-dependent varying environmental conditions and performance requirements make it difficult to collect comparable and repeatable data for parameterization and validation compared to stationary processes. The basic feasibility of functional thermal modeling based on lumped-parameter network models of a mobile machine has been demonstrated by Zimmerman and Busquets for a displacement-controlled mini-excavator [4] – [7] and Kwon et al. for the hydraulic drivetrain of a hybrid vehicle [8] – [10]. The aforementioned works have in common that they calculate the thermal behavior using coupled time-dependent differential equations to solve conservation of mass and energy, which result from lumped-parameter network models. Therefore, the entire heating process must be calculated over countless small-scale time steps in order to finally determine local temperatures in thermal equilibrium. For systems of low complexity, results can thus be obtained in a short to moderate time. However, with increasing complexity of the system, this computational approach can lead to simulation times of many hours up to days or weeks due to very small time steps. In addition, this is scaled by the stiffness of the system and the dynamics of the process which can drastically affect the stability of the thermal-hydraulic simulation and lead to even longer simulation times or crashes. For complex systems, the development and optimization with this calculation approach is therefore only possible with great simplifications or very high expenditure of time. However, manufacturers want to model their machines as accurately as possible and have short simulation times for extensive optimizations and adjustments. Therefore, for an economically efficient simulation of the thermal behavior of the entire machine, it is desirable to develop a calculation approach that significantly reduces the simulation time without losing accuracy.

The following paper presents an efficient simulation approach for thermal lumped parameter models with dynamic loads. It directly determines the local temperatures within the state of thermal equilibrium, taking into account the dynamic work process, without calculating the dynamic heating process. Following the statistical analysis of performance using an elementary hydraulic circuit in

[11], the methodology is now applied to a complex network model of a 24 t wheel loader. In order to demonstrate the potential of the approach, the simulation results and times are compared with the classic transient calculation approach.

2. SIMULATION APPROACH

For the design of mobile machines, knowledge of the transient behavior of component and fluid temperatures, in particular the heating process, is not necessary in most cases. One exception here is the improvement in cold start behavior, for example. To optimize efficiency, operability and reliability only temperatures at which the machine operates in the thermal equilibrium range are taken into account. Therefore, a computational approach for thermo-hydraulic network models has been developed to directly determine the temperatures in thermal equilibrium without time-consuming solution of time-dependent differential equations of energy conservation due to varying temperatures (**Figure 2**).

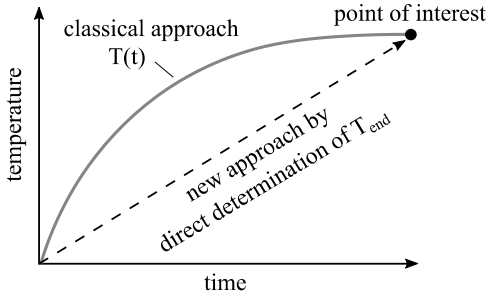


Figure 2: Developed steady-state approach compared to classical approach

In contrast to the classical transient simulation, the fluidic domain and thermal domain are decoupled. The fluidic domain is simulated using a functional-energetic network model, as is also used to calculate the kinematics of a mobile machine. The thermal domain is then calculated separately, using the same thermal parameterization as in classical coupled thermo-hydraulic network models. To determine process parameters and fluid properties, a considered duty cycle (dc) only needs to be simulated once instead of several times using the functional-energetic model (fluidic domain). This is carried out using the system simulation software SimulationX®. The time-dependent characteristics of parameters and properties are then used as boundary conditions for a steady-state calculation to determine the thermal equilibrium, which is realised with MATLAB®.

2.1. Derivation of the thermal model

The objective of the thermal model is to determine the steady-state operating temperatures of any system for an arbitrary repeating cycle. This state is called thermodynamic (or thermal) equilibrium and is defined by the 1st law of thermodynamics. The condition is reached when the internal energy U no longer changes over time:

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \sum_i \dot{Q}_i + \sum_j \dot{W}_j + \sum_k \dot{H}_{in,k} + \sum_l \dot{H}_{out,l} = 0 \tag{1}$$

with $\dot{Q}$ as external heat flow, $\dot{W}$ as technical power and $\dot{H}$ as enthalpy flow into and out of the system. This also means that all temperatures within the system must be constant over the change in the time step dt. **Figure 3** illustrates an established state of equilibrium for an exemplary periodically repeated duty cycle with a defined cycle time t_{dc}. The state is reached after about 3000 s. Looking at the

behavior of the temperature curve for a single duty cycle within the equilibrium, it can be seen that the temperatures are still varying.

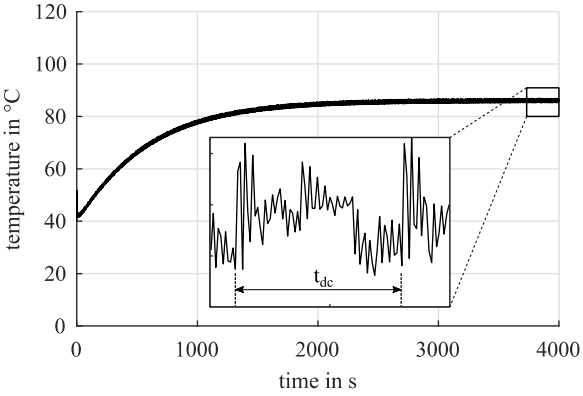


Figure 3: Dynamic temperature behaviour for an exemplary duty cycle

This is due to the fact that the state of equilibrium is never reached at any time t within the cycle, if the considered time step in (1) is $\mathrm{d}t \ll t_{dc}$. This leads to the conclusion that the thermal equilibrium is a cycle averaged consideration of temperatures and energy over several cycles with $\mathrm{D}t = t_{dc}$. Consequently, the 1st law of thermodynamics can be expressed by process averaged quantities.

$$\frac{\mathrm{D}\overline{U}}{\mathrm{D}t} = \sum_i \bar{\dot{Q}}_i + \sum_j \overline{\dot{W}_j} + \sum_k \bar{\dot{H}}_{in,k} + \sum_l \bar{\dot{H}}_{out,l} = 0 \tag{2}$$

The previously transient duty cycle is therefore considered as steady-state. This implies that there are process averaged temperatures (not time averaged) belonging to the quantities in (2). For a correct averaging of those quantities and solving of the local process averaged temperatures the differentiation of physical state variables have to be taken into account. They are divided into extensive and intensive quantities. Extensive quantities, like mass or energy, can be averaged directly over time due to the applicable conservation laws. Intensive quantities, like temperature or density, are always directly related to an extensive variable to calculate another extensive quantity. For a better understanding the heat term of the averaged enthalpy flow $\bar{\dot{H}}$ is given here as an example, which is calculated by the product of the constant specific heat capacity c_p, the mass flow $\dot{m}$ and the fluid temperature T_f.

$$\bar{\dot{H}} = \frac{1}{t_0 - t_{end}} \cdot \int_{t_0}^{t_{end}} c_p \cdot \dot{m}(t) \cdot T_f(t)\, \mathrm{d}t \tag{3}$$

Since $\dot{m}$ and T_f are directly related to each other and varying over time with $\mathrm{d}t \ll t_{end} - t_0$, only integration of the product is permissible according to the laws of physics. However, following the aforementioned assumption that there are process averaged temperatures, (3) can be simplified to

$$\bar{\dot{H}} = \frac{\tilde{T}_f}{t_0 - t_{end}} \cdot \int_{t_0}^{t_{end}} c_p \cdot \dot{m}(t)\, \mathrm{d}t \tag{4}$$

by introducing a process averaged fluid temperature $\tilde{T}_f$. For the presented calculation approach this assumption is only applied to temperatures. All other intensive quantities are not assumed as average

and are determined by the transient functional simulation model. A detailed description of the determination of all necessary variables can be found in [11]. Through the described procedure the thermal steady state can be represented by averaged variables depending on process-averaged temperatures.

Since any number of local temperatures and heat emissions can be defined, the sole application of (2) is not sufficient to determine them all, because it only describes the global thermodynamic behavior of the entire system. Therefore, it is necessary to define a set of equations which are describing the local behavior of any points in the system. The first two equations are derived by the local heat transfer at any given points in the system.

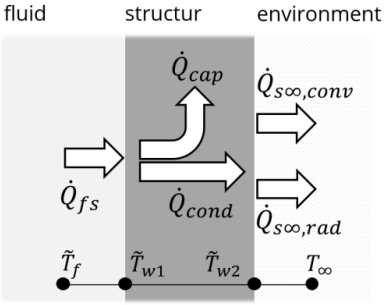


Figure 4: Local heat transfer out of the system

Figure 4 shows the heat transfer out of the fluid into the mechanical structure and out of this structure into the environment for a consideration with lumped parameters. In the heating phase the convective heat flow out of the fluid $\dot{Q}_{fs}$ divides into two parts, the structure heating flow $\dot{Q}_{cap}$ and the thermal conduction $\dot{Q}_{cond}$. When it comes to the emission into environment, the latter is divided into radiation and convection. In thermodynamic equilibrium $\dot{Q}_{cap}$ becomes zero in a process averaged consideration, as the average temperature of the structure is constant. Thus, the local heat transfer out of the system for any point i within the thermal steady state can be described by the following relationships.

$$\bar{\dot{Q}}_{fs,i}(\tilde{T}_{f,i}, \tilde{T}_{w1,i}) = \bar{\dot{Q}}_{cond,i}(\tilde{T}_{w1,i}, \tilde{T}_{w2,i}) \tag{5}$$

$$\bar{\dot{Q}}_{cond,i}(\tilde{T}_{w1,i}, \tilde{T}_{w2,i}) = \bar{\dot{Q}}_{s\infty,conv,i}(\tilde{T}_{w2,i}, T_\infty) + \bar{\dot{Q}}_{s\infty,rad,i}(\tilde{T}_{w2,i}, T_\infty) \tag{6}$$

This defines all energy flows out of the system and the local wall temperatures can be determined using the local fluid temperatures. However, the actual thermodynamic state of the latter is still undefined. In order to be able to calculate the local fluid temperatures, the causal relationships within the considered system must be known. These are described by the enthalpy flow, or rather by the change in enthalpy flow between two points under consideration (**Figure 5**).

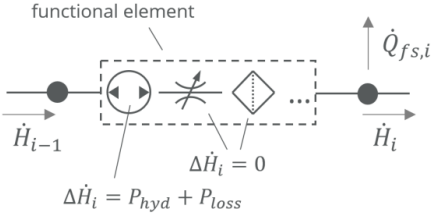


Figure 5: Change of state of the fluid between two considered points

In principle, the enthalpy only changes through the supply or removal of energy across the system boundary of the fluidic domain. As there are always functional elements such as pumps or valves between these points, there is always a change in enthalpy, as heat is dissipated into the mechanical structure. Pumps, for example, transfer the mechanical energy of the shaft into the fluid and thus feed energy into the fluidic domain. A pressure loss within the fluidic domain, as in a valve or a pipe, does not lead to a change in enthalpy. This merely shifts energy from the volume work to the heat portion of the enthalpy. Therefore, the causal relationship between two points under consideration is described by:

$$\bar{\dot{H}}_i\left(\tilde{T}_{f,i}\right) = \bar{\dot{H}}_{i-1}\left(\tilde{T}_{f,i-1}\right) + \Delta\bar{\dot{H}}_i - \bar{\dot{Q}}_{fs,i}\left(\tilde{T}_{f,i}, \tilde{T}_{w1,i}\right) \tag{7}$$

This means that the state of the previous point in the system must be known. Therefore, the determination of all local temperatures within the thermal equilibrium results in a matrix of all considered points, described by (5), (6) and (7), as well as the global condition in (2). For an example of a hydraulic circuit with i considered points and none enthalpy flow into or out of the system, the following set of equations results.

global:
$$\frac{\mathrm{D}\bar{U}}{\mathrm{D}t} = \sum_i \bar{\dot{Q}}_i + \sum_j \overline{\dot{W}_j} = 0$$

local:
$$\begin{pmatrix} \bar{\dot{Q}}_{fs,1} = \bar{\dot{Q}}_{cond,1} & \dots & \bar{\dot{Q}}_{fs,i} = \bar{\dot{Q}}_{cond,i} \\ \bar{\dot{Q}}_{cond,1} = \bar{\dot{Q}}_{s\infty,conv,1} + \bar{\dot{Q}}_{s\infty,rad,1} & \dots & \bar{\dot{Q}}_{cond,i} = \bar{\dot{Q}}_{s\infty,conv,i} + \bar{\dot{Q}}_{s\infty,rad,i} \\ \bar{\dot{H}}_1 = \bar{\dot{H}}_i + \Delta\bar{\dot{H}}_1 - \bar{\dot{Q}}_{fs,1} & \dots & \bar{\dot{H}}_i = \bar{\dot{H}}_{i-1} + \Delta\bar{\dot{H}}_i - \bar{\dot{Q}}_{fs,i} \end{pmatrix}$$

2.2. Determination of process parameters

In order to receive all necessary transient data to perform a physically correct process averaging to solve the resulting thermal matrix a functional-energetic simulation is needed. The characteristic of this model is basically the same as for kinematic simulations. There are no additional thermal elements necessary and a constant global fluid temperature is used for the calculation. If the considered system is exactly initialised only one duty cycle has to be calculated for the determination of the required process data. To avoid initialisation errors, the cycle can be simulated twice and the data extracted from the second cycle. The data determined in this way are purely functional parameters such as power $\vec{P}(t)$, pressure $\vec{p}(t)$ or mass flow $\vec{\dot{m}}(t)$. For the derivation of the actual fluid parameters, like the density $\vec{\rho}(t)$, a map of the fluid $\rho\left(p, T_f\right)$ is used. The actual benefit of this type of determination becomes apparent in the following section. In combination with the previous described thermal model the work flow of the developed approach shown in **Figure 6** results.

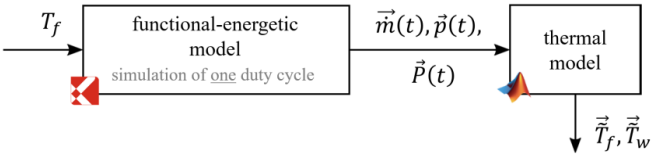


Figure 6: Initial model by combining the functional and thermal model

As the combined model can be initialised with any fluid temperature, the resulting process averaged temperatures within the thermal equilibrium are calculated by the process quantities corresponding

to this initial temperature. This leads to incorrect power values and power losses in the system and thus to incorrect temperatures calculated by the thermal model. In fact they have to be determined by the process parameters corresponding to the resulting temperatures. Therefore, a correction loop needs to be implemented. Thereby, the fluid temperature calculated by the thermal model is set as the new initial temperature for the next step, i.e. the following functional-energetic simulation of the single duty cycle. The loop ends when a certain relative deviation δ between the initial and newly determined fluid temperature is fallen short of, for example $\delta < 0.1$ %. In summary, the developed calculation approach works as shown in **Figure 7**.

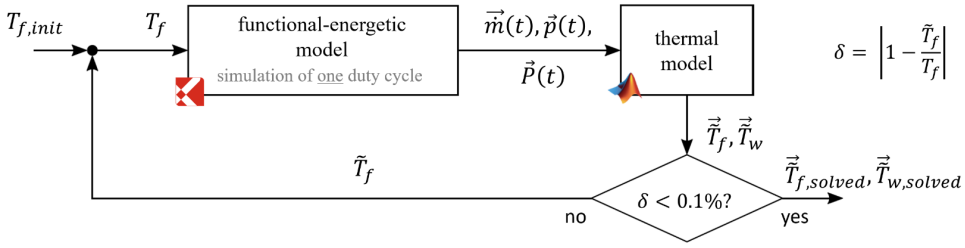


Figure 7: Resulting model by correct determination of fluid and process parameters

As already mentioned, the fluid parameters are derived by a map of the fluid $f(p, T_f)$ for each point in the system under consideration. When using the initial combined model, this is obsolete as only the initial fluid temperature can be used. The correction loop now uses the simulated fluid temperatures $\vec{\tilde{T}}_f$ to calculate the fluid parameters in the following run of the loop. This guarantees a solid solution of these quantities if the deviation δ is small enough.

3. SETUP DEMONSTRATION SYSTEM

The realisation and testing of the methodology just described is carried out on a 24 t wheel loader. This machine was prototypically build as part of the TEAM-project [12] with the objective to analyse and adapt the system structure at a functional level in order to improve the overall efficiency of the machine. The schematic system structure of the wheel loader is shown in **Figure 8**.

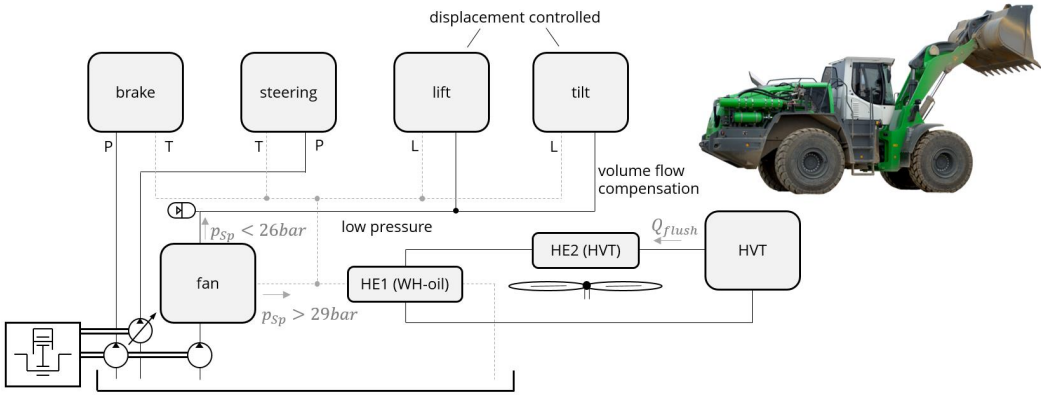


Figure 8: Demonstration object used for the methodology introduced

The test object has special functional features compared to conventional series machines. Firstly, the combustion engine has been optimised in terms of its efficiency for its operating range. The travel drive of the wheel loader is realised via a hydro-mechanical variable transmission (HVT). Furthermore, the working hydraulics (WH) is realised by a displacement controlled system. The

lifting and tilting functions are controlled by separate pumps. **Figure 8** also shows that there is a separate hydraulic circuit for the HVT and a separate hydraulic circuit for the working hydraulics and auxiliary units. The only connection between the circuits is through thermal interaction via the oil-oil heat exchanger HE1. Here, the oil from the working hydraulics is cooled by the oil from the transmission circuit. The transmission oil itself is previously cooled via the air-oil heat exchanger HE2 and circulates in a closed circuit. The axial fan used for this is driven by a hydraulic motor via the fan subsystem. The pumps of the brake, steering and fan subsystems draw hydraulic oil from the reservoir. In the fan subsystem, the entire pump volume flow is used to charge the hydraulic accumulator when the accumulator pressure falls below 26 bar. If the threshold value is exceeded, the volume flow is routed into the reservoir via HE1. The accumulator branch is used for the supply and removal of volume flow for the closed displacement-controlled circuits of the lifting and tilting subsystems. Only the leakage from the pump units is passed on to the reservoir via the heat exchanger. Due to the system complexity described, which includes both valve- and displacement-controlled structures and their complex interaction, this demonstrator is particularly suitable for demonstrating the simulation methodology developed.

4. MODELLING AND RESULTS

For a comprehensive comparison between the classical transient and the developed steady-state calculation method, the thermo-fluid network model was first set up for the classical approach. For this purpose, the existing functional-energetic model was adapted with regard to a thermo-fluidic simulation and extended by thermal elements.

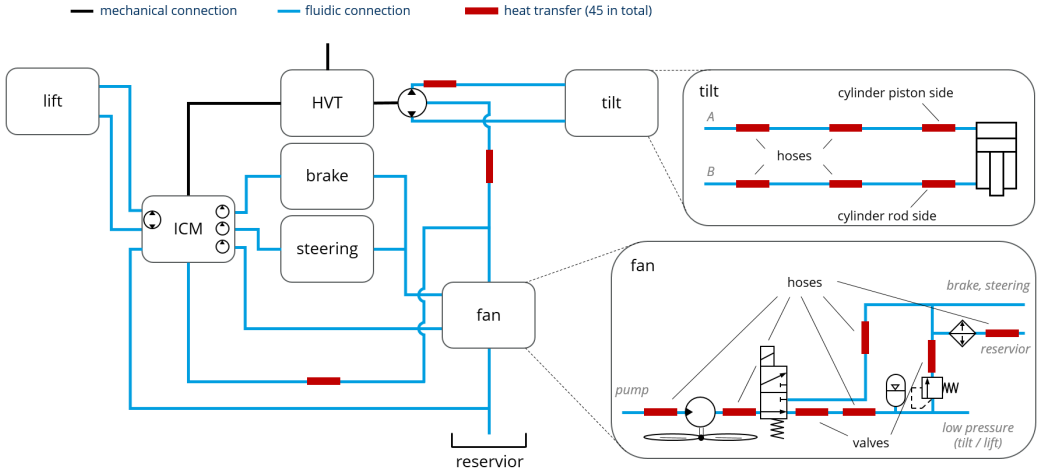


Figure 9: Simplified scheme of the thermo-fluidic model with exemplary subsystems

Figure 9 illustrates the simplified scheme of the complex system. For a better understanding, two exemplary subsystems with their thermal discretisation are shown. For each discretised point, i.e. thermal element, the fluid and wall temperatures are calculated. In order to reduce the size of the model and thus improve the performance, some elements are combined, e.g. hoses connected in series without thermally relevant elements in between. The entire system consists of a total of 45 thermal points that are analysed. In each point the heat transfer is modeled via analytical approaches and empirical correlations, which are commonly used in lumped parameter network models and described in detail in [13]. The heat transfers were validated by loading the individual subsystems separately until the local thermal equilibrium was reached. The results of the validation can be found in [14]. A complete validation and comparison to the measured results of the entire machine is not possible due

to the lack of loss maps for the pumps, which are absolutely essential for this. Furthermore, the machine cannot yet perform any cycles automatically, which is another constraint for correct validation. For this reason, only the two simulation methods are compared. For this purpose, the validated heat transfers and the use of constant pump efficiencies are sufficient.

For the calculation approach developed, the functional-energetic simulation model for determining the dynamic process quantities is the same as the model mentioned for the classical approach, with the exception of the thermal elements. These are modelled with exactly the same parameterisation in the decoupled thermal model.

To further reduce the calculation time of the transient model, a simplified cycle was used in which the vehicle is stationary and only the lifting and tilting functions are used. Nevertheless, all other subsystems also heat up due to the interconnected system structure.

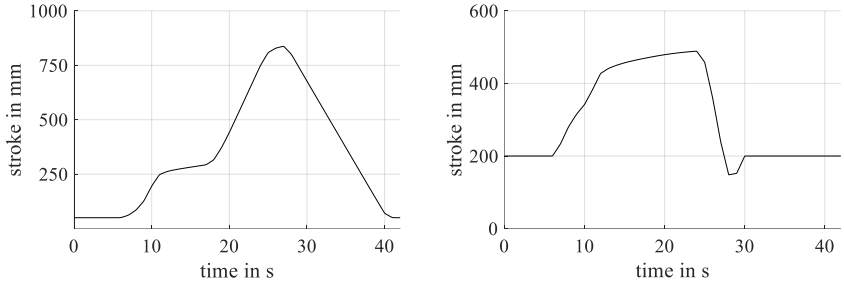


Figure 10: Trajectories of lifting (l.) and tilting (r.) cylinders of the duty cycle used

The simplified duty cycle used takes 42 s and represents a typical motion of the working hydraulics of a wheel loader, such as during a y-cycle (**Figure 10**). Using the cycle shown, the simulations are now carried out with both methods and the results are compared with each other.

Since the objective of the steady-state approach is the reduction of simulation time while maintaining accuracy, the latter must first be proven. Therefore, the fluid temperatures of all 45 thermal points indicated in **Figure 9** are compared for the two simulations methods (**Figure 11**). Looking at the quantitative differences, it can be seen that the absolute deviation between the temperatures calculated is less than 2 K in most cases. Only the displacement controlled subsystems (lift and tilt) show higher deviations of up to 4 K.

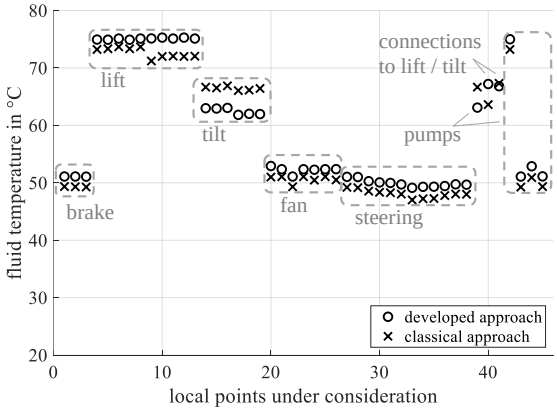


Figure 11: Comparison of calculated fluid temperatures for all points under consideration

In addition to the small quantitative deviations, the qualitative distribution of the temperatures also shows very good correlation. The only exception to this is the lift subsystem. This is due to the fact that the power losses of the pump in a displacement-controlled system vary in magnitude in the respective flow direction, which is not yet taken into account in the thermal model. The same effect occurs in the tilt subsystem, but is much smaller due to the lower difference in magnitude of the power losses and the much shorter process duration when tipping the bucket. To summarise, it can be said that both the quantitative and qualitative differences are acceptably small. However, there is still potential to improve the accuracy of the developed methodology.

Once the acceptable accuracy has been proven, the focus can now be placed on the reduction of simulation time. Due to the complexity of the system and the high level of detail in terms of the large number of thermal heat transfer elements, the presented model is enormously computationally complex. In addition, due to stability problems with the classical transient method, some fluid capacities had to be greatly increased in order to be able to dampen dynamic effects sufficiently. Although this leads to a shorter computing time of a cycle, it also causes a delayed behavior of the temperatures for reaching thermal equilibrium. This means that a single cycle can be calculated faster, but significantly more cycles must be simulated to reach steady state. To ensure best possible comparability the fluid capacities were also adjusted in the functional-energetic model of the developed method.

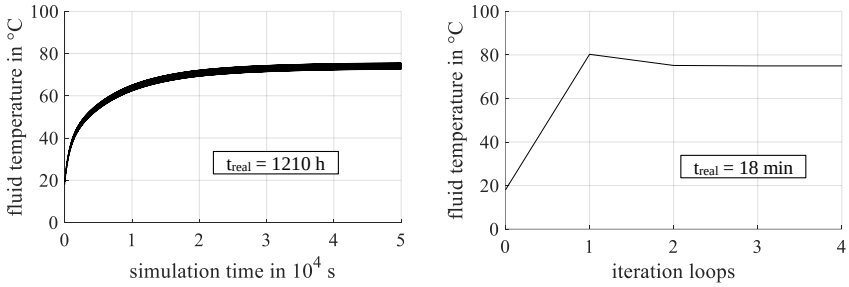


Figure 12: Comparison of computational times of the classical (l.) and the steady-state (r.) approach

The transient heating process of a point within the lifting subsystem calculated by the classical approach is shown in the left diagram of **Figure 12**. The thermal equilibrium is reached after approximately $5 \cdot 10^4$ s, which corresponds to a real time of around 1210 hours. In contrast, the iteration process for the same point calculated by the developed method is shown on the right diagram. The decoupled simulation requires 4 iterations, e.g. 3 correction loops, until the fluid temperature in the thermodynamic steady state is determined. The entire simulation with the decoupled steady-state approach only takes about 18 minutes. Compared to the classical approach, this means that it is more than 4000 times faster. This shows that the new methodology can drastically reduce the simulation time of complex hydraulic systems. This comparatively very short calculation time with acceptable accuracy opens up completely new possibilities in the model-based design and optimisation of mobile machines and furthermore for any hydraulic system.

5. CONCLUSION AND OUTLOOK

In this paper a simulation approach is presented, which reduces the dynamic heating process of a 24 ton wheel loader to the computation of the steady-state final state using functional dynamic quantities as boundary conditions. This is based on the approach that the thermodynamic equilibrium is a process-averaged consideration. Using the analytical and empirical equations of heat transfer together with state changes of enthalpy flow and functional dynamic process quantities allows an iterative calculation of process-averaged temperatures within this equilibrium. A comparison with the classical

approach, i.e. using time-dependent differential equations to solve mass and energy conservation, showed a substantial reduction in simulation time by a factor of more than 4000. The proof of only minor qualitative and quantitative deviations from the results of the classical approach underlines that the presented approach represents an efficient methodology for the design and optimisation of mobile machines. Furthermore, it can be applied to any hydraulic system as long as only operating temperatures are relevant. To further improve the approach shown, future work will focus on improving the accuracy of displacement-controlled systems.

ACKNOWLEDGEMENT

The project “Analysis, modeling and consideration of thermal interactions in the design and optimization of drive systems for mobile machinery” (Ref. No. AiF 20993 BR/1) was financed and supervised by the Forschungskuratorium Maschinenbau e.V. – FKM, Lyoner Straße 18, 60528 Frankfurt am Main. In the scope of the Program to promote Industrial Collective Research it was funded by the German Federation of Industrial Research Associations (AiF) with means of the Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag.

NOMENCLATURE

c_p	Specific heat capacity	J/(kg·K)
$\dot{H}$	Enthalpy flow	W
$\dot{m}$	Mass flow	kg/s
P	Power	W
p	Pressure	bar
$\dot{Q}$	Heat flow	W
T	Temperature	°C
t	Time	s
U	Internal Energy	J
$\dot{W}$	Technical power	W
δ	Relative threshold	
ρ	Density	kg/m³
HE	Heat exchanger	
HVT	Hydro-mechanical transmission	
ICM	Internal combustion machine	
WH	Working hydraulics	

REFERENCES

[1] Sidders J A, Tilley D G, Chapple P J (1996) Thermal-hydraulic performance prediction in fluid power systems. Part I: Journal of Systems and Control, vol. 49, pp. 231-242

[2] Johansson B, Ölvander J, Krus P (2001) Thermal Modelling of an electrohydrostatic actuation system. In Proceedings of Recent advances in aerospace actuation systems and components, Toulouse, France, pp. 49-54

[3] Michel S, Schulze T, Weber J (2014) Energy-efficiency and thermo energetic behaviour of electrohydraulic compact drives. In Proceedings of the 9th International Fluid Power Conference (IFK), Aachen, Germany, pp. 162-177

[4] Zimmerman J, Ivantysynova M (2010) Reduction of Engine and Cooling Power by Displacement Control. In Proceedings of the 6th FPNI PhD Symposium, West Lafayette, USA, pp. 339-352

[5] Zimmerman J, Ivantysynova M (2011) 40% Fuel Savings by Displacement Control leads to lower Working Temperatures - A Simulation Study and Measurements. In Proceedings of the 52nd National Conference on Fluid Power

[6] Busquets E, Ivantysynova M (2012) Cooling power reduction of displacement controlled multi-actuator machines. In Proceedings of the 7th FPNI PhD Symposium on Fluid Power, Regio Emilia, Italy, pp. 453-465

[7] Busquets E, Ivantysynova M (2013) Temperature Prediction of Displacement Controlled Multi-Actuator Machines. International Journal of Fluid Power, vol. 14, pp. 25-36

[8] Kwon H, Ivantysynova M (2017) System and Thermal Modeling for a Novel On-Road Hydraulic Hybrid Vehicle by Comparison With Measurements in the Vehicle. ASME/BATH 2017 Symposium on Fluid Power and Motion Control, Bath, UK

[9] Kwon H, Keller N, Ivantysynova M (2018) Thermal Management of Open and Closed Circuit Hydraulic Hybrids. In Proceedings of the 11th International Fluid Power Conference (IFK), Aachen, Germany

[10] Kwon H, Sprengel M, Ivantysynova M (2016) Thermal modeling of a hydraulic hybrid vehicle transmission based on thermodynamic analysis. Energy, vol. 116, pp. 650-660

[11] Pohl E, Osterland S, Weber J (2022) An efficient simulation approach to determine the thermal equilibrium of drive systems in mobile machinery. IEEE Global Fluid Power Society PhD Symposium, Neapel, Italy

[12] Schneider M (2017) Effizienzsteigerung mobiler Arbeitsmaschinen durch vollvariable, 4-quadrantenfähige Antriebe und eine bedienerentkoppelte Betriebsstrategie. Dissertation, Technische Universität Dresden

[13] Verein deutscher Ingenieure – VDI-Gesellschaft Verfahrenstechnik und Chemieingenieurswesen (2013) VDI-Wärmeatlas, Springer-Vieweg, Berlin

[14] Pohl E (2023) Analyse, Modellierung und Berücksichtigung der thermischen Wirkzusammenhänge bei der Auslegung und Optimierung von Antriebssystemen mobiler Arbeitsmaschinen. Final report (Collaborative Industrial Research, Ref. No.: 20993 BR/1), Technische Universität Dresden

A HYDRO-MECHANIC VIBRATION ABSORBER WITH ADJUSTABLE OPERATING FREQUENCY

peer reviewed

Helmut Kogler*, Bernd Winkler, Andreas Plöckinger

Linz Center of Mechatronics GmbH, Altenberger Straße 69, 4040 Linz, Austria
*Corresponding author: Tel.: +43 732 2468 6059; E-mail address: helmut.kogler@lcm.at

ABSTRACT

Unwanted resonances in industrial plants often result in unpleasant noise or may even end up in a reduced lifetime of mechanical structures and components. Such vibrations are often reduced by well-established vibration absorbers, which are designed and precisely tuned to one specific frequency. Furthermore, often the natural frequency of, for instance, a mechanical structure is not accurately known and must be identified by an experimental modal analysis. Another limitation of conventional single-frequency absorbers is that in some applications the frequency of the unwanted vibrations varies over lifetime, which reduces the effectiveness of the initially designed absorber. Moreover, in many cases the frequency of the undesirable vibration changes with the operating point of the plant, which requires the installation of multiple vibration absorbers with different operating frequencies. In this paper a new hydro-mechanical vibration absorber with a continuously adjustable operating frequency is presented, which eliminates the shortcomings mentioned above. In contrast to conventional absorbers the stiffness of the hydro-mechanical spring mass absorber is realized by a gas-loaded hydraulic capacitance, which depends on the mean operating pressure. The concept of the presented absorber is investigated theoretically and validated by simulation experiments. The results are discussed in detail and an outlook on further steps in development is provided.

Keywords: vibration, resonance, mitigation, adaptive, hydraulic, absorber

1. INTRODUCTION

Most industrial processes are characterised by a cyclic operation. The inevitable reaction forces often result in generally unwanted vibrations in the basements, suspensions or housings of the plant. The vibrations typically appear in the range from below one cycle per second up to several hundred Hertz, like for instance due to unbalance effects in high speed rotors. In some cases the vibrations may act as an excitation for other dynamical mechanical systems located close to the actual working process. Furthermore low frequency resonances of buildings and bridges may result from broad band excitation effects like earthquakes or wind gusts. If the frequency of the excitation is close to or even coincide with the natural frequency of an exposed dynamic system, then the vibration may cause a tremendous resonance effect, which may lead to either undesirable noise or even to a reduced life time of the machine or structure. There even exist a few cases, where some specific operating points of a machine result in a structural instability, which may end up in the worst case of a destruction of the plant.

As a counter measure against such unwanted resonance effects so-called vibration absorbers are widely used, which basically represent an oscillating system acting against the resonance of a vibrating primary system or structure, see for instance [1, 2, 3, 4, 5, 6]. Most vibration absorbers are able to operate at a single frequency and, thus, are designed for one specific working point. Therefore, the natural frequency of the absorber must fulfill certain conditions for the desired suppression of the vibrations. For many decades researchers developed numerous strategies for an optimized absorber

design. The most relevant methodology often used as a reference is presented in [7]. Other publications from recent years dealing with design optimization can be found, for instance, in [8, 9, 10], where in the latter also energy harvesting with a vibration absorber is discussed.

In fact, the natural frequency of the primary system can be identified accurately by more or less laborious measurements, however, the perfect tuning can often only achieved by an absorber with variable frequency. For instance, in [11] an adaptive vibration absorber with smart leaf springs is presented, where the stiffness of the absorber is controlled by the bending of the leaf springs. Another contribution dealing with a vibration absorber with adjustable inertia using a cone based continuously variable transmission can be found in [12]. Also in [13] an active vibration absorber with adjustable frequency using an electromagnetic actuator assembled at the end of a cantilever beam is investigated, and a similar approach is presented in [14]. The specific problem of a varying natural frequency of the primary system due to different working conditions is discussed in [15] by the use of multiple tuned mass dampers for a parallel machining robot.

In this paper the concept of a hydro-mechanical vibration absorber (HMA) with adjustable operating frequency is investigated by theoretic considerations. The basic design aspects with regard to optimal tuning and damping are based on the established literature [7]. Furthermore, simulations carried out with an exemplary parameter set confirm the expected functionality. Moreover, not only the mitigation effect for a weakly damped primary system is shown, also the applicability for moderate instabilities is presented.

2. BASIC CONCEPT OF THE HYDRO-MECHANIC VIBRATION ABSORBER

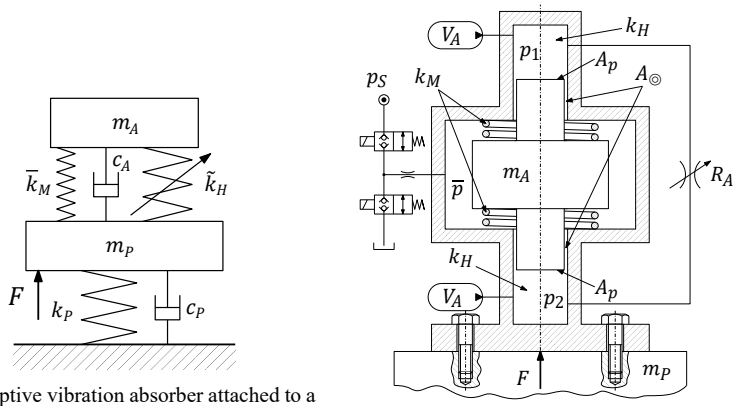
The basic functional scheme of an adaptive vibration absorber is illustrated in Fig. 1a attached to a primary damped spring mass oscillator (m_P, k_P, c_P) resonating due to a harmonic excitation force F. The absorber mass m_A is interconnected to the mass m_P of the primary system by a combination of two springs and a damper. Beside a constant spring coefficient $\overline{k}_M$ an additional variable stiffness $\tilde{k}_H$ is used in order to realize an adjustable resonance frequency of the absorber. In a so-called hydro-mechanic vibration absorber (HMA), as illustrated in Fig. 1b, the absorber mass resonates between two pressure chambers with gas-loaded accumulators. The nonlinear characteristics of the gas springs in the accumulators depend on the mean operating pressure $\overline{p}$ and result in the variable stiffness $\tilde{k}_H$. The hydraulic reactive force of the absorber $F_A = (p_1 - p_2)\, A_p$ is transmitted by the housing to the primary mass. The housing is pressurized by two directional on/off seat type valves to a constant mean pressure $\overline{p}$, which propagates slowly through the annulus gaps ($A_{\circledcirc}$) between the piston and the sleeve from the central chamber to both working chambers. In order to guarantee a proper dynamic operation of the absorber the relation $A_p \gg A_{\circledcirc}$ of the relevant cross-section areas must be fulfilled. Assuming that the dynamic leakage flow through the annulus gaps $A_{\circledcirc}$ is negligible, then only the fluid flow through the resistance R_A between the two pressure chambers results in a considerable damping effect.

The hydraulic stiffness of one chamber incorporating the gas-spring of the accumulator reads

$$k_H = \frac{A_p^2 n \overline{p}}{V_A} \left(\frac{p_0^G}{\overline{p}} \right)^{-\frac{1}{n}}, \tag{1}$$

where the compressibility of the oil is neglected for simplicity. Since the displacement of all individual springs k_M and k_H is unique, their interconnection results in a parallel[1] arrangement of springs. Thus,

[1] In a serial arrangement of multiple springs the force is unique.



(a) Adaptive vibration absorber attached to a primary resonating system

(b) Adaptive hydro-mechanic vibration absorber (HMA)

Figure 1: Adaptive absorber concept

Table 1: Parameters of a reference configuration

Parameter	Value
primary mass	$m_P = 100\,\mathrm{kg}$
primary stiffness	$k_P = 9.87 \cdot 10^6\,\frac{\mathrm{N}}{\mathrm{m}}$
primary resonance frequency	$\omega_P \approx 50 \cdot 2\pi\,\frac{\mathrm{rad}}{\mathrm{s}}$
absorber mass	$m_A = 5\,\mathrm{kg}$
piston diameter	$d_P = 40\,\mathrm{mm}$
accumulator volume	$V_A = 0.3\,\ell$
gas pre-load pressure	$p_0^G = 20\,\mathrm{bar}$
polytropic exponent	$n = 1.3$
hydraulic damping resistance	adjustable

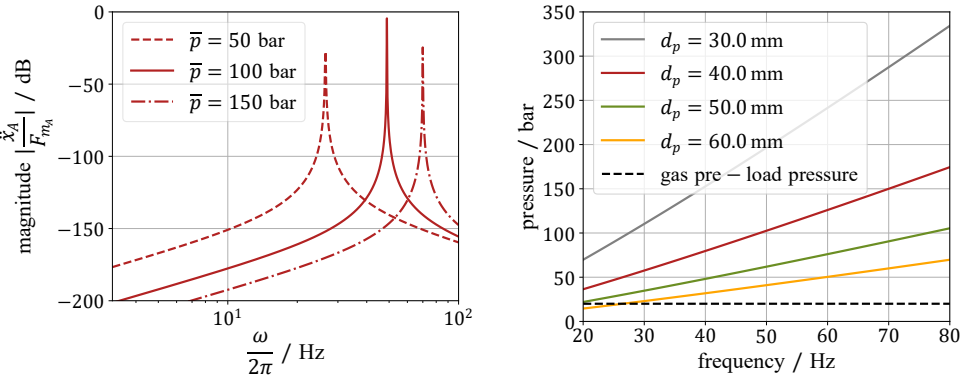
the resonance frequency of the absorber depending on the operating pressure $\overline{p}$ calculates to

$$\omega_A = \sqrt{\frac{\tilde{k}_H + \overline{k}_M}{m_A}} = \sqrt{\frac{2}{m_A}\left(\frac{A_p^2 n \overline{p}}{V_A}\left(\frac{p_0^G}{\overline{p}}\right)^{-\frac{1}{n}} + k_M\right)} \tag{2}$$

and the operating pressure depending on the resonance frequency ω reads

$$\overline{p} = \left(\frac{V_A m_A \left(p_0^G\right)^{\frac{1}{n}}}{2A_p^2 n}\omega^2\right)^{\frac{n}{n+1}}, \tag{3}$$

where k_M is neglected for simplicity. Both relations are illustrated in Fig. 2 for a parameter set according to Tab. 1, which results in an operating range between approximately 25 Hz up to 80 Hz for meaningful pressure values $\overline{p}$.



(a) Frequency responses for $d_P = 40\,\text{mm}$ **(b)** Piston diameters

Figure 2: Operating range

In Fig. 2b the mean pressure $\overline{p}$ related to the resonance frequency for different piston diameters $d_p = \sqrt{\frac{4A_p}{\pi}}$ is presented. The dashed black line indicates the gas pre-load pressure of the accumulator p_0^G, which must be sufficiently below the minimum operating pressures of both chambers in order to guarantee an appropriate operation. For instance, the absorbers with the parameters according to Tab. 1, but designed with larger piston diameters represented by the green and orange line, respectively, would be too close to or even below the gas pre-load pressure, which means that the accumulators are nearly empty in this frequency range. On the other hand, the design with a piston diameter $d_p = 30\,\text{mm}$ requires a high mean pressure for a high frequency operation, which would result in a too high ratio of operating pressure to pre-load pressure.

3. MODELING

The mathematical model of the adaptive HMA from Fig. 1b reads

$$\begin{bmatrix} \dot{x}_P \\ \dot{v}_P \\ \dot{x}_A \\ \dot{v}_A \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \begin{bmatrix} v_P \\ \frac{1}{m_P}\left(A_p\,(p_1 - p_2) - c_P v_P - k_P x_P + F\right) \\ v_A \\ \frac{1}{m_A} A_p\,(p_2 - p_1) \\ \frac{E_\mathcal{F}}{V_A\left(\left(\frac{p_0^G}{p_1}\right)^{\frac{1}{n}}\left(\frac{E_\mathcal{F}}{np_1}-1\right)+1\right)}\left(A_p\,(v_A - v_P) - \alpha\frac{Q_N}{\sqrt{p_N}}\sqrt{p_1 - p_2}\right) \\ \frac{E_\mathcal{F}}{V_A\left(\left(\frac{p_0^G}{p_2}\right)^{\frac{1}{n}}\left(\frac{E_\mathcal{F}}{np_2}-1\right)+1\right)}\left(-A_p\,(v_A - v_P) + \alpha\frac{Q_N}{\sqrt{p_N}}\sqrt{p_1 - p_2}\right) \end{bmatrix} \tag{4}$$

with the dynamic states x_P, v_P for the position and the velocity of the primary mass, x_A, v_A are the position and the velocity of the absorber mass and p_1, p_2 are the pressure states in the absorber chambers. As mentioned above the mechanical stiffness of the absorber k_M is neglected for simplicity. Furthermore, the cross-sectional areas of the annulus gaps are neglected ($A_\circledcirc = 0$) and, thus, no dynamic leakage is considered. The hydraulic resistance R_A between the absorber's pressure chambers is

considered as a proportional valve orifice (Q_N, p_N) with a square root characteristics and the opening parameter α. The effective compressibility modulus of the fluid is considered with $E_{\mathcal{F}} = 12000$ bar. For simplicity no viscous friction of the absorber mass is considered. The system from Eq. (4) will be considered later for the simulations, but for the design analysis a linearized model is used, which calculates to

$$\begin{bmatrix} \dot{x}_P \\ \dot{v}_P \\ \dot{x}_A \\ \dot{v}_A \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k_P}{m_P} & 0 & 0 & 0 & \frac{A_p}{m_P} & -\frac{A_p}{m_P} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{A_p}{m_A} & \frac{A_p}{m_A} \\ 0 & -\frac{A_p}{C_A} & 0 & \frac{A_p}{C_A} & -\frac{1}{R_A C_A} & \frac{1}{R_A C_A} \\ 0 & \frac{A_p}{C_A} & 0 & -\frac{A_p}{C_A} & \frac{1}{R_A C_A} & -\frac{1}{R_A C_A} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} x_P \\ v_P \\ x_A \\ v_A \\ p_1 \\ p_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m_P} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} F \tag{5}$$

with the dynamic matrix **A** and the excitation force F as an input. The capacitance of the absorber chambers reads

$$C_A = \frac{V_A}{n\overline{p}} \left(\frac{p_0^G}{\overline{p}} \right)^{\frac{1}{n}} = \frac{k_H}{A_p^2}, \tag{6}$$

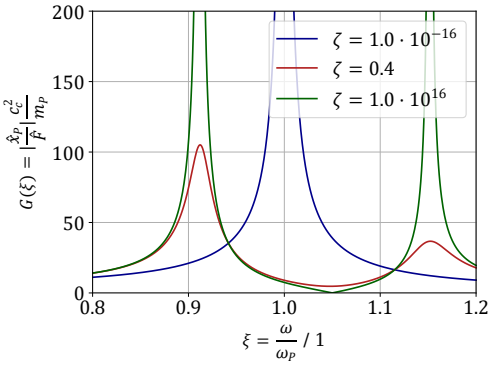
where, as in Eq. (1), the fluid is considered as incompressible ($E_{\mathcal{F}} \to \infty$) compared to the softness of the gas spring in the accumulators for simplicity. The parameter A_p denotes the cross-sectional area of the absorber piston and R_A is the linearized damping resistance between both pressure chambers of the absorber. For simplicity and in accordance with [7] the viscous friction of the primary system c_P is neglected in the design process. Since the cross-sectional areas of the annulus gaps $A_{\circledcirc}$ are assumed to be zero, thus, the pressure build up in the middle chamber of Fig. 1b and the corresponding damping effect are negligible. Thus, the transfer function of the displacement of the primary system due to the excitation force F follows to

$$G(s) = \frac{\hat{x}_P}{\hat{F}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} (s\mathbf{E} - \mathbf{A})^{-1} \begin{bmatrix} 0 & \frac{1}{m_P} & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}, \tag{7}$$

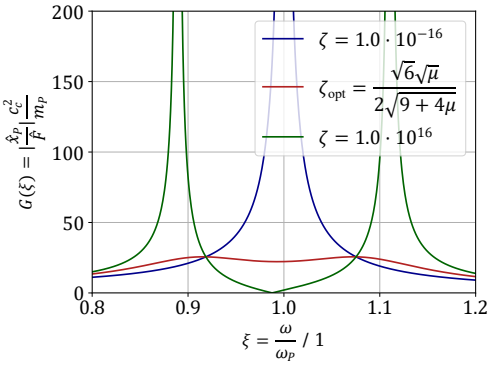
which results in a quite long symbolic expression impossible to be presented here. In order to keep the considerations simple the following substitutions $s = i\omega$, $\omega = \omega_P \xi$, $\omega_A = \omega_P \varphi$, $V_A = \frac{2A_p^2 n\overline{p}}{m_A \omega_A^2} \left(\frac{p_0^G}{\overline{p}} \right)^{-\frac{1}{n}}$, $m_A = m_P \mu$, $k_P = \omega_P^2 m_P$, $\omega_P = \frac{c_c}{2m_P}$, $R_A = \frac{c_c \zeta}{A_p^2}$ and $c_P = 0$ are used to rewrite this transfer function (7) in a dimensionless form

$$G(\xi) = \frac{\hat{x}_P}{\hat{F}} \frac{c_c^2}{m_P} = -\frac{4\left(\mu\varphi^2\xi - 2i\varphi^2\zeta + 2i\xi^2\zeta\right)}{\mu\varphi^2\xi^3 - 2i\mu\varphi^2\xi^2\zeta - \mu\varphi^2\xi - 2i\varphi^2\xi^2\zeta + 2i\varphi^2\zeta + 2i\xi^4\zeta - 2i\xi^2\zeta}. \tag{8}$$

The frequency response of Eq. (8) is depicted in Fig. (3a) for three different values ζ, which constitutes the damping according to the absorber resistance R_A. At low damping values of R_A, which is indicated by the blue line, only one resonance peak is present. This means, with regard to Fig. 1b, that the channel between both pressure chambers represents a short-circuit and, thus, the absorber is dynamically in an idle state and, furthermore, is not interacting with the primary system. Consequently, the degree of freedom of the absorber is lost and, thus, the transfer function has only one resonance frequency according to the original, undamped primary system without the HMA at the dimensionless frequency $\xi = \frac{\omega}{\omega_P} = 1$. On the other hand, the green line shows the frequency response with a very high resistance R_A, which has now two resonance peaks. The fact that both subsystems, the primary system and the absorber system, have nearly the same natural frequency results in a split-up



(a) Eq. (8) for 3 different damping values ζ at $\varphi = 1.05$



(b) red: optimized tuning φ_{opt} and damping ζ_{opt}

Figure 3: Transfer functions $G(s)$ of an HMA with $\mu = 0.05$ for different damping values ζ

of the single resonance peak into two eigenfrequencies, which are approximately mirrored around the dimensionless frequency $\xi = 1$. The third (red) curve in Fig. 3a shows a frequency response with a resistance value between both mentioned extremal values. It is remarkable that all three lines are passing through two distinct points, which are completely independent of the damping value ζ. They are called fixed points, which are located at the frequencies

$$\frac{\omega_I}{\omega_P} = \sqrt{\frac{\mu\varphi^2}{4} + \frac{\varphi^2}{2} - \frac{\sqrt{\mu^2\varphi^4 + 4\mu\varphi^4 + 4\mu\varphi^2 + 4\varphi^4 - 8\varphi^2 + 4}}{4} + \frac{1}{2}} \tag{9}$$

$$\frac{\omega_{II}}{\omega_P} = \sqrt{\frac{\mu\varphi^2}{4} + \frac{\varphi^2}{2} + \frac{\sqrt{\mu^2\varphi^4 + 4\mu\varphi^4 + 4\mu\varphi^2 + 4\varphi^4 - 8\varphi^2 + 4}}{4} + \frac{1}{2}}. \tag{10}$$

According to the design process from [7] the main strategy of the presented design paradigm is to bring both fixed points to the lowest possible level by an optimal tuning $\varphi = \frac{\omega_A}{\omega_P}$ and, furthermore, to place the maxima of the transfer function into these fixed points by adjusting the damping value ζ. For the considered primary system with the HMA this method succeeds with the optimal tuning

$$\varphi_{\text{opt}} = \sqrt{\frac{2}{\mu + 2}} \tag{11}$$

and the optimal damping

$$\zeta_{\text{opt}} = \frac{1}{2}\sqrt{\frac{6\mu}{4\mu + 9}}, \tag{12}$$

which is illustrated by the red curve in Fig. 3b. The resulting maximum magnitude calculates to

$$|G_{\text{opt}}|_{\max} = 4\sqrt{\frac{2 + \mu}{\mu}}. \tag{13}$$

Using Eq. (11), the optimal mean operating pressure in the HMA calculates to

$$\overline{p} = \left(\frac{V_A m_A \left(p_0^G\right)^{\frac{1}{n}}}{\left(\frac{m_A}{m_P} + 2\right) A_p^2 n} \omega_P^2 \right)^{\frac{n}{n+1}}, \tag{14}$$

which, in case of $\mu = \frac{m_A}{m_P} = 0$, tends to the result of Eq. (3). Since in most cases $\mu \ll 1$ is valid, it becomes obvious that the operating pressure $\overline{p}$ is robust with respect to the optimal tuning, which simplifies a proper design of the absorber significantly. With regard to Eq. (12) the relation for the resistance R_A reads

$$R_A = \frac{m_P}{A_p^2} \sqrt{\frac{6\mu}{4\mu + 9}} \omega_P. \tag{15}$$

In hydraulics a linear resistance is not easy to realize, rather the quadratic characteristics of an orifice. Since the orifice equation is nonlinear with respect to the pressure drop Δp, an operating point of the intended resistance must be considered, like

$$q = \frac{\Delta p}{R_A} = Q_N \sqrt{\frac{\Delta p}{p_N}}. \tag{16}$$

The pressure drop Δp in the operating point can be assessed by evaluating the transfer function for the chamber pressure due to the excitation force

$$G_p(\xi) = \frac{\hat{p}}{\hat{F}} A_p = \frac{i\mu\varphi^2\xi^2\zeta}{2i\mu\varphi^2\xi^2\zeta - \mu\varphi^2\xi^3 + \mu\varphi^2\xi + 2i\varphi^2\xi^2\zeta - 2i\varphi^2\zeta - 2i\xi^4\zeta + 2i\xi^2\zeta} \tag{17}$$

according to the system dynamics from Eq. (5). Using again the conditions for optimal tuning and optimal damping from Eqs. (11) and (12), then the maximum dynamic pressure magnitude in a pressure chamber calculates to

$$\tilde{p}_{\max} = \frac{1}{A_p} \sqrt{\frac{12\,(\mu + 2)\,(\mu + 3)}{56\mu^2 + 255\mu + 288}} F, \tag{18}$$

which can be used for solving Eq. (16) for the nominal valve size

$$Q_N = \frac{\sqrt{2\tilde{p}_{\max} p_N}}{R_A}, \tag{19}$$

which in turn can be realized by a type of proportional valve orifice. Finally, the order of magnitude of the power dissipation over the resistance R_A can be assessed by

$$\overline{P} = \frac{1}{T} \int_0^T q_{R_A}(t) \Delta p(t) \mathrm{d}t \approx 4 \frac{\tilde{p}_{\max}^2}{R_A}, \tag{20}$$

which is converted into heat.

4. SIMULATIONS

The dynamic behaviour of the absorber design is investigated by simulations as presented in the following. On the one hand the response due to an external excitation is analyzed, which is the most likely case in industrial applications. On the other hand the response resulting from a moderate instability in form of a negative damping value c_P of the primary system is shown.

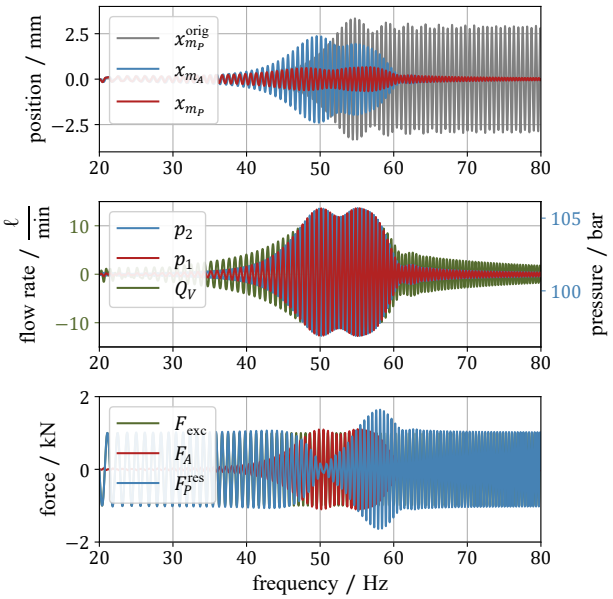
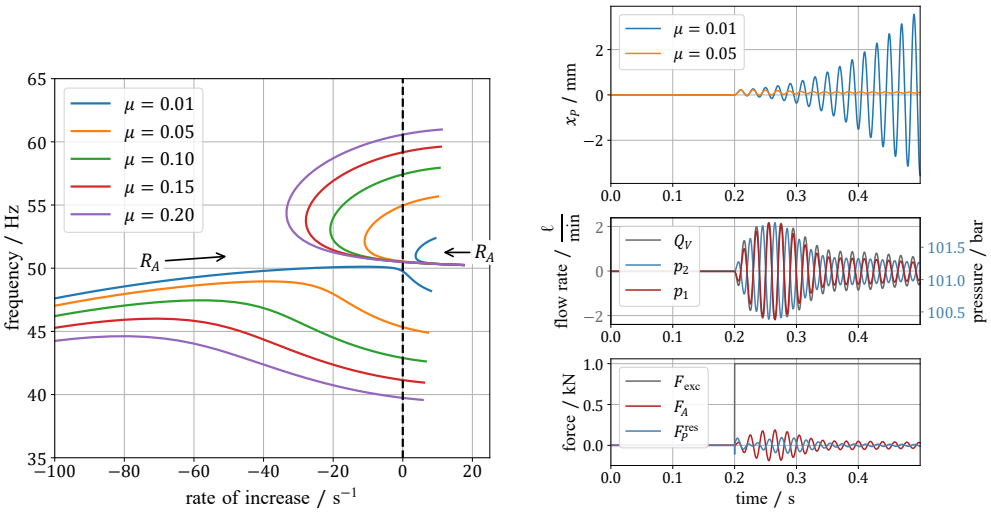


Figure 4: Response due to an external chirp excitation force F

4.1. External Excitation

In the simulation results presented in Fig. 4 the nonlinear system from Eq. (4) was excited by a chirp modulated force F_{exc} in order to investigate the system behaviour in a certain frequency range. In the upper diagram the grey signal $x_{m_P}^{\text{orig}}$ indicates the displacement of the primary mass without an absorber. The resonance frequency is designed to 50 Hz, thus, in this frequency range the magnitude is growing. In the simulations no viscous damping is considered for the primary system ($c_P = 0$). For this reason the magnitude remains constant at frequencies above the natural frequency of the primary system, since no power is dissipated. It must be remarked that in the simulations the chirp process was carried out within a simulation time of 2 seconds for illustration purposes. Thus, for a slower frequency rate a larger magnitude in the position $x_{m_P}^{\text{orig}}$ must be expected. The red line x_{m_P} in the upper diagram shows the displacement magnitude of the primary mass with the absorber and the blue signal x_{m_A} represents the magnitude of the absorber displacement. In the middle diagram the pressures in both absorber chambers as well as the flow rate through the damping resistance are illustrated. In the lower diagram the excitation force, the reaction force from the absorber and the resulting force on the primary mass as the sum from the excitation and the absorber are depicted.

The simulations taught that the system with the hydro-mechanical vibration absorber is well damped in the relevant frequency range. Furthermore, comparing the displacement of the damped primary mass in the upper diagram of Fig. 4 with the frequency responses in Fig. 3b, the maxima of the red optimized curve below and beyond the natural frequency of the primary system ($\xi = \frac{\omega}{\omega_P} = 1$) can be clearly identified. Thus, the nonlinear dynamics reflect the assumptions for the linear design process. However, this might not be the case anymore, if the excitation force exceeds a certain magnitude, which is not discussed in this contribution. According to Eq. (20) derived from the linear model the mean power dissipated over the resistance R_A is assessed to approximately 88 W and the measured dissipation power from the simulation with the nonlinear model results in approximately 95 W, which is in fact the same order of magnitude, but so far it cannot be assured that this holds for other dimensions.



(a) Stability chart with respect to ζ respectively R_A **(b)** Simulation results for two different values of μ

Figure 5: Unstable primary system with negative viscous damping c_P

4.2. Moderate Instability

In some applications a structural instability may lead to growing oscillations, like illustrated by the blue line in the upper diagram of Fig. 5b for the displacement of the primary mass x_P, which is caused by a negative damping of the primary system $c_P = -3500\,\frac{\mathrm{Ns}}{\mathrm{m}}$. In Fig. 5a the stability of the system at different mass ratios μ with regard to the damping resistance $R_A = 0, \ldots, 10^{12}\,\frac{\mathrm{Pa{\cdot}s}}{\mathrm{m}^3}$ is analyzed. Positive abscissa values denote the unstable region. At a resistance value $R_A = 0$ the system is unstable and oscillates at approximately 50 Hz with a diverging magnitude. Increasing R_A results in a reduced rate of increase, however, when the absorber mass ratio is too low ($\mu = 0.01$) the system cannot be stabilized at all. With a larger absorber mass like, for instance, with $\mu = 0.05$ the instability due to the negative damping can be prevented with certain values of R_A. But, if the value of the resistance is increased further, then the system is becoming unstable again with different frequencies, which also occurs with larger absorber masses.

5. DISCUSSION

The basic structure of the HMA is a resonating mass between two pressure chambers and in case of need two mechanical springs. Basically, the resonance frequency is determined by the relation of mass and resulting stiffness of the different springs. Both pressure chambers are equipped with gas-loaded accumulators and their stiffness is adjusted by the mean operating pressure of the gas spring resulting in an adaptive natural frequency.

The absorber is encapsulated in a closed housing, which is mounted to the primary system to be dampened. The momentum of the absorber mass is converted to pressures in both chambers and the reactive force is transmitted by the housing, which must be designed significantly stiffer than the rigidity of the resulting parallel arrangement of the mechanical and gaseous springs.

In the absorber concept from Fig. 1b the piston is gliding in a sleeve, which is lubricated by the dy-

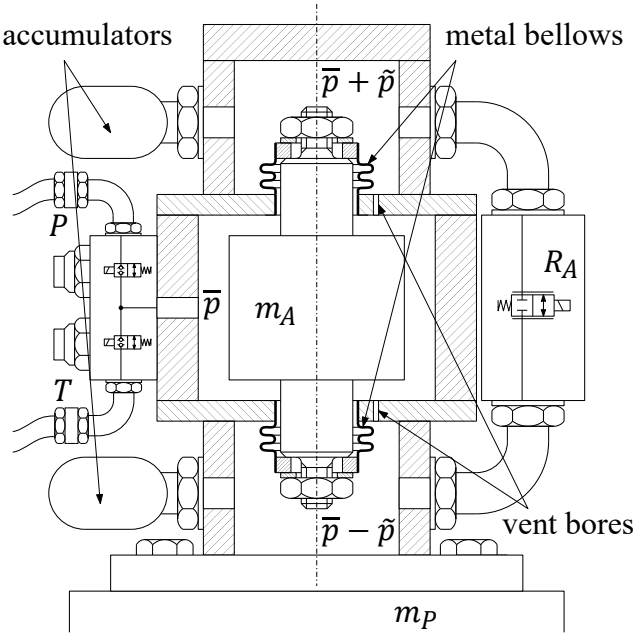


Figure 6: A realization study

namic leakage flow through the annulus gaps with the cross-sectional areas $A_{\circledcirc}$. In the design procedure presented above, this leakage flow was completely neglected, however, the cross-sectional areas $A_{\circledcirc}$ must be related to the nominal flow coefficient of the resistance R_A. Because, the clearance for lubrication and/or due to manufacturing tolerances between piston and sleeve may result in a considerable damping effect. Thus, the value of the resistance R_A must be obviously increased, which in turn can lead to very small and sensitive openings in the resistance valve. In order to avoid the problem of too high cross-sectional areas $A_{\circledcirc}$ a different design for an exemplary realization of an HMA is depicted in Fig. 6, where the absorber mass is supported by two metal bellows for guiding the piston and sealing the chambers. The mean pressure in the absorber is adjusted by two seat-type valves for increasing or decreasing the operating pressure $\overline{p}$, which propagates now through vent bores from the centered mean pressure chamber to both dynamic pressure chambers. The dimension of the vent bores can be designed independently from the absorber design, since requirements for guiding and lubrication can be dropped. Furthermore, in many cases the dynamic displacement of the absorber mass is in the sub-millimeter range, thus, the major design restrictions for the metal bellows are the dynamic pressure $\tilde{p}$ and the gravity force of the absorber mass. In such a case, the metal bellows do not significantly contribute to the main stiffness of the absorber ($k_M \ll k_H$), which is determined mainly by the gas springs in the pressure chambers. However, the soft metal bellows are still able to maintain the mean center position of the absorber mass. For applications where the operating frequency varies a control of the mean operating pressure $\overline{p}$ is necessary. For applications with one constant operating frequency, a hydraulic supply is not even required, because the absorber is pressurized just once during the installation.

The higher the magnitude of the external excitation force, the higher the expected magnitudes of the dynamic states in the absorber, i.e., larger displacement of the absorber mass and larger dynamic pressure values $\tilde{p}$. The assessment of the dynamic states presented in the considerations above require the knowledge or at least the order of magnitude of the excitation force, which must be somehow identi-

fied by corresponding measurements. The assessment of the dissipated power according to Eq. (20) leads to lower values compared to the simulations with the nonlinear systems dynamics, but the order of magnitude is similar. If already the assessed power dissipation at the resistance R_A is considerably high, then a cooling strategy could be necessary.

In some applications the excitation frequency varies during operation. Basically, the optimal damping according to Eq. (12) leads to a robust system behaviour with regard to the excitation frequency like illustrated in Fig. 3b. But this is not true anymore, when the natural frequency of the primary system varies, which can occur in certain applications. In such a case the resonating frequency of the absorber and, thus, the mean operating pressure $\overline{p}$ must be adapted in order to maintain an optimized behaviour. There exist different strategies for adapting the operating point, which are out of scope of this contribution, however, in such a case a hydraulic supply and a signal processing unit are required. Then additional pressure transducers can be applied directly to the different pressure chambers in order to observe the effectiveness of the absorber.

The final paragraph is devoted to the potential applicability of the presented absorber. Since it is encapsulated in a closed housing, the absorber can be easily assembled to any vibrating spring mass system independent of the environmental conditions. Only the suitable oil temperature could be a limitation for certain applications, but also in such cases probably cooling or even heating could be an opportunity. If applied to a vibrating structure, then the absorber must be located in an antinode of resonance in order to achieve its best performance. Basically, the concept is not limited in size or frequency, however, the smaller the dimensions of the absorber the higher the influence of parasitic effects like, for instance, manufacturing tolerances or friction effects. Furthermore, if commercial components are used for the realization, then also the available dimensions may play certain role. But, compared to conventional absorbers the additional parameters of the variable hydraulic stiffness give much room and flexibility in the design process, which makes the concept interesting for a wide range of applications.

6. CONCLUSION AND OUTLOOK

In this paper the basic concept of a hydro-mechanic vibration absorber with adjustable operating frequency was presented. The absorber consists of a lumped mass resonating between two pressure chambers with an adjustable stiffness depending on the mean operating pressure. The nonlinear spring characteristics are realized by the gas springs of hydraulic accumulators, where the compressibility of the oil can be completely neglected. The main design aspects were derived mathematically and, furthermore, the considerations for optimal tuning and damping based on linear systems theory were carried out with regard to the established literature. Simulation experiments taught that the basic functionality of the absorber could be shown in the more realistic case of nonlinear system dynamics. Moreover, the stabilizing effect of the absorber for an unstable primary system could be demonstrated. The major design aspects for a realistic application as well as the assessment of dissipated power were discussed elaborately. A realization of a prototype and the verification of the concept by measurements constitute next steps in development. Furthermore, a strategy for an online adaption of the operating frequency is part of future work.

ACKNOWLEDGEMENT

This work has been supported by the COMET-K2 Center “Center for Symbiotic Mechatronics” of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

REFERENCES

[1] Ioi T, Ikeda K (1978) On the dynamic vibration damped absorber of the vibration system. Bulletin of JSME 21(151):64–71, DOI 10.1299/jsme1958.21.64

[2] Yuen T, Balan L, Mehrtash M (2019) Implementation of an Absorber Design for Vibration Control in Automation Systems. Procedia Manufacturing 32:578–584, DOI 10.1016/j.promfg.2019.02.255

[3] Bschorr O, Raida HJ (2022) One-Way Vibration Absorber. Acoustics 4(3):554–563, DOI 10.3390/acoustics4030034

[4] Wang M, Xu H, He D, Wang T, Zhang J (2022) Design of a damped vibration absorber to control the resonant vibration of roll. Mechanical Systems and Signal Processing 178:109,262, DOI 10.1016/j.ymssp.2022.109262

[5] Trujillo-Franco LG, Flores-Morita N, Abundis-Fong HF, Beltran-Carbajal F, Dzul-Lopez AE, Rivera-Arreola DE (2022) Oscillation Attenuation in a Building-like Structure by Using a Flexible Vibration Absorber. Mathematics 10(3):289, DOI 10.3390/math10030289

[6] Daman AAA, Guntur HL, Susastro (2016) The influence of dynamic vibration absorber to reduce the vibration of main system with 2-DoF. In: AIP Conference Proceedings, Author(s), DOI 10.1063/1.4965762

[7] Den Hartog JP (1947) Mechanical Vibrations. McGraw-Hill Boo Company, Inc.

[8] Esen I, Koc MA (2015) Optimization of a passive vibration absorber for a barrel using the genetic algorithm. Expert Systems with Applications 42(2):894–905, DOI 10.1016/j.eswa.2014.08.038

[9] Hao Y, Shen Y, Li X, Wang J, Yang S (2020) H_∞ optimization of Maxwell dynamic vibration absorber with multiple negative stiffness springs. Journal of Low Frequency Noise, Vibration and Active Control 40(3):1558–1570, DOI 10.1177/1461348420972818

[10] Yuan M, Jin Y, Liu K, Sadhu A (2022) Optimization of a Non-Traditional Vibration Absorber for Vibration Suppression and Energy Harvesting. Vibration 5(3):383–407, DOI 10.3390/vibration5030022

[11] Guo X, Zhu Y, Qu Y, Cao D (2022) Design and experiment of an adaptive dynamic vibration absorber with smart leaf springs. Applied Mathematics and Mechanics 43(10):1485–1502, DOI 10.1007/s10483-022-2905-6

[12] Punyakaew S (2022) Adaptively tuned vibration absorber using cone based continuously variable transmission. Cogent Engineering 9(1), DOI 10.1080/23311916.2022.2102044

[13] Rincon CG, Alencastre J, Rivera R (2021) Active Vibration Absorber for a Continuous Structure Model. engrXiv (Engineering Archive) DOI 10.31224/osf.io/r3ujn

[14] Timorian S, Valasek M (2022) Dual Frequency Vibration Absorber. In: Computational Mechanics 2022, University of West Bohemia

[15] Zhang J, Xie F, Ma Z, Liu XJ, Zhao H (2023) Design of parallel multiple tuned mass dampers for the vibration suppression of a parallel machining robot. Mechanical Systems and Signal Processing 200:110,506, DOI 10.1016/j.ymssp.2023.110506

ENERGETIC OPTIMIZATION OF AN EXISTING CLAMPING POWERPACK BY SYSTEM AND CONTROL CONCEPT ANALYSIS AND ADAPTION OF THE HYDRAULIC FLUID VISCOSITY

Johannes Gattinger[1]*, Markus Schwimmer[1]

[1] *WEBER-HYDRAULIK GMBH, Siemensstraße 17, 84109 Wörth a. d. Isar*

* Corresponding author: Tel.: +49 173 9310487; E-mail address: johannes.gattinger@weber-hydraulik.com

ABSTRACT

Many hydraulic systems that have been in operation for several years are planned to be used further on some time. As nowadays the design objectives and possibilities are a lot more focused on energy efficiency than at the time these machines were build, retrofit activities such as energetic optimization are often useful in existing installations.

The clamping powerpack of an automated milling station is examined according to the control strategy of the pump motor by systematic measuring and recording of all relevant system parameters like operating status of the motor, activation status of control valves and pressure switches as well as the different hydraulic pressures in the system.

The system evaluation showed that the motor is driven at constant speed on demand by the controller of the milling station but with a time limited switch-on time, which leads to high losses due to oil flow over the systems relief valve. By improving the hydraulic system design, due to the addition of a pressure accumulator, and changing the control signal of the motor to a separate pressure switch, the energy consumption of the clamping powerpack is reduced by 89 %. In addition, the hydraulic oil viscosity was found to have no significant influence on the efficiency of the powerpack, which was presumed because of the specific application, where pressure, not flow, is the predominant parameter.

Keywords: clamping, power unit, efficiency, system, control concept

1. INTRODUCTION

Major advantages of hydraulic systems are their very high reliability and durability which leads to very long machinery lifetimes. Therefore, much machinery installed years ago is planned to be used further on for some more time. Nowadays, energy efficiency is a much more important design parameter then it was at the time these machines were planed and manufactured. The potential of systematically energetic optimization of existing installations is significant and some measures can be realized with manageable effort. [1]

This study focusses on the optimization of the control concept and system architecture of a clamping powerpack of a milling machine where high amount of wasted energy was assumed.

2. METHODS

The investigations were done with a hydraulic powerpack of an automated milling station which consists of three combined production cells that were fed with workpieces by a robot. The production cells are initially 3-axis milling machines which were upgraded with a fourth axis and a hydraulic clamping unit. The hydraulic oil supply for the clamping and the locking of the fourth axis is realized

via three separate powerpacks (one for each production cell). The hydraulic circuit diagram of the powerpacks is shown in **Figure 1**. For the evaluation of the control logic and understanding of the system behavior pressure sensors (Trafag 8252.34.2517) were installed at A and B ports of the clamping and the locking section and at the port MP for the pump pressure. Additionally, all signals of the pressure switches, the voltage of the valve solenoids and the voltage signal from the motor contactor are recorded with the DAQ system USB-AD16f (bmcm). For the investigation of the efficiency the power consumption was measured with the energy meter EASTRON SDM72D-M. The pulse output of the active energy from the energy meter was count on the digital input channel of the AD16f. The oiltemperature is recorded with a Parker SCT-150-0-02, connected to the DAQ via a SCM-200.

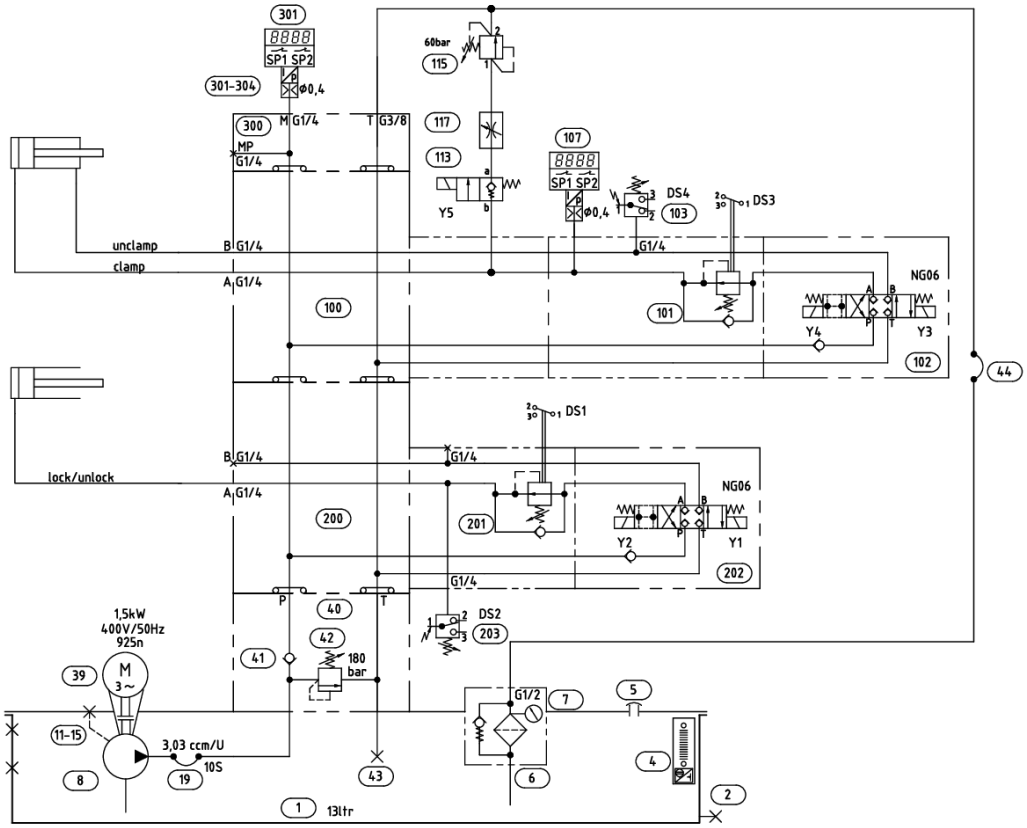


Figure 1: hydraulic circuit diagram of the powerpack

The first step of the evaluation was to record a typical cycle at the powerpack. This is absolutely necessary to understand the machine process and the logic of the higher-level machine control. In **Figure 2** a recording of the original cycle is shown. A cycle starts/ends with unclamping of the workpiece followed by a clamping and rotating the fourth axis to rinse the clamping jaw before clamping the new workpiece. The clamping pressure is adjustet to nearly 80 bar. The pressure to lock the rotation of the fourth axis is set to 65 bar. The main pressure relieve valve is set to 175 bar. At every clamping/unclamping and locking/unlocking operation the motor starts by request of the machine controller for providing required hydraulic energy. To ensure immediate availability of the pressure the motor is started 500 ms before the valve is activated. The motor stop is time-based after

5 s. This control concept leads to a calculated hydraulic energy consumption of 1,4 Wh ($E = \frac{p \cdot Q}{0,8} \cdot \Delta t$) [2] at every clamp or lock operation regardless of the actual energy demand. The motor is running 60 % of the cycle time of 483 s. while lot of energy is dissipated by the flow over the pressure relieve valve the oil temperature reaches a static value of 90 °C.

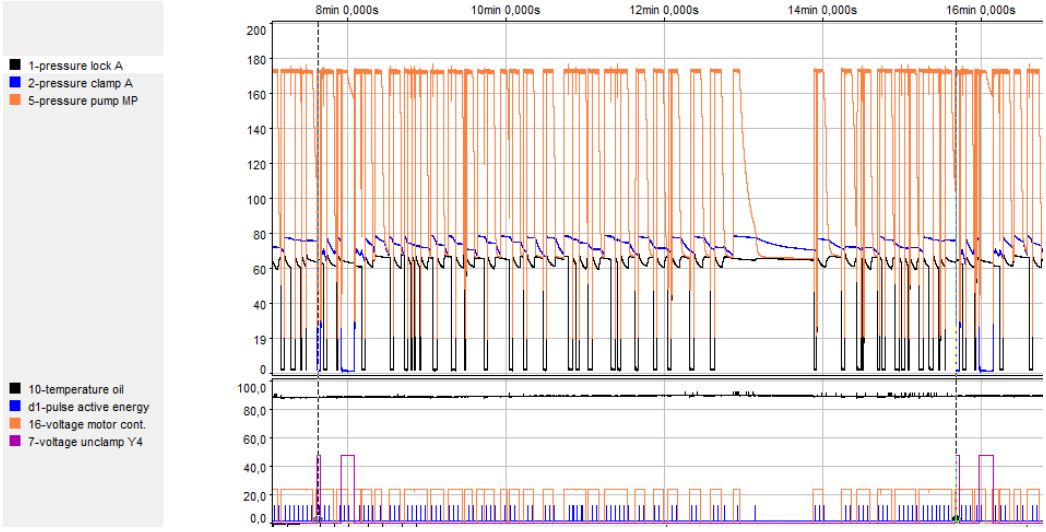


Figure 2: recording of powerpack cycle before optimization

By knowing the behaviour of the powerpack respectively the control logic from the recordings of the relevant system parameters it is totally clear that there hast to be found a solution to prevent the motor from pumping most of the time just over the system pressure relief valve. Several solutions would be possible but one of the easiest in this application is to integrate a small accumulator (0,75 l) at the port MP. The clamping and locking pressure are fed from the accumulator which just needs a few reloads within a cycle because the cylinders make very little stroke. Besides the modification of the hydraulic system a little modification on the electric side has to be done to realize the proper function of the higher-level machine control and the reload function for the accumulator. The higher-level machine control gives an output to the motor contactor every time a state of clamp/unclamp or lock/unlock should change. With no change in the motor control logic the pump would run almost the same duration over the relief valve than before. To solve that issue, without changing the control software, a separate pressure switch is installed at MP and the switching output is connected to the motor contactor replacing the signal line from the machine controller. These modifications are done with little effort on the hydraulic and electric system and without changes of the higher-level machine control logic. The control signals and software that are related to the functional safety of the machine (e. g. the signals of the pressure switches) are not changed.

The optimization results in a reduction of the oil temperature and thus the chance to select a lower oil viscosity to enable further savings through ideal oil selection. Two different tests are carried out to check the influence of viscosity on the efficiency of the power pack. In the first test, the temperature dependence of the viscosity is utilized and the power consumption is determined as a function of the oil temperature in order to make the influence of the viscosity visible. This test is done with the initial system design, where a temperature range from room temperature to 90 degrees can be examined. With the used oil AVIA FLUID HLPD 32 in this temperature range a viscosity range of 76 mm²/s – 6,7 mm²/s is tested. The second test is done with the optimized system and two different oils with the viscosity grades 32 and 10 (AVIA FLUID HLPD 32 and AVIA FLUID RSL 10), which leads to the

viscosities of 76 mm²/s and 19 mm²/s at room temperature. The viscosities as a function of the temperature are calculated with the formulas of Ubbelohde and Walter respectively DIN 51563 [3, 4, 5].

3. RESULTS

A recording of a cycle of the optimized system is shown in **Figure 3**. The motor runs an average of 2,5 times per cycle to reload the accumulator whereas clamp/unclamp operation needs most of the stored hydraulic energy. To realize the lock/unlock operations within the cycle only one reload per cycle is necessary. The oil temperature remains at 22 °C, which corresponds to the room temperature.

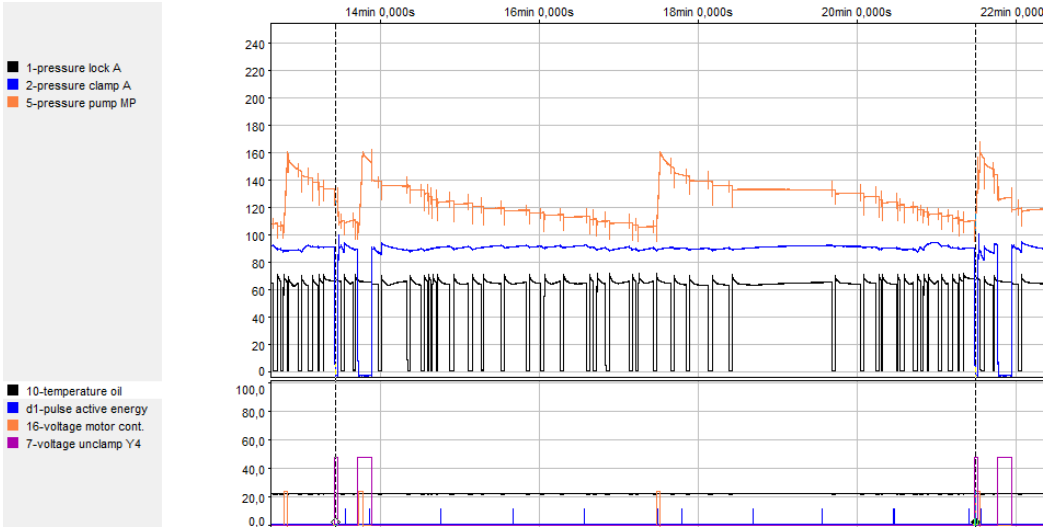


Figure 3: recording of powerpack cycle after optimization

The overall energy consumption of the powerpack is reduced from 96 Wh to 11 Wh per cycle which corresponds to a power reduction from 693 W to 79 W in relation to the cycle time. By analysing the period when the motor is not running the power requirement of the solenoid valves is determined to 66,5 W. That means that the power requirement of the motor in the optimized system is only 12,5 W on average.

The evaluation of the measurements to check the influence of viscosity on the power requirement did not show any significant effect. **Figure 4** shows the results of the first test with the original system at increasing temperature. There seems to be a slight reduction in power consumption with increased temperature, but the coefficient of determination is very low. With the conversion of temperature into viscosity there is also no significant dependency on the power consumption identifiable.

The overall effect of the optimization is shown in **Figure 5**. A huge decrease of the power consumption is realized by the optimization of the hydraulic system and the control strategy. The change of the hydraulic oil to a lower viscosity grade at the optimized system shows no significant influence on the power consumption. The chart includes measured values at different temperatures which is acceptable because no significant effect of the temperature was observed. By reducing the number of measured values in the chart of the optimized system with HLPD32 to a temperature of 24 °C the result is qualitatively equal. The only difference would be the omission of the lower whisker, due to the omission of the minimum measured value, which was measured at 36 °C. It should be noted that the energy meter used cannot resolve more precisely than 1 Wh and the values measured

on the optimized system vary between 10-12 Wh per cycle regardless of the temperature (viscosity).

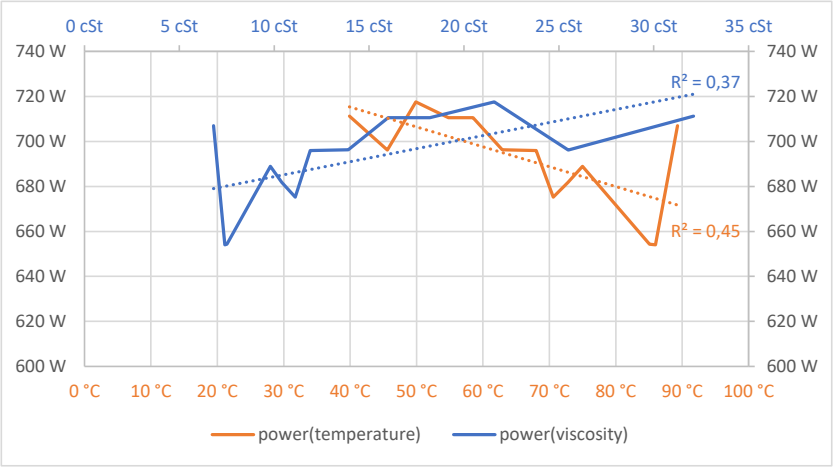


Figure 4: influence of fluid viscosity on the power consumption of the clamping system.

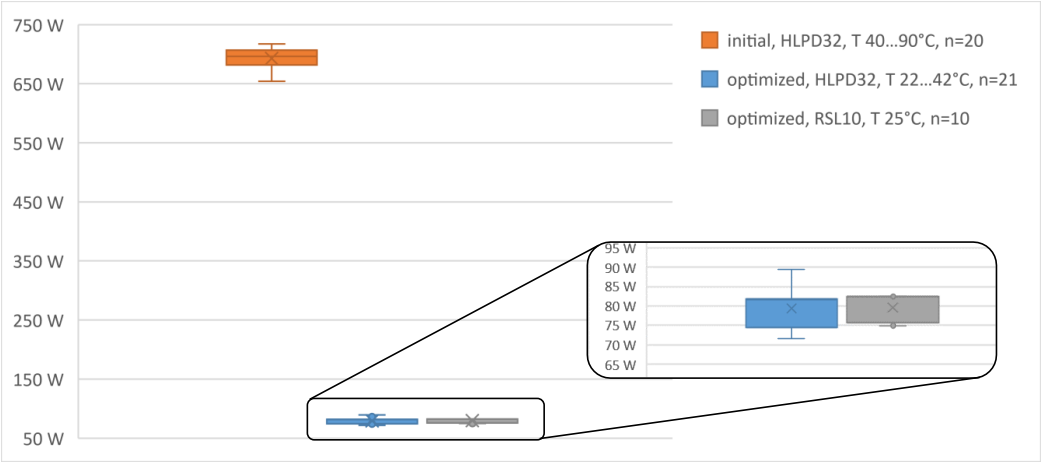


Figure 5: power consumption of the initial clamping system, the optimized system and the optimized system with a lower viscosity grade oil.

4. CONCLUSION AND OUTLOOK

A comprehensive analysis of the work process of the clamping unit and a critical examination of the implemented control strategy made it possible to identify considerable energy-saving potential. The power requirement was reduced by 89 % by making minor changes to the hydraulic system and the control of the clamping unit. The resulting reduction of the oil temperature from 90 °C to room temperature also extends the service life of the seals and the hydraulic oil, thereby increasing machine availability through reduced maintenance and repair times. The number of motor starts and the operating time of the motor were also reduced, which increases the service life of the motor, the relay and the contactor.

Unfortunately, no significant results were obtained for the influence of oil viscosity on energy

efficiency. The reasons for this include the fact that the hydraulic performance of the clamping system in question is mainly defined by the required pressure and less by the volume flow and therefore the viscosity induced flow losses in the hydraulic lines are relatively low. It may be possible to prove an influence of viscosity with a higher number of measurements and the use of an energy meter with higher measurement accuracy. However, the present results show that any additional energy savings due to oil viscosity in this specific case are negligible, since the motor's average remaining power requirement is only 12,5 W.

REFERENCES

[1] T. Radermacher, M. Merx, A. Sitte, V. Boyko und M. Unger (2021) Potenzialstudie Energie-/Kosteneinsparung in der Fluidtechnik. Umweltbundesamt, Dessau-Roßlau

[2] J. Dobrzanski (2017) Formelsammlung Hydraulik. Parker, Kaarst

[3] DIN 51563 (1976) Prüfung von Mineralölen und verwandten Stoffen; Bestimmung des Viskosität-Temperatur-Verhaltens, Richtungskonstante m. Beuth Verlag GmbH, Berlin

[4] H. Matthies und K. Renius (2014) Einführung in die Ölhydraulik. Springer Fachmedien, Wiesbaden

[5] L. Ubbelohde (1936) Zur Viskosimetrie. Hirzel, Leipzig

USE OF BROADBAND SILENCERS IN HYDRAULIC CIRCUITS TO REDUCE PULSATIONS

Peter Kloft*, Torsten Kusserow

HYDAC Technology GmbH, Research & Development, Justus-von-Liebig-Straße 10, 66280 Sulzbach/Saar, Germany

* Corresponding author: E-mail address: peter.kloft@hydac.com

ABSTRACT

New axial piston pumps with fixed displacement volumes, such as the high efficient power package (HEPP) pumps, operate at variable rotational speeds and thus variable flow rates in a wide range. Using intelligent pump control systems, small and light packages are possible. In contrast, the frequency of the fluid pulsations vary due to the rotational speed and have to be controlled avoiding stress or wear in the hydraulic system components (e.g. seals of tubing) and potentially impacting noise emissions.

Especially in mobile machines, the noise of a combustion engine drown out the noise of the hydraulic system. By replacing the combustion engines by electric drives, the fluid born noise (FBN) becomes audible.

Due to the excitation frequencies, a broadband silencer for high-pressure applications is required. This is realized by a Multi-Helmholtz-Resonator (MHR) concept within one cylindrical volume. The design will be adapted for pump-specific frequency characteristics in aircraft hydraulic conditions and fine adjusted during test verification by means of an adjustable tube-inside a pressure shell.

Keywords: Pulsation, Fluid Born Noise, Hydraulic Dampener, Silencer, Piston Pumps, Gear Pumps, Excitation frequencies, Resonance, Multi Helmholtz Resonator

1. INTRODUCTION

HYDAC Technology designs, produces and sells all kind of hydraulic accumulators such as diaphragm-, bladder-, piston- and metal bellows accumulators covering a standard pressure range up to 1,200 bar and volumes from 0.07 dm3 to more than 3,000 dm3. In addition to a classic integration of accumulators into hydraulic systems, the basic use as a hydraulic damper defines the HYDAC product category of hydraulic dampers, where the specific cause of the pulsation is taken into account.

The pressure ripples occurring in hydraulic systems vary in cyclical or one-off problems like those from displacement pumps, valve actuation or system start-up and shutdown. Based on the specific problems and the corresponding excitation frequencies, different hydraulic damper types are recommended. A distinction is generally made between the usage of the capacitive effect of gas volumes on the one hand and damping principles of the hydraulic liquid on the other hand.

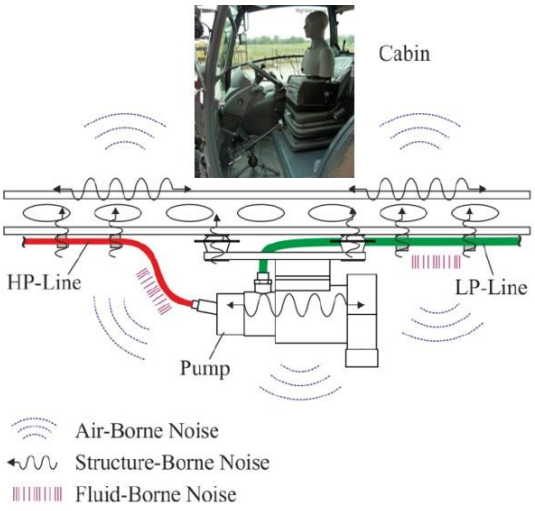


Figure 1: Noise measurement in a mobile application – figure modified taken from [1]

Two types of liquid dampers are frequently used and well known theoretically, Helmholtz Resonators and inline expansion chambers. HYDAC's silencer concept is a mixture of both types increasing the attenuation levels for the given envelope space. The functional technology is recommended for excitation frequencies above 175 Hz, which is typical for industrial axial piston pumps with corresponding displacement elements and rotational speeds. The presented broadband technology extends the damping effect from dedicated frequencies to a wider frequency range, which meets the requirements of variable speed pumps.

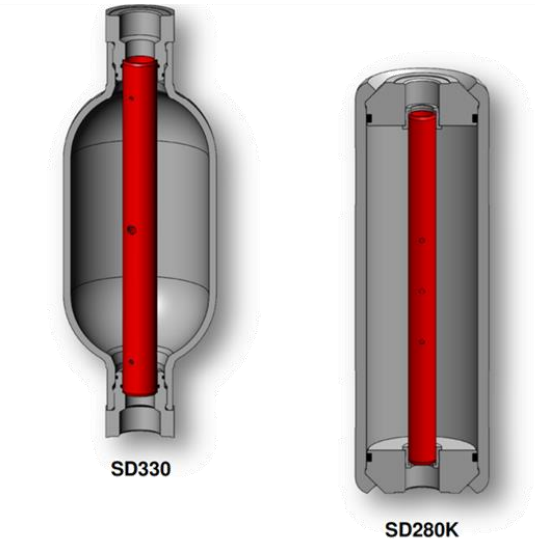


Figure 2: HYDAC industrial broadband-silencer types SD330 and SD280K

HYDACs broadband silencer design was adapted for pump-specific frequency characteristics in aircraft hydraulic conditions. Fundamental knowledge for the design of corresponding damper types was developed during investigations in the framework of aviation research projects funded by the

ministry of economics of Germany (BMWi) [3, 4]. A main driver for the insertion tube design is the wide excitation frequency range of the axial piston pumps used. Tests accompanying the simulation to validate the complex internal processes are essential for product development of these damper types and model updating of the simulation.

2. DAMPING PRINCIPLES - BROADBAND SILENCER TECHNOLOGY

The basic principle of an inline Helmholtz-Resonator for HYDACs silencer technology was developed for more than 20 years and has been used successfully in industry ever since. Figure 2 illustrates the differences in damping level amplitude and bandwidth for an inline chamber and a single Helmholtz Resonator [1]. The silencer combines the possibility of an inline integration of the hydraulic damper in the hydraulic circuit with the higher damping levels of a Helmholtz-Resonator compared to an inline chamber.

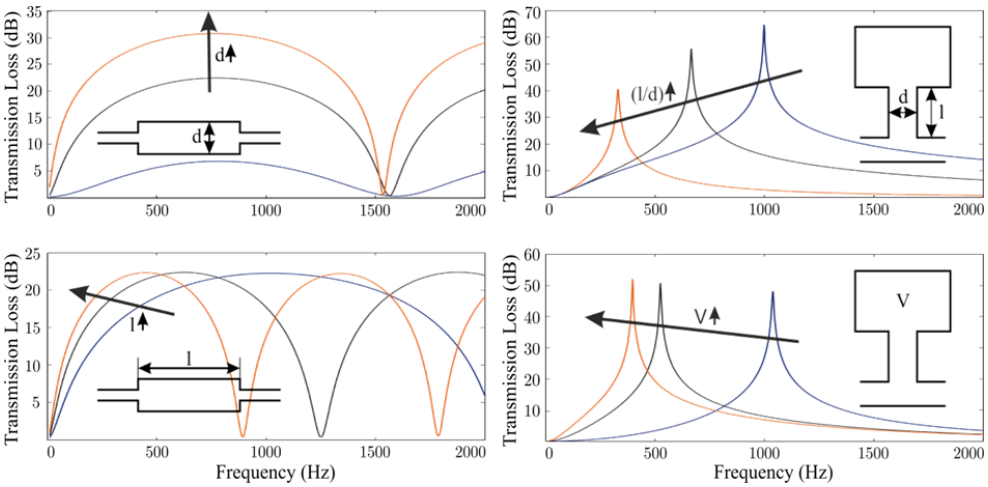


Figure 2: Different damping principles - figure taken from [1]

This is realized by integration of an insertion tube in an accumulator shell, separating one cylindrical volume representing the stiffness of a Helmholtz Resonator. Every radial bore of the insertion tube is linked to the neck of a Helmholtz Resonator and the cylindrical fluid inside acts as a point mass. Inertia and stiffness define the natural frequency of this single vibration absorber. The combination of different bores within one tube is the basis of HYDACs broadband silencer concept, achieving a sufficient hydraulic damping level over the frequency range due to superimposed Resonator effects. Due to the wider frequency bandwidth of aviation hydraulic systems, the number of required necks was extended to a sufficient level compared to the proven industrial design. This additionally requires the shift of the resonator frequencies to higher levels. The damping principle via parallel Helmholtz Resonators of the broadband silencer concept is hereby very robust against temperature fluctuations and the associated fluctuations of the speed of sound in the hydraulic fluid. This is a significant advantage in the aviation sector, where an operating temperature range from -55 to over 100 °C can be assumed.

Besides the positive impact on the fluid borne noise of the hydraulic system and the corresponding noise level within the passenger cabin, the service life and reliability of the existing hydraulic components inherent in the system can be increased.

3. TEST SETUPS

As mentioned before, the validation by test of the broadband silencer design is a main and important driver for the product development. The experimental analysis of hydraulic dampers requires a dedicated sinusoidal pressure ripple source, which is applied to the test object. The adjustability of the ripples amplitude and frequency are core requirements.

In general, two different types of hydraulic damper tests are possible. The silencer can be adapted directly to the axial piston pump to realise a high coupling of the damper to the pressure ripple source (see Figure 3), neglecting the inner length of the pump. Behind the damper, a measurement pipe with different pressure sensor positions allows the damping analysis for different frequency ranges. A big advantage is the elimination of influences of the transmission behaviour of further components in between. Otherwise, this could have a negative impact on the pressure ripple amplitude in the silencer (oscillation node). Nevertheless, it is not possible to determine reflection effects of the test bench setup in order to consider them when determining the damping performance.

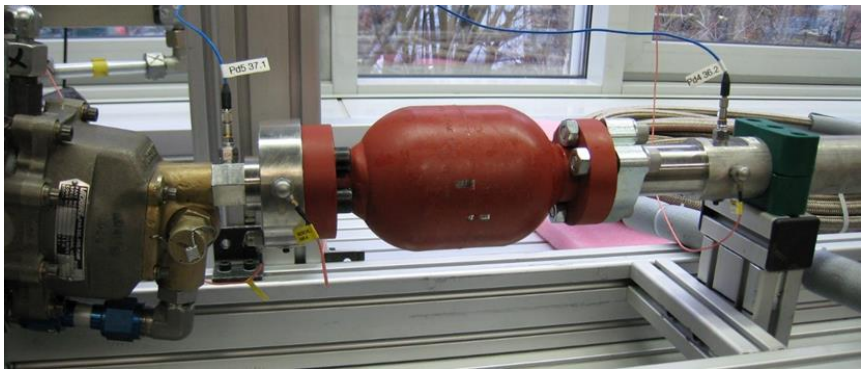


Figure 3: Aviation pump and directly flanged HYDAC standard silencer - FST TUHH test bench

For the possible high excitation frequencies of the pulsation ripples of up to 1.500 Hz in aviation high efficient power package (HEPP) systems, a modification of the test bench was necessary in addition to the broadband silencer design itself. In collaboration with the Institute of Aircraft Systems Engineering (FST) at the Hamburg University (TUHH), relevant components were modified, such as the measurement pipes and the rotary valve design (see Figure 4 and Figure 5).

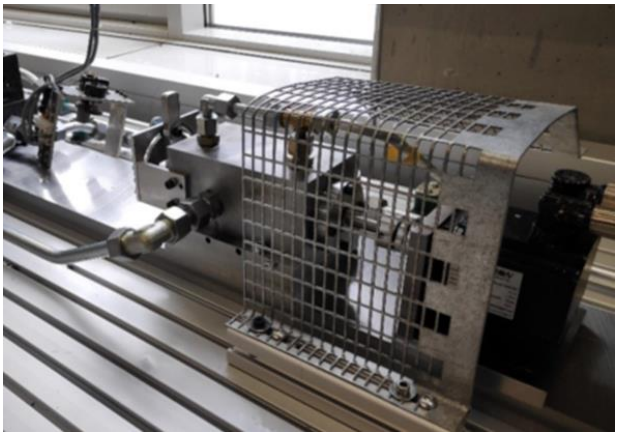


Figure 4: Integration of advanced rotary valve - FST TUHH test bench

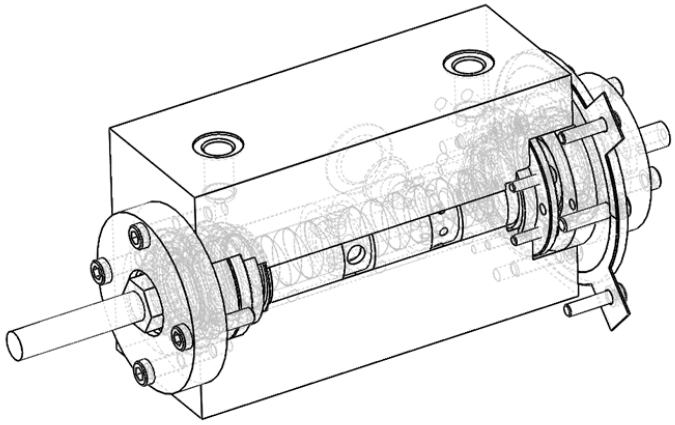


Figure 5: Semi-cut CAD line model of modified rotary valve

Alternatively, an additional measurement pipe can be placed in front of the silencer for high laminar inflow and additional sensor positions in front of the hydraulic damper (see Figure 6 and Figure 7). The flow ripple source is realized by the integration of a rotary valve for effecting the pulsation. The pump pulsations of the hydraulic power unit used upstream of the rotary valve is hereby levelled by the pipe distance to the hydraulic power source and an additional damper. Due to the distance between the pressure ripple source (rotary valve) and the test silencer, an impact on the measured pressure amplitudes for the different sensor positions remains in this test setup configuration. Different varying pulsation frequencies cause the variation of the wavelength of the pressure ripples and require the adaption of the sensor positions of both measurement pipes for the observed frequency range, which is illustrated in Figure 7. The movement of the oscillation nodes is clearly recognisable and important for the evaluation and interpretation.

Based on the second modified test setup with two measurement pipes, different damper volumes and insertion tube designs for the broadband silencer concept for aviation applications were tested and analysed at the Institute of Aircraft Systems Engineering FST at Hamburg University of Technology [7]. Figure 8 illustrates qualitatively the effect of a cylindrical resonance volume increase for the same insertion tube. It can be stated in principle, that a larger resonance volume increases the damping performance. Due to the logarithmic representation of the graphs, this applies in particular to the lower frequencies. For higher frequencies, the integration of hoses between hydraulic pump and damper is beneficial.

A further investigation of the variation of the fluid flow for laminar levels showed no major impact on the measured damping performance of the broadband silencers. In general, a laminar inflow is essential for a proper damping behaviour of this Multi-Helmholtz Resonator concept. Jumps in the area cross-section and corresponding turbulent flow decreases the performance and have to be avoided during hydraulic system conception and silencer integration.

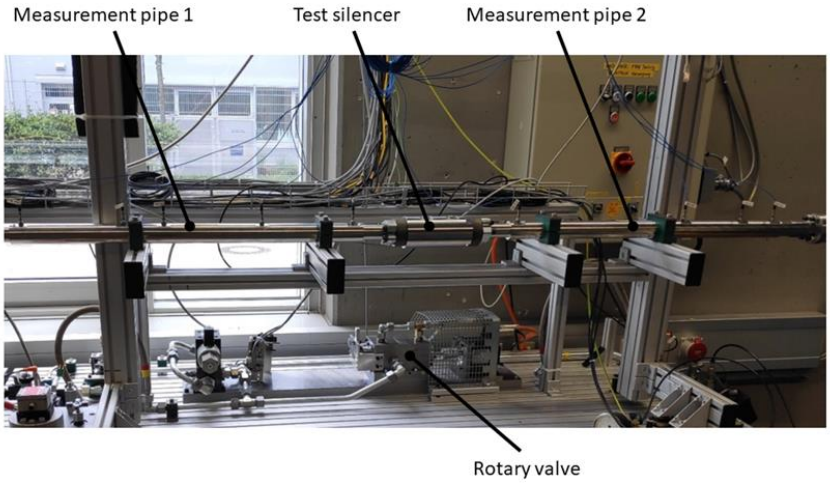


Figure 6: Test setup with two measurement pipes - FST TUHH test bench – figure taken from [7]

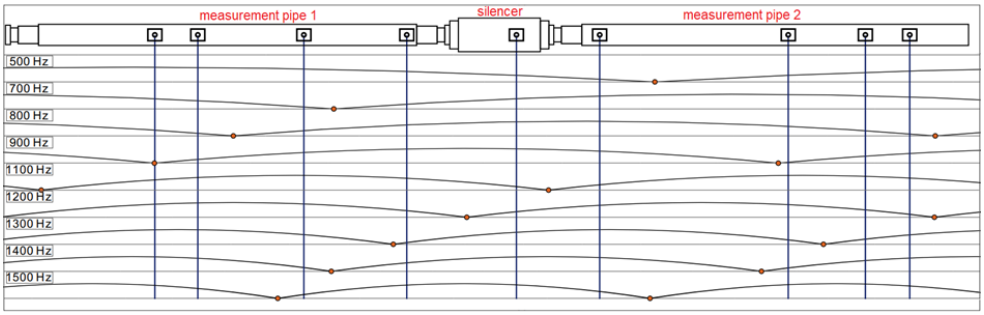


Figure 7: Pressure sensor positions of test setup and wavelength influence on signal strength - FST TUHH test bench – figure modified taken from [7]

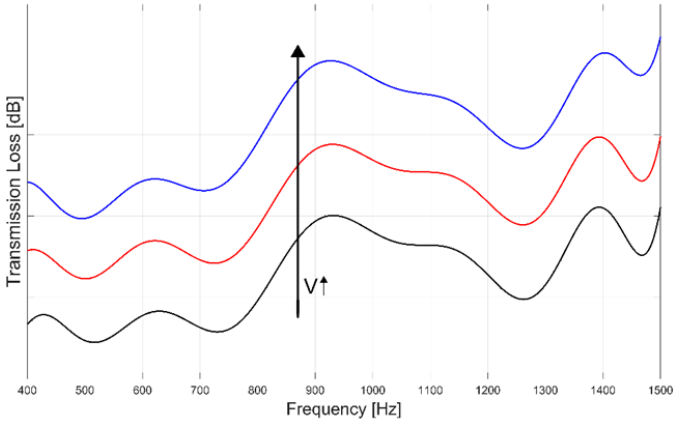


Figure 8: Qualitative influence of resonance volume on silencer damping performance

4. SIMULATION

Since 1997 HYDAC uses Hydroplus, a software developed at the University of Stuttgart. Unfortunately, this software was not updated anymore and therefor HYDAC switched to a new software package. The University of Bath offers a pressure ripple analysis software package (PRASP). A Matlab/Simulink® toolbox models FBN characteristics of hydraulic circuits. The fluid-borne noise characteristics of positive displacement hydraulic pumps can be measured using the Secondary Source test method. This forms the basis of British Standard BS6335 pt. 1 and ISO 10767-1: 2015 for the rating of pump fluid-borne noise. Hydraulic motors, valves and hoses can also be tested using this method [5].

Already during the aviation research projects in collaboration with the FST TUHH [3, 4], the modelling capability of the number of bores and corresponding Helmholtz chambers was studied for a filter silencer using PRASP (see Figure 9) and compared with corresponding test results [6]. This investigation was limited to two bores for one resonance volume. The modelling complexity of the superposition of many bores for one resonance volume for the broadband silencer concept is all the greater, as neither the direct combination of the bores in the model nor the separate consideration of the volume for each borehole reflects practice. In order to improve the current modelling and design development methods, targeted investigations on different insertion tubes are necessary. This concerns above all the effective capacity of the cylindrical resonance volume in the simulation model that is available for each bore with superimposed utilisation. This is a decisive issue for the methods of design finding, particularly with regard to the variance of the bore diameters and consequently the respective frequency effect of each individual resonator.

Besides the complex transmission behaviour of the broadband silencer itself, the hydraulic circuit with all components has to be implemented in a simulation model. The best possible positioning for maximum effectiveness in the system is the general aim of the system simulations. In general, for decoupling and improved damping, especially at higher frequencies, a hose between the pump and damper is common and expedient. But the conflicting objectives for the optimum distance in the frequency range remain the same as in the test setup. In addition to the aforementioned position dependence of oscillation nodes for different excitation frequencies, the temperature also has an influence on the density of the fluid and thus the speed of sound. This change in turn causes an adaption of the wave structure. Hereby, the requirements in the aviation sector clearly outweigh those of stationary industrial hydraulic systems. Consequently, the simulation-based investigation of main influencing factors is relevant for the success of the damper integration on system level. The best damper design loses its favourable effect if it is located at an oscillation node.

This requires the most accurate possible description of the transmission characteristics in the system design, including pump and hose modelling. This is to be able to estimate the influence on the hydraulic sound velocity on system level reasonably. With the help of PRASP and its integrated empirical values in hose modelling, the system-relevant verification process is to be advanced before the respective system is set up. Research is still needed, especially with regard to the visualisation and consideration of turbulent flow phenomena.

The overriding factor here is always the reduction of time-consuming measurement campaigns through predictions that are as precise as possible and ideally, only the need of a final prototype measurement is the declared objective.

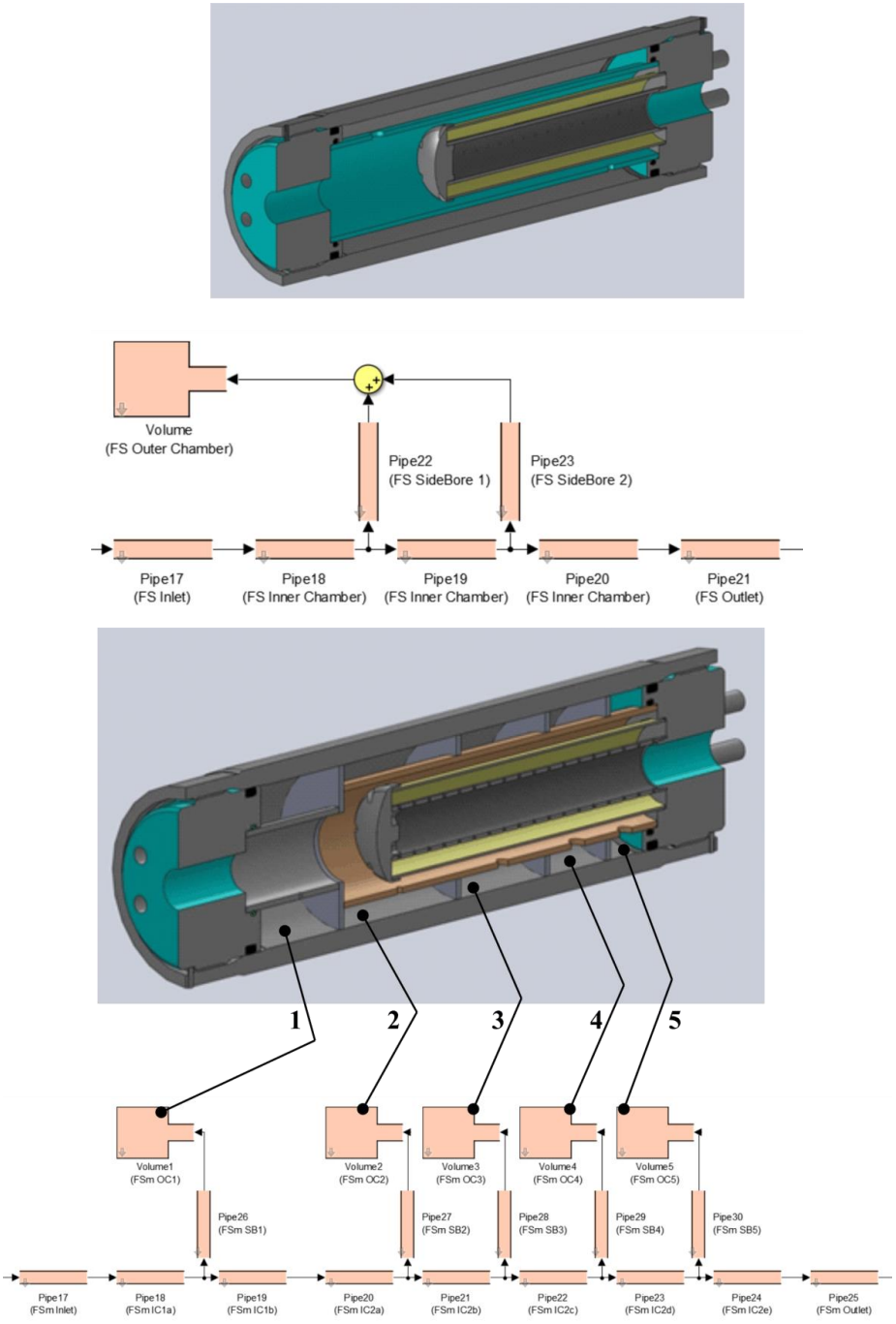


Figure 9: Filter-Silencer PRASP model with one Helmholtz chamber and two bores (top) and five Helmholtz chambers and five bores (bottom) - figures taken from [6]

5. ADDITIVE MANUFACTURING

During the R&D programs ELENA [3] and SINHUS [4] HYDAC produced first samples of the dampeners by additive manufacturing. The first sample was a Helmholtz Resonator which was manufactured in titanium by additive manufacturing and high strength corrosion resistant steel (CRES) by conventional processes (electron beam welding).



Figure 8: Helmholtz Resonator - conventional and additive manufacturing

The sample from CRES showed a mass of 0.400 kg compared with the titanium one of 0.230 kg. The material costs for both are similar, because the mass of the steel bar used for the conventional design is much higher than the amount of powder used in the stereolithography process (STL).

The main advantage of additive manufacturing is the chance to choose complex designs, which cannot be produced by conventional processes or only with an excessive effort.

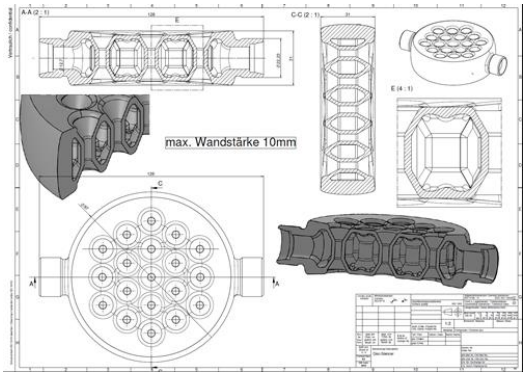


Figure 9: Disc Silencer – produced by additive manufacturing (STL)

The increased productivity of the STL process combined with a high process stability and a reduction of costs for the metal powders leads to a commercially attractive process opening the field for completely new designs.

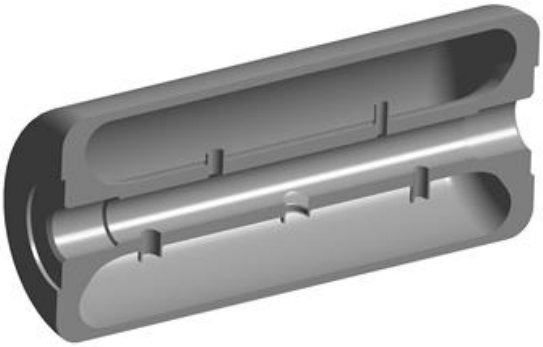


Figure 10: Broadband Silencer – produced by additive manufacturing (STL)

6. SUMMARY AND CONCLUSION

HYDAC has more than 20 years of experience using silencer technology in different hydraulic circuits and applications, mostly in industrial and mobile applications. This high reliable and energy efficient passive damping technology has been adapted to the requirements in aviation applications, e.g. high efficient power packages (HEPP). Low maintenance efforts lead to a minimum of service life costs. Dispensing a sealing system and without the necessity of gases there is no wear and no contamination of the hydraulic fluid and the corresponding hydraulic system. The different hydraulic connection options for the used inner free diameters of the system allow a minimization of additional pressure drops.

The simulation modelling of the complex damping behaviour of Multi-Helmholtz Resonators is still ongoing research due to its interaction phenomena. Further variation options for the neck lengths of the insertion tube bores and the corresponding single Helmholtz Resonators will be given by the integration and investigation of ALM-manufactured insertion tube designs.

REFERENCES

[1] Waitschat, A. et al. (2014): Hydraulic System Noise Attenuators for Aircraft Application: Study of a “Disc-Silencer” Design, Recent Advances in Aerospace Actuation Systems and Components, Toulouse, France

[2] Hoffmann, D. (1976): Die Dämpfung von Flüssigkeitsschwingungen in Ölhydraulikleitungen, VDI Forschungsheft 575, Lübeck

[3] Kloft, P. (2015): Final Report, Erweiterte Luftfahrtsysteme für effektive neue Architekturen (ELENA) - Hydraulic Noise Reduction (HNR) – LuFo IV https://www.tib.eu/de/suchen/id/TIBKAT:863749372/Erweiterte-Luftfahrtsysteme-f%C3%BCr-Effektive-Neue?cHash=75ed5e2964dfd6abeebc2b5b81b13eb5

[4] Kloft, P. (2020): Final Report, Hydraulische System Innovationen (HydI) – Silent Hydraulics (SINHUS) – LuFo V https://www.tib.eu/de/suchen/id/TIBKAT:1744590966/Erweiterte-Luftfahrtsysteme-f%C3%BCr-Hydraulic-System?cHash=92b723c424aa0dff15fd5bf773b20a34

[5] Sander, M. (2023): Vermessung Breitband-Pulsationsdämpfer, analysis report

[6] Waitschat, A. (2016): Study of an Inline Silencer with Integrated Filter-Element & Multi-Helmholtz-

Resonator-Inlays, Technical Report, Issue 01, FST-PB-2016-122

[7] Sander, M. (2023): Vermessung Breitband-Pulsationsdämpfer, analysis report

Chapter 17

Simulation

AN APPROACH TO THE EVALUATION OF THE ENERGY EFFICIENCY OF MACHINES BASED ON DIGITAL TWINS AND SIMULATION METHODS

Rüdiger Kampfmann[1]*, Nils Menager[1], Thomas Ehehalt[1], Michael Liedhegener[1]

[1]*Bosch Rexroth AG, Zum Eisengiesser 1, 97816 Lohr am Main, Germany*

* Corresponding author: Tel.: +49 9352 18-6415; E-mail address: Ruediger.Kampfmann@boschrexroth.de

ABSTRACT

This paper presents an approach to evaluating the energy efficiency of hydraulic machines using digital twins and simulation methods. It emphasizes the importance of considering the entire life cycle of a product and minimizing power consumption of a hydraulic system during its development. Additionally, sustainability activities at Bosch Rexroth are presented and upcoming legal requirements are discussed. By utilizing digital twins, which contain simulation models, the dynamic and efficiency behavior of the machines can be investigated. Through a visualization of energy flow based on simulation results, opportunities for energy savings are identified. Overall, this contribution showcases the benefits of digital twins in optimizing energy consumption and promoting sustainability in the industrial sector.

Keywords: System simulation, energy efficiency, sustainability, carbon footprint

1. INTRODUCTION

Sustainability has been gaining in importance in all areas of society for several years. Therefore, OEMs, machine operators, government and society expect extensive activities from machinery suppliers. Bosch Rexroth has taken this challenge and pursues a holistic approach from cradle to grave. This means that the whole life cycle of a product is considered, i.e., from raw material mining up to recycling. Therefore, tools for the quantification of sustainability measures like product carbon footprint were developed and are used during product development and manufacturing.

In the industrial sector, the operating phase of a machine is of particular importance, as it has a massive impact on the CO2 emissions, whereas in the sector of consumer goods, it is often the manufacturing phase that is more significant. Hence, the power consumption of machines used in industrial production should be already considered and minimized during development. This applies especially to hydraulic machines such as presses and injection molding machines, which often have a high energy consumption.

However, a component-based approach alone is not meaningful. Instead, the interaction of the entire machinery must be considered, and the energy consumption is strongly dependent on the used cycle. Without the usage of digital twins, the energy consumption can hardly be calculated before commissioning, especially since norm cycles are not always defined. Therefore, Bosch Rexroth has recognized this demand and provides digital twins containing simulation models for many products, which allow to investigate the dynamic and efficiency behavior. An automatic visualization of the energy flow based on the simulation results helps to gain further insights and to uncover savings opportunities. The machinery builder can compare different machinery setups regarding energy

consumption and thus find an optimal setup.

In this contribution, first the current societal and legal framework conditions regarding sustainability and energy efficiency are examined. Subsequently, the possibilities of energy-efficient drives from Bosch Rexroth are presented, and the necessity of a simulation-based approach is motivated. Afterwards the Rexroth Simulation Library BRSL is introduced, which provides a variety of digital twins for Rexroth products. Following this, an application is presented, which allows a clear visualization of the system's energy flows. Subsequently, the capabilities of the simulation-based approach for the energetic analysis are demonstrated using a practical example, namely the retrofit of a block press. Therefore, the importance of visualizing the energy flow is particularly emphasized. Finally, the main contents of the article are summarized, and an outlook on future activities is given.

2. LEGAL ASPECTS REGARDING SUSTAINABILITY AND ENERGY EFFICIENCY

Companies are faced with expectations regarding sustainability from various sides. Many investors demand compliance with ESG (Environment, Social and Governance) criteria. In addition to good working conditions, employees seek for a meaningful corporate purpose. Achieving the climate targets requires significant efforts to reduce energy consumption. In the current energy price environment efficiency increases are strictly necessary in order to maintain competitiveness. End customers are becoming increasingly responsible and consider the exact ecological footprint of a product in their purchasing decisions. Summing it up, the subject area of sustainability covers a huge variety of aspects that are important in mechanical and plant engineering. Among others these are resource efficiency, safety and health, emission reduction, energy efficiency and ethical aspects. To provide a comprehensive overview of the entire field of sustainability and all the legal framework conditions would go beyond the scope of this article. Therefore, only the most important points for the fluid industry in Germany from the authors' point of view are discussed below.

A crucial factor in gaining an overview of this multitude of aspects is a systematic approach. Therefore, the creation of a Life Cycle Assessment (LCA) in accordance with DIN EN ISO 14040 [1] yields a good starting point. The norm specifies that an LCA should consider the entire life cycle of a product or service, from the extraction of raw materials through production and use to disposal. The aim is to assess and reduce environmental impacts. The standard serves as a guide for companies and organizations to improve their environmental performance and make more sustainable decisions. This involves an input and output analysis of value streams. One of these value streams is, for example, energy consumption. Bosch Rexroth has developed internal tools for this purpose, in order to display the resource expenditure and thus the ecological footprint as accurately as possible from cradle to gate.

The use of environmentally hazardous substances is often prohibited or restricted by law. For example, the currently discussed ban of PFAS materials, which are very important for sealing technology. In terms of total CO2 emissions, the operating phase is more important than the production phase for industrial equipment. For this reason, there are government funding programs for new systems with higher energy efficiency as well as statutory minimum requirements in mechanical and plant engineering. In Germany, for example, the retrofit of hydraulic systems with energy-efficient variable-speed drives could be subsidized by Federal Office of Economics and Export Control [2].

The most important legal requirements result from the eco-design directive 2009/125/EC of the European Union [3]. This directive is intended to ensure that energy related products on the European market meet certain minimum energy efficiency requirements. To this end, certain product groups are defined for which regulations are successively being established.

The requirements arising from this regulation are steadily increasing. Certain requirements for fluid power systems are not yet specified. Therefore, the eco-design regulation (EU) 2019/1781 [4] for electric drives is the most important for fluid industry, since most hydraulic pumps within industry hydraulics are driven by asynchronous drives. In addition to the requirement of efficiency class IE4 for direct-fed asynchronous motors, the current version also defines minimum requirements for asynchronous motors in frequency converter operation. It is conceivable that in future the system limits will be defined more broadly and that the hydraulic pump will also be considered. This would mean that like the product class compressor, requirements would also be made on the efficiency of a hydraulic power unit. In addition to regulations, the eco-design directive also allows self-regulation if the sector targets are met.

A component-based approach to the evaluation of energy efficiency is not considered appropriate in the field of fluid technology. For example, if a very efficient hydraulic power unit is used in a valve-controlled system with very high throttling losses, the overall efficiency is still low. Another challenging aspect of the overall system analysis is that there are only defined standard cycles for a few machine types and the actual energy consumption also depends on the used cycle. For injection molding machines characteristic cycles were defined with EUROMAP60 [5]. But there is no standard cycle that applies to all hydraulic systems, such as the WLTP in the automotive industry, due to the great variance and different applications.

Hence, neither a component-based approach nor absolute efficiency requirements at system level are appropriate. However, a mandatory energy assessment of the overall system could lead to the required energy savings. This could be carried out on a real test system or virtually with the usage of digital twins.

3. ENERGY EFFICIENT DRIVES FROM BOSCH REXROTH

The classification of the basic circuits of electrohydraulic drives according to Backé offers a systematic approach to the different drive technologies of the fluid industry [6]. For this purpose, a differentiation is made between systems with impressed volume flow and systems with impressed pressure on the one hand and resistance or displacement controls on the other hand.

All these four drive concepts have their system-related advantages and disadvantages and their respective application purpose. For example, secondary controlled drives are usually used for winches, whereas throttle-controlled drives are used for injection molding machines. In general, resistance-controlled drives offer higher dynamics compared to displacement-controlled drives, but also lower energy efficiency. As a fullliner, Bosch Rexroth offers a wide range of products that can be combined to create energy-efficient hydraulic drives.

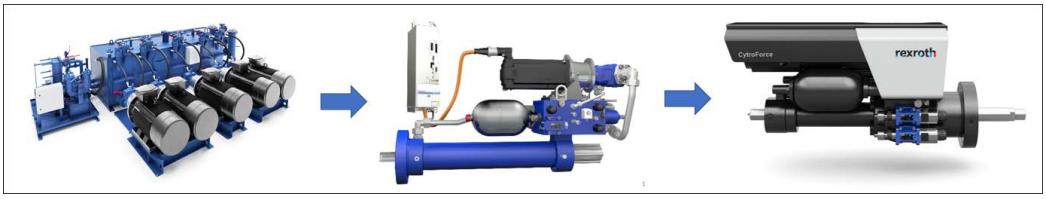


Figure 1: Technological development in hydraulic drives

The long-term trend that can be observed in the fluid industry is shown in **Figure 1** above. Development began with central hydraulic power units, which provide the power centrally for all consumers. The next stage of development was to provide on-demand volume flow directly at the actuator. Finally, this led to highly integrated compact actuators, which can hardly be distinguished

from an electric actuator by their interface.

Furthermore, various new approaches have also emerged from the academic environment to increase energy efficiency, for example high-pressure hydraulics, digital hydraulics, or dissolved control edges.

However, detailed energy considerations for specific industrial applications go far beyond the theoretical analysis of drive concepts. For this purpose, different product classes such as asynchronous motors or synchronous motors or different types of pumps must be compared with each other down to individual details such as series and nominal sizes. Ultimately, this means that for the real energy assessment of the overall system, the analysis must be carried out down to the individual components used.

Regarding energy efficiency, two questions usually need to be answered when retrofitting an existing system or developing a new machine: Does the new drive concept meet the dynamic requirements? How big are the energy savings compared to the existing concept? This evaluation and cross-comparison between two drive systems is only possible either with real prototypes or with the help of simulation. Therefore, Bosch Rexroth has decided to start the development of a simulation library with the aim of providing digital twins for all relevant products.

4. BRSL – REXROTH SIMULATION LIBRARY

4.1. Overview

The Rexroth Simulation Library BRSL is a complete library for analyzing the static, dynamic and kinematic behavior of machines and systems. It allows the creation of a virtual representation of a system based on so-called digital twins of the components. The library contains both generic components, which can represent the behavior of components from different manufacturers using data sheets, and pre-parameterized models of Bosch Rexroth products. These pre-parameterized models are ready to use by default. They are fully validated by Bosch Rexroth and selected directly via the type code (see **Figure 2**). On the one hand, this has the advantage of eliminating the time required for parameterization, while at the same time preventing incorrect parameterization.

The BRSL is developed in the tool-independent modeling language Modelica [7], which results in direct compatibility with several widely used simulation tools available on the market. Using the FMI (Functional Mockup Interface) modeling standard [8], BRSL models can also be easily integrated into non-Modelica tools. The BRSL is officially released as a product, i.e. it is available to all users who want to benefit from the numerous well-known advantages of simulation usage.

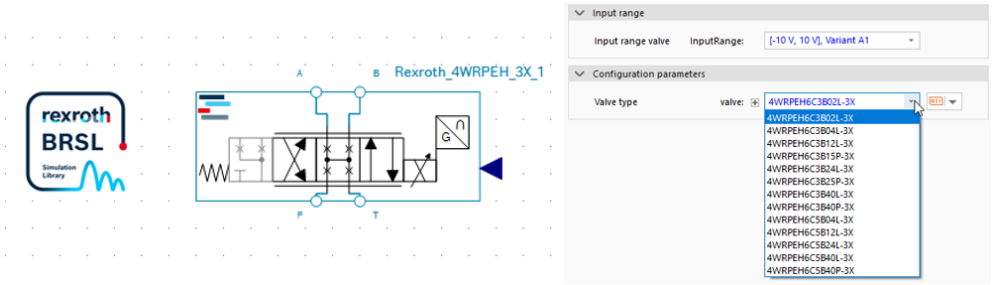


Figure 2: Parameter window for pre-parameterized Rexroth 4WRPEH valve model

4.2. Requirements for analyses of energy efficiency and carbon footprint

The BRSL models include all the physical effects required for the use in a 1D dynamic system simulation. This includes not only static operating limits of the components, but also dynamic effects such as friction and damping, pressure build-up behavior, pressure- and load-dependent switching times as well as frequency responses.

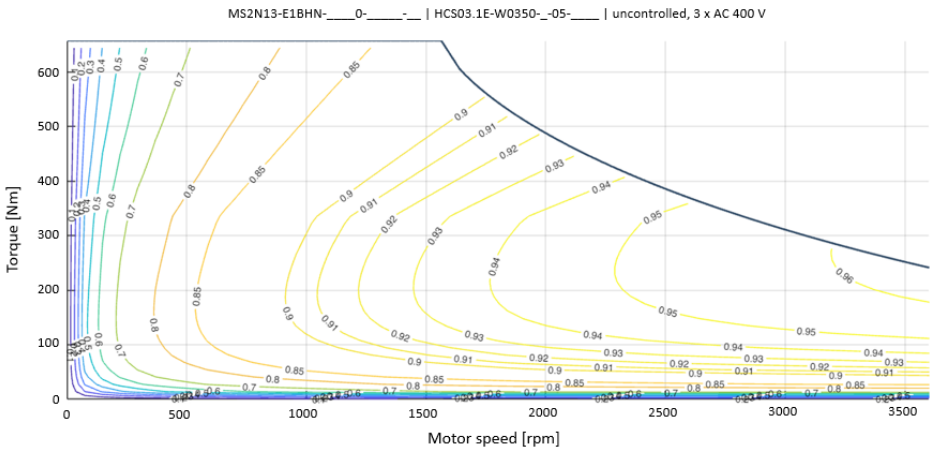


Figure 3: Efficiency characteristics for Rexroth MS2N13 motor with Rexroth HCS03 frequency converter over speed and torque

Of particular importance for energy efficiency and carbon footprint analyses is the inclusion of component losses. For this purpose, the individual product models within the BRSL contain detailed efficiency characteristics that include the efficiency as a function of relevant influencing variables. In the case of a hydraulic variable displacement pump, for example, these are pressure, speed and swivel angle. For an electric motor, however, these are speed and torque (see **Figure 3**). This includes all quadrants, so that recuperation effects can also be considered. The hydraulic losses within the components are also calculated for the generic hydraulic components; these can be calculated from the product of pressure loss and volume flow.

5. A MATLAB APPLICATION FOR ENERGY VISUALIZATION

5.1. Overview

Although the losses of the individual components are already included in the simulation models and hence in the simulation results, it is very difficult to analyze the energy household directly within the simulation tool. For this purpose, Bosch Rexroth has developed a MATLAB WebApp that clearly visualizes the energy losses of both the overall system and the contained components separately. The application allows to analyze one single system and to identify critical losses, but it is also possible to compare two systems with each other. In this case, energy savings and carbon footprint reductions become clearly visible.

As basis for the visualizations, simulation result files in CSV format are used. Currently, the tool is specifically tailored to the use of the BRSL, however, an extension to simulation models of other tools is feasible. For the analyses presented in this paper, simulations were carried out using models based on BRSL within the simulation software SimulationX by ESI Group.

5.2. Description of the different views of the application

After the start, a new empty project is automatically created. If it is desired to continue working on a previous analysis, this project can be loaded by using the corresponding button at the top of the application. In any case, the tool will start with the main project view, which allows to enter project-specific data. In addition to the project name, this also includes information on the project participants and the contact persons involved. The next step is to load the simulation results. It is possible to either load one file or – in case that two systems shall be compared – two different result files.

For the following analysis, the tool contains several views, each of which allows a different perspective on the data. The first view provides an overview of the input and output power of the individual components. These are displayed both over time and as a mean power value. This view, using the data from the example used in this paper, is shown in **Figure 6** in the Examples Section. In addition to the power, the energy consumed or emitted can also be displayed. Similarly, both the energy over time for the individual components and the mean value for a component are shown.

Furthermore, it is possible to display the energy sources and energy sinks separately in the form of pie charts (see **Figure 7** in the Example Section). This enables to determine the share of each component in the overall energy household of the system or machine. Another option is to display the data separately by domain (mechanic linear, mechanic rotary, hydraulic, electric). This emphasizes the system's energy converters. At the same time, it is possible to clearly show the resulting losses.

Finally, the tool contains a quantitative evaluation of the system, which, however, requires further input data. Firstly, the number of cycles per hour, the number of hours the machine runs per day and the number of days of operation per year must be specified. Secondly, it is necessary to specify the price per kWh of energy and the CO2 emissions in grams per kWh. Based on this, the annual energy consumption, the resulting costs and the CO2 emissions are calculated. If two systems are compared, the values just mentioned are calculated for both systems and the difference value, i.e. the saving, is shown. When also providing the investment costs for a retrofit with conversion of the system, for example, the ROI (Return on Investment) is derived directly from the savings. This view is shown in **Figure 8** in the Example Section.

To summarize, the application allows it to easily evaluate one or more machines in terms of energy consumption and, if necessary, to compare them with each other to quantify the savings potential.

6. APPLICATION TO A PRACTICAL EXAMPLE

This article presents the application of simulation and energy analysis for the economic evaluation of a retrofit of block press. This is a very common case in practice: there is already an existing system and the question arises whether it is economical converting the system to a more energy efficient one. To answer this, the investment costs must be compared with the financial savings resulting from higher energy efficiency. Therefore, however, it is first necessary to know the exact energy consumption of both systems, which can be determined using simulations. The following example is derived from a real industrial project.

6.1. Reference system – Valve-Driven System

The reference system is a press with a conventional valve drive. A schematic diagram is shown in **Figure 4**. It should be noted that this is a reduced view for the sake of clarity and that the simulation model used is more detailed. From a high-level perspective, the model consists of a hydraulic power unit, a valve for controlling the volume flow, a double-rod cylinder, a process model and a corresponding control algorithm.

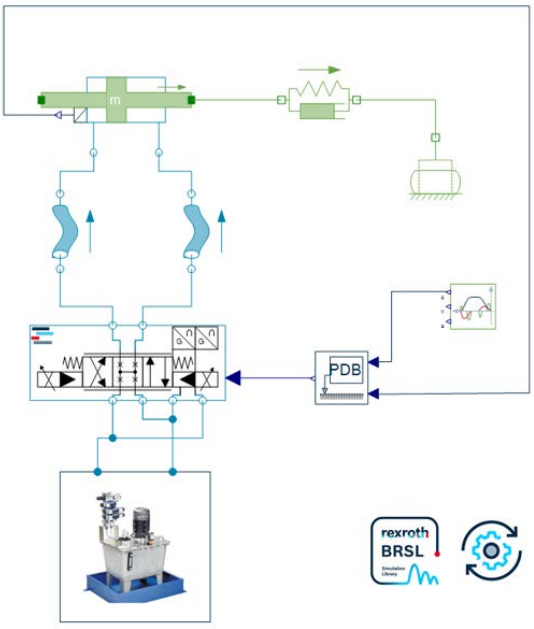


Figure 4: Schematic view of the reference system (simplified illustration)

The hydraulic power unit consists of a direct-fed asynchronous motor including efficiency characteristics, an axial piston pump with swivel angle control (Rexroth A4VSO), a hydraulic accumulator and a check valve. The process consists of pressing the blocks. To do this, the cylinder extends to an end position within a defined time and then returns to the start position. Including the waiting time for inserting the new material, one cycle takes approximately 4.5 seconds.

This system is suboptimal from an energy point of view, as the hydraulic power unit constantly provides a pressure of 200 bar, regardless of what is required by the process. The pressure that is not required is throttled at the valve. For this reason, the valve-driven system could be replaced by a more energy-efficient cylinder direct drive.

6.2. Optimized system – Cylinder Direct Drive

To optimize the system, the valve drive is replaced by a frequency-controlled pump drive in a closed circuit system. A simplified, schematic view of the optimized system is shown in **Figure 5**.

The drive consists of a synchronous motor (Rexroth MS2N13) with a suitable converter (Rexroth HMS01), which drives an axial piston pump for the closed circuit (Rexroth A4CSG). All components include efficiency characteristics. Two boost pumps are required in addition to the axial piston pump. Boost pump A is used for the pressure supply to the swiveling unit of the axial piston pump (not shown in simplified model view), boost pump B provides the necessary supply pressure for the closed loop system. Hence, the pump directly drives the double-rod cylinder, which interacts with the process. The cylinder, the process model and the control command values are identical to the reference model. Finally, the model also includes the necessary equipment for the closed circuit and hydraulic piping.

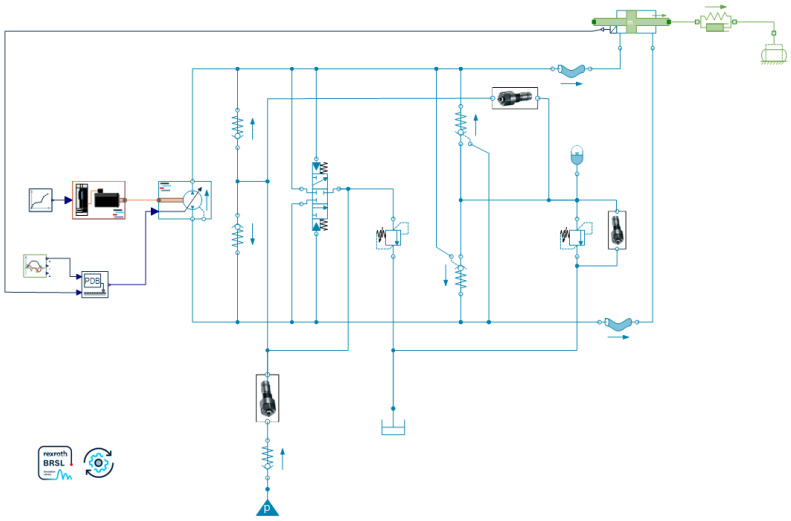


Figure 5: Schematic view of the optimized system (simplified illustration)

6.3. Results

The two systems described above, i.e. the reference system and the optimized system, were simulated with SimulationX and the results were saved in CSV format. Based on the simulation results, it was first verified that the requirements for accuracy and dynamics are fulfilled also in the optimized system setup.

The result files were then loaded into the MATLAB application for analyzing the energy household of the systems.

Power view

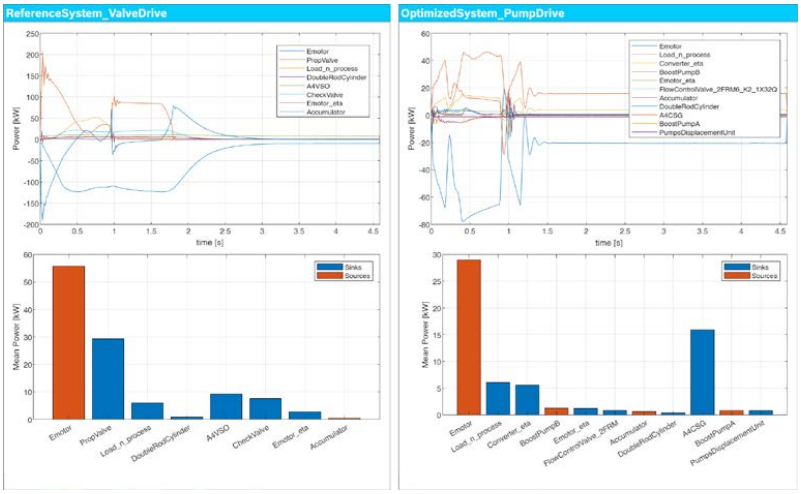


Figure 6: Power input and output of components (left: reference system, right: optimized system); upper graphs show the consumption over time, lower charts contain mean power values of components

The evaluation of the outputs (see **Figure 6**) shows that the average power of the electric motor, which is the main energy source in both systems, can be reduced from 58 kW to 29 kW as a result of the retrofit. This corresponds to a reduction of 50%. In the optimized system compared to the reference system, however, two additional energy sources (with relatively small impact) are available with the two boost pumps. The main consumer in the reference system is the proportional valve, with a mean power consumption of 29 kW. In the optimized system, the pump (A4CSG) has the highest share, but with 16 kW mean power this is below the consumption of the valve in the reference system.

Splitting by energy sources and energy sinks

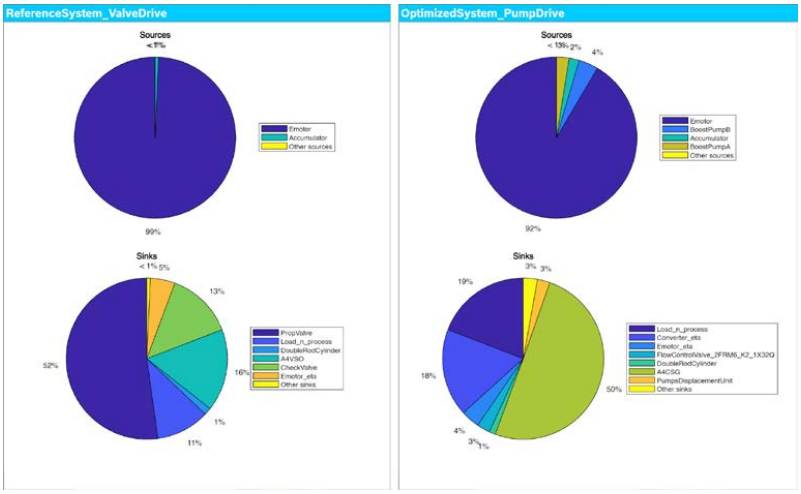


Figure 7: Visualization of energy sources and energy sinks using pie charts (left: reference system, right: optimized system)

Figure 7 shows the energy share of the individual components, broken down into energy sources and energy sinks. It becomes again visible that the electric motor is the main source of energy in both systems. In the reference system, the motor accounts for 99% of all energy sources, while in the optimized system it accounts for 92% (the two boost pumps account for 3% and 4% respectively).

In the case of energy sinks, it is desirable that the process has the highest possible share and that there are as few other, undesirable energy sinks as possible. The improvement in the optimized system is clearly visible in the diagrams. In the reference system, the process (Load_n_process) has a share of 11%, whereas in the optimized system it has increased to 19%. As previously mentioned, the main energy sinks for the two systems are the proportional valve and the axial piston pump, respectively. Both have a share of around 50% each.

It should be emphasized that this view does not show absolute values, but the relative share of energy sources or sinks in the respective household. The fact that the optimized system consumes less overall power was already apparent in the power view described above. This also means that a proportion of 50% in the optimized system means less energy consumption in absolute terms than a proportion of 50% in the reference system. This becomes also clear in the quantitative evaluation of both systems, which is described in the following section.

Quantitative evaluation of the savings

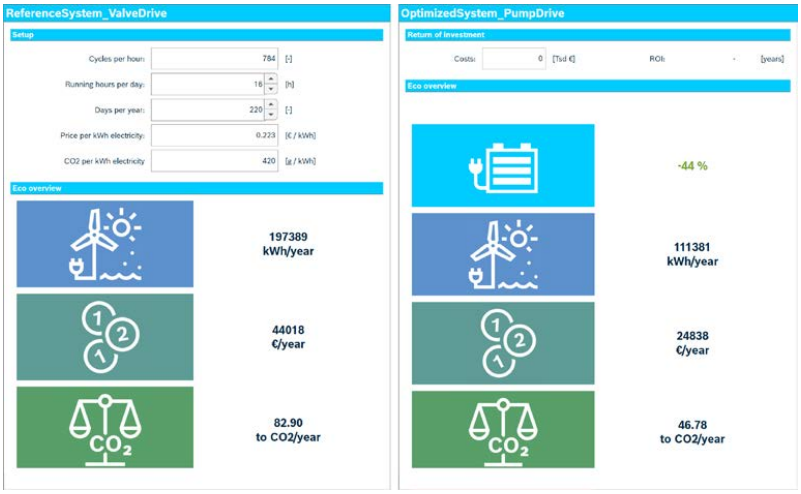


Figure 8: Quantitative comparison of the two systems with presentation of the savings (left: reference system, right: optimized system)

An overview of the quantitative comparison of the two systems is shown in **Figure 8**. In the upper left part of the figure, definitions regarding the machine runtime are specified. The process used in this example has a duration of slightly more than 4.5 seconds per cycle. Assuming continuous production, this results in 784 cycles per hour. It is further assumed that production takes place in two-shift operation (16 hours per day) and that the machine operates 220 days per year. The assumed price of 22.3 ct per kWh of energy and the CO2 emissions of 420 grams per kWh are also set there.

Based on the simulated energy required for the systems and the aforementioned data, the retrofit results in energy savings of 86008 kWh per year. This corresponds to yearly financial savings of 19180 €. Furthermore, 36.12 tonnes of CO2 per year can be saved. The total energy saving as a result of the retrofit is 44 %. These results correspond to the values measured on the real systems before and after the retrofit.

If the investment costs for the retrofit are additionally entered in the application, it is possible to calculate directly how fast the conversion will amortise financially. However, the economic view in relation to the ROI is not considered in this example.

7. SUMMARY AND OUTLOOK

7.1. Summary

In this contribution, the importance of increasing the energy efficiency of hydraulic systems and how this can be achieved through the usage of digital twins was discussed. The general trend towards sustainability, which is also being demanded by more and more customers, will confront industrial companies with significant challenges in the future. In addition, legal requirements are constantly increasing. A Life Cycle Assessment according to DIN EN ISO 14040 offers a systematic approach and important legal framework conditions are contained in eco-design directive 2009/125/EC of the European Union. This includes requirements for the energy efficiency of energy-related products. It was argued that a general component-based approach to the energy efficiency of hydraulic systems is not appropriate. Instead, when assessing the energy efficiency of a system, it is necessary to

consider the interaction of all components, particularly taking into account the actual loads due to the machine cycle. Simulation methods are a suitable option for a simple and cost-effective realisation of this. With the simulation library BRSL, Bosch Rexroth provides digital twins for Bosch Rexroth components that enable virtual energy analyses without the use of real prototypes. Furthermore, a MATLAB application that evaluates energy flow based on simulation results was presented. The functionality of this approach, consisting of simulations with BRSL components and a subsequent evaluation with the MATLAB application, was demonstrated using a real industrial example, the retrofit of a block press. The quality of the simulated results was compared with real measurements before and after the retrofit and could be confirmed.

7.2. Outlook

In future, the presented approach will be used for other industrial projects at Bosch Rexroth to increase the energy efficiency of hydraulic systems. To enable this, the development of the BRSL simulation library will be continued in order to close existing gaps in the product portfolio and to be able to offer a digital twin for new products directly. Currently, the BRSL is only offered as one package containing all components and only Modelica-based tools are supported. To ensure wider availability, Bosch Rexroth is actively involved in the IDTA (Industrial Digital Twin Association) working group to enable tool-independent provision of component models via the simulation subaspect of the Asset Administration Shell [9].

However, hydraulic systems do not only contain components from Bosch Rexroth. Competitors are therefore also invited to provide simulation models for their products in order to offer machine builders a supplier-independent assessment of the energy consumption of systems.

The topic of sustainability is of particular importance to Bosch Rexroth. This is why Bosch Rexroth is continuously working on innovative methods for optimizing the energy efficiency of machines and systems. To support this, also new tools and methods such as the simulative approach for evaluating the energy efficiency of machines described in this contribution will be continuously developed.

REFERENCES

[1] DIN EN ISO 14040: 2021-02. Environmental management - Life cycle assessment - Principles and framework. Berlin: Beuth Verlag

[2] Bundesamt für Wirtschaft und Ausfuhrkontrolle (2024) Energie- und Ressourceneffizienz in der Wirtschaft. https://www.bafa.de/DE/Energie/Energieeffizienz/Energieeffizienz_und_Prozesswaerme/energieeffizienz_und_prozesswaerme_node.html. Accessed 11 Jan 2024

[3] Directive 2009/125/EC of the European Parliament and of the Council of 21 October 2009. Establishing a framework for the setting of ecodesign requirements for energy-related products. Official Journal of the European Union, L 285, 10-35.

[4] Regulation (EU) 2019/1781 of the European Parliament and of the Council of 1 October 2019. Ecodesign requirements for electric motors and variable speed drives. Official Journal of the European Union, L 272, 1-32.

[5] EUROMAP (2013). EUROMAP60 – Injection Moulding Machines – Energy Efficiency. Technical Recommendations. https://www.euromap.org/technical-issues/technical-recommendations. Accessed 11 Jan 2024

[6] Findeisen D, Hellduser S (2015). Ölhydraulik: Handbuch der hydraulischen Antriebe und Steuerungen. Springer Vieweg

[7] Fritzson P, Bachmann B, Moudgalya K et al (2018). Introduction to Modelica with Examples in Modeling, Technology, and Applications. Linköping University Electronic Press

[8] Junghanns A, Blochwitz T, Bertsch C et al (2021). The Functional Mock-up Interface 3.0 – New Features Enabling New Applications. In: Proceedings of the 14th International Modelica Conference, September 20-24, 2021, Linköping, Sweden. https://ecp.ep.liu.se/index.php/modelica/article/view/178/138. Accessed 11 Jan 2024

[9] Industrial Digital Twin Association (2022). IDTA 02005-1-0 Provision of Simulation Models – Specification. https://industrialdigitaltwin.org/wp-content/uploads/2023/01/IDTA-02005-1-0_Submodel_ProvisionOfSimulationModels.pdf. Accessed 11 Jan 2024

A NOVEL SAAS DEVELOPMENT PLATFORM FOR FLUID POWER STANDARD DRIVES

Heiko Baum[1]*

[1]*FLUIDON Gesellschaft für Fluidtechnik mbH, Jülicher Straße 338a, 52070 Aachen, Germany*

* Corresponding author: Tel.: +49 241 980 935 61; E-mail address: heiko.baum@fluidon.com

ABSTRACT

Today, it is state of the art to use 0D/1D simulations in the development of complex fluid technology drives, as this is the only way to evaluate the dynamic interactions of the system components. However, when modifying existing drives, developers often refrain from simulative validation because they consider the changes to be minor and/or rely on the positive experiences of the past. Consequently, design errors are only discovered in practical use, leading to production disruptions and costly troubleshooting.

A new type of SaaS development platform closes this gap and is aimed at companies with limited capacities and budgets. The platform enables companies to engage engineering service providers in creating customized design workflows. The paper illustrates the platform's application in a mobile hydraulic drive example, detailing the orchestration of pre- and post-processing tasks through a web browser interface. The familiar Excel spreadsheet used for static calculations continues to serve for parameterization source, maintaining the user's established design process while leveraging the precision of 0D/1D simulation. The simulation results are automatically converted into a format familiar to users, either as an Excel spreadsheet or a PowerPoint presentation for documentation and sales support.

Keywords: SaaS development platform, orchestrated design workflow, FMU, Spreadsheet driven simulation, automatic result reporting

1. INTRODUCTION

Today, the use of 0D/1D simulation is largely standard in the development of complex fluid power drives. This is the only way to test and evaluate the dynamic interactions of all installed components as a system. The buzzwords "digital twin" and "virtual commissioning" are not on everyone's lips for nothing. The situation is different with fluid power "standard drives". These drives, which are manufactured in small batches, can be found in mechanical and plant engineering, the raw materials industry or in mobile machinery or agricultural machinery. In terms of their total number, however, these drives form a quantitatively significant group of fluid power drives.

If a standard drive is to be modified at the customer's request or if a new product generation is planned, simulative validation of the changes is generally still not carried out. When asked about the reasons for this, the justification given is that the new design is only a minor modification or scaling of an existing drive for which there is extensive positive practical experience. The project engineer is then often satisfied with static calculations, e.g. in an Excel spreadsheet. However, dynamic interactions between the components are not considered in this way.

Practical experience shows that this is grossly careless. The simple scaling of components or even a change to the piping layout requested by the customer poses considerable risks regarding the dynamic behavior of the fluid power drive. If, for example, components need to be scaled up, their operating points will shift. It must then be considered that the operating behavior of many fluid power

components changes non-linearly when the operating point is shifted. Even if the same components are only spatially arranged differently or the piping is modified in another way, the dynamic interactions between the components change. These influences can only be considered insufficiently or not at all by quasi-static calculations.
If design errors are only recognized during practical use - i.e. too late - the consequences are no less serious than with complex systems. Production lines or machines in which the drives are installed come to a standstill and must be repaired at great expense. A simulation for troubleshooting, which is needed quickly and is therefore expensive, is often the only way to identify the problem and develop a solution.
Many of these problems could be avoided if 0D/1D simulation could be used more often in the design of supposedly simple fluid power drives. The simulation ensures that a new or reconfigured drive meets the customer's technical boundary conditions and requirements.
To achieve this goal, it is necessary to make 0D/1D simulation as accessible as possible for all those users who have yet little experience in simulation. What was missing until now was a development platform specialized for this user group.

2. A FLEXIBLY SCALABLE SAAS DEVELOPMENT PLATFORM

Cube could be just the right platform, as it has been deliberately designed to allow experienced users to prepare pre-configured workflows for the development, testing and marketing of fluid power drives. These workflows can then be used by less experienced users by giving them access to Cube. The workflows release users from routine tasks and allow them to concentrate better on their main task of interpreting the results. The following section outlines the basic procedure.

Each Cube account has a workspace in which users can create projects, either from scratch or by using templates. The most important features of this workspace are outlined in **Figure 1**.

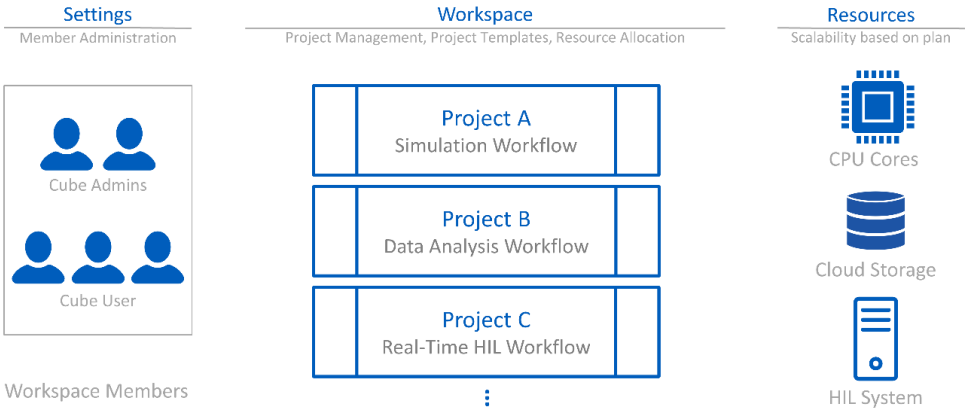


Figure 1: Workspace and Collaboration

In general, a wide range of tasks can be implemented in Cube projects. However, this paper focuses specifically on the presentation of workflows for the repetitive execution of tasks in the field of the simulative design of fluid power drives. For example, as shown in Figure 1, Project A could be a workflow for the simulation of fluid power drives, Project B could be a workflow for processing simulation results or measurement data from the test field, and Project C could be used to configure a simulation model on the Cube RT system required for the development of control systems running a real-time digital twin.
Every Cube user who has been invited to a workspace by an administrator is a member of that workspace. Members can work on their own projects or on a joint project. The workspace area gives

members an overview of the distribution of available resources. Resources are CPU cores, storage space on the cloud platform or Cube RT systems. The resources are generally located in a shared pool and can be allocated to individual projects as required.
As is usual with modern SaaS applications, a Cube account is always linked to a "plan". The plan describes the different pricing or subscription models available to users, what additional resources or additional services cost and what other services or support options are included. The basic "Cube Free" plan, in which a single workspace with a single project can be created, is free of charge.

Figure 2 summarizes the important benefits for a company when introducing the Cube SaaS application. A SaaS application is available immediately after the initial registration, which enables companies to introduce the new solution quickly. This also means a lower initial investment compared to traditional software development and deployment, as companies do not need to purchase expensive hardware or software licenses. Instead, they pay monthly or annual subscription fees.
SaaS applications are accessible via the Internet, which means that users can access their account at any time, from different locations and with different end devices. This is an advantage for modern company-specific working time models. In addition, thanks to the intuitive user interface typical of browser applications, users can find their way around the SaaS application quickly and without a great deal of training.

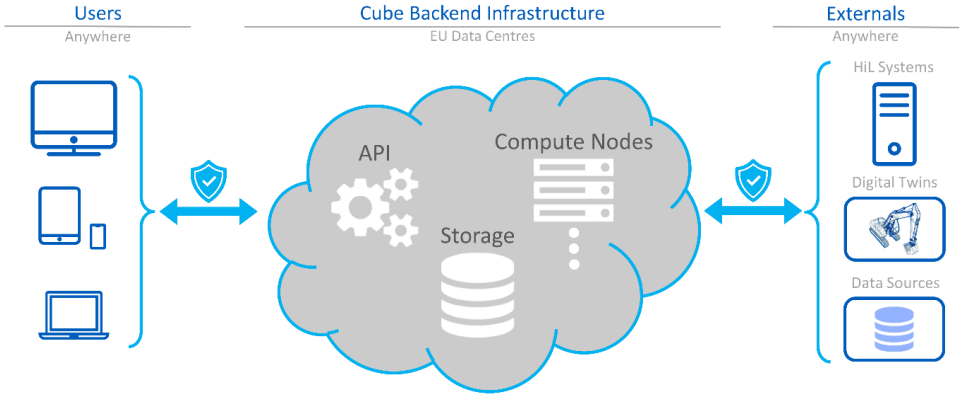


Figure 2: Flexibility, Scalability, and Security

Scalability and high availability are further advantages of SaaS applications (centre section of Figure 2). Storage space or computing power can be added at short notice by changing the plan, allowing companies to react flexibly to increased requirements without having to carry out expensive hardware upgrades.
In addition, SaaS applications guarantee high availability, as the operators of the data centers invest in redundancy and high-availability solutions. SaaS applications are also updated automatically, which relieves companies of the need to take care of maintenance and updates themselves and ensures that its users always have access to the latest functions and security patches.
Another important point in favor of SaaS applications is data security. In the case of Cube, the data is stored exclusively in European data centers. In addition to the security measures of the data centers, all areas of Cube that can be accessed via the Internet connection are secured separately in accordance with the state of the art. This measure provides companies with a level of security that generally goes beyond what would be practical for smaller companies or organizations to implement if they had to provide it themselves.

An interesting additional feature of the Cube development platform is shown on the right-hand side of Figure 2. Cube provides interfaces to access external devices or data sources from the SaaS

applications. This is used, for example, to communicate with Cube's own RT system [1], access measurement data or exchange live data with other cloud applications.

The steps required to execute a workflow on Cube are described qualitatively in **Figure 3**. In preparation, the user has created a project in the workspace, set up a workflow in it, loaded any required data or simulation models and specified the CPU cores required for execution. The prepared project does not yet block any of the available computing resources. Scripts, simulation models and data only take up storage space in the cloud. Only when a compute node on which the workflow is to be executed is created by the user (Figure 3, top left), the specified number of CPU cores is allocated. Once the compute node is ready for use, the project is started.

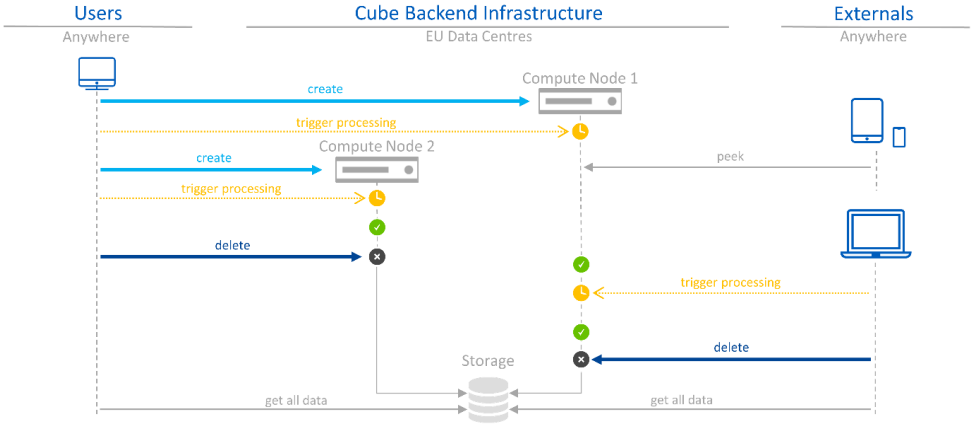


Figure 3: Computing Power on Demand

Both the start of the workflow and the interaction with the running workflow can be carried out using any end device, as illustrated once again in the right-hand section of Figure 3. For example, the tablet or mobile phone is used to check whether the execution has been completed. If necessary, the configuration of the workflow is now updated, new data is loaded into the workflow and a new execution is started.
If the results meet the requirements, the occupied computing resources can be released. To do this, the user deletes the compute node, and the CPU cores are once again available to all workspace members. The data generated during execution is automatically moved from the compute node to the cloud storage, where the user can continue to access it at any time.
If free computing resources are still available in the workspace during the execution of compute node 1, further projects can be executed, as shown in Figure 3 using the example of compute node 2.

3. DEVELOPMENT AS A TEAM ACHIEVEMENT

The possibility for several users to work together on projects in the Cube workspace offers companies the opportunity to utilize contemporary forms of development collaboration. One of these concepts, "development as a team achievement", and the resulting opportunities for companies in product development are presented below.

The concept outlined in **Figure 4** emphasizes the idea that the development of products, services or projects is most effectively achieved through the collaboration and coordinated work of a team rather than individual efforts. The concept thus reflects the reality of many engineering projects and the importance of teamwork in engineering. This is because engineering projects often require experts from different disciplines to work together to solve complex problems and drive technical innovation.

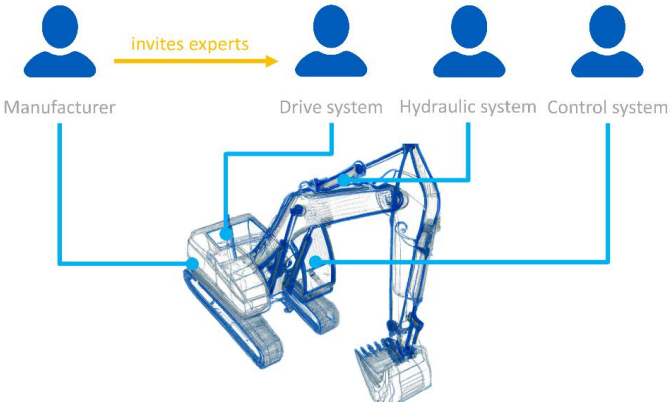


Figure 4: Engineering as a Team Achievement

The Cube development platform promotes this approach by automatically making the company's internal experts, as members of the workspace, members of the team. If necessary, the company can also integrate external service providers into the team by inviting them to join the workspace.

External service providers can leverage the Cube development platform to offer their solutions not only to a single customer but to multiple companies. (**Figure 5**).

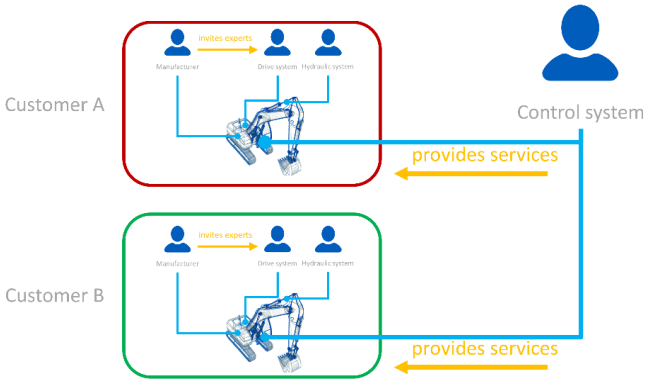


Figure 5: Engineering as a Service

A service provider for control systems can, for example, make his algorithms or simulation model available to companies. Cube provides interfaces to protect the specific expertise of the provider. This enables a business model in which service providers set up workflows, prepare Python modules, provide measurement data or create visualizations for Cube's animation task and make them available to Cube users. This can be done directly within an existing project or as a well-documented project template for the Cube development platform.

4. A SIMULATION WORKFLOW EXAMPLE

As a typical example of a ready-to-use and easily parameterizable Cube workflow, the simulation of a standard mobile hydraulic drive is presented in the following section. Such a hydrostatic drive (**Figure 6**) usually comprises at least one pump and several motors.

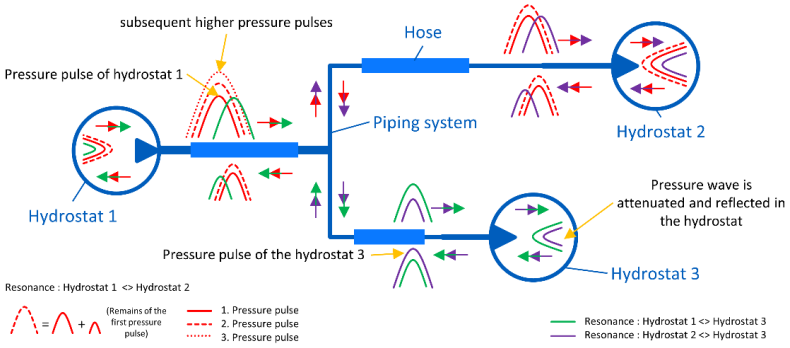


Figure 6: Overview of the pressure oscillation situation of a hydrostatic drive train

A setup with only two motors was selected for the example. The hydrostats are connected by a piping system that, depending on the topology of the vehicle, branches out into several segments of different lengths, which in turn often consist of a combination of steel and hose lines. For the sake of simplicity, only the pressurised part of the piping system is considered.
To operate the drive in the optimum efficiency range or to enable special operating conditions, all hydrostats are adjusted individually. This results in a broadband pulsation excitation of the piping system, which can cause pressure oscillation problems not only between the source and the consumers, but also between the consumers themselves [2].

In the "classic" design of the piping system, it is almost impossible for reasons of time and resources to experimentally find the configuration that leads to the lowest pressure pulsation load on the piping system from the large number of possible variants. This is where the Cube workflow (**Figure 7**) comes into play to determine the best configuration for the piping system.

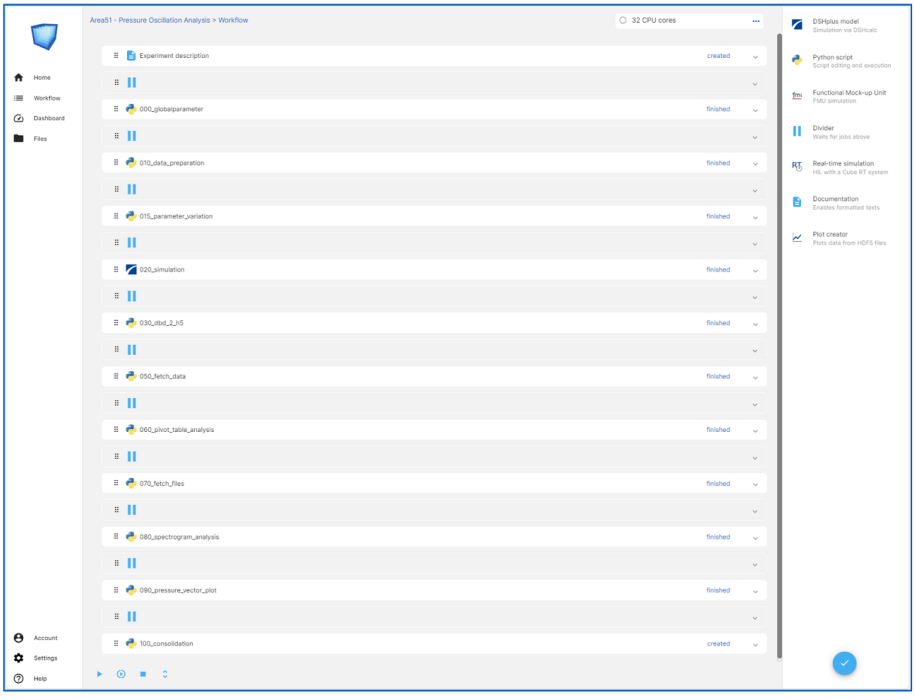


Figure 7: Pressure oscillation workflow of the hydrostatic drive train

The workflow consists of eleven tasks, which are arranged in their logical working order. Except for task *020_simulation*, which includes the simulation model of the hydrostatic drive, all other tasks are described by Python scripts.
Such Python scripts can be freely created by the user, who can also incorporate (install) their own personal Python module library into the project. However, the Cube environment also offers a collection of predefined Python modules that users can integrate into their own scripts.
Task *020_simulation* of the example workflow is a dedicated DSHplus module that can import DSHplus simulation models directly [3]. If this task were replaced by Cube's FMI module, the workflow could run any Linux 64bit Functional Mock-up Unit (FMU) model generated by an FMI-compliant development tool [4]. Of course, the model parameters of such a FMU must be accessible to carry out the following parameter study!

At the beginning of a parameter study, it is necessary to define the limits within which the component parameters are varied and to determine how the results are to be evaluated. **Figure 8** shows an Excel workbook that is set up for this purpose. Excel was chosen as the data input medium to fulfil the requirement that the workflow should also be usable by non-simulation experts.

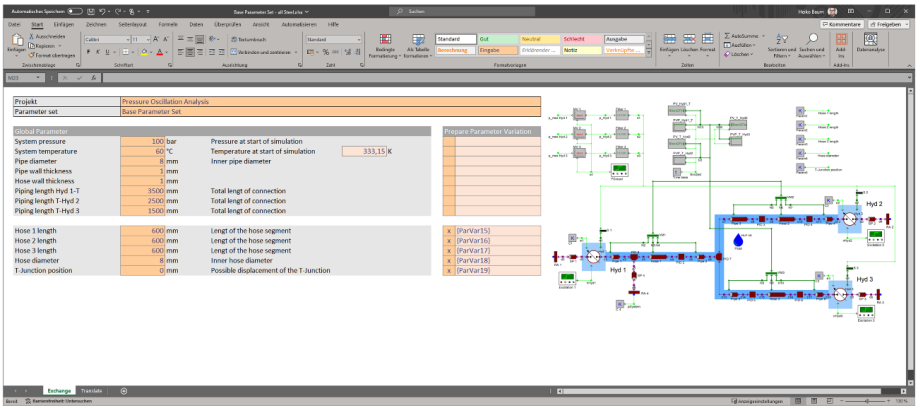


Figure 8: Parameterization of the hydrostatic drive train

For installation space reasons, the overall lengths of the piping system must remain constant. However, the lengths of hoses 1 to 3 can be designed variably between 400 mm and 1000 mm. Formula relationships in the components automatically adjust the lengths of the remaining steel pipes accordingly. Other possible variations include internal hose diameters of 12 mm and 16 mm, as well as the positioning of the T-branch, where the pipe coming from the source (hydrostat 1) to the two consumers (hydrostat 2 and 3) can be moved from -250 mm to +125 mm from its starting position.
The medium pressure of the system is set to 100 bar. All three hydrostats are modelled as nine-piston variable displacement units. This means that not only the basic excitation frequency of the pump is considered, but also higher pump orders as excitation of the pipe system.

Figure 9 shows a screenshot of the Python script of task *015_parameter_variation* responsible for the variation of the design parameter. The source code view of the task is intentionally displayed to illustrate how straightforward it is to setup parameter variations using publicly available Python libraries.
The script lines 33 to 38 are used to enter the parameter range to be varied. The script generates a full-factorial experimental design [5] with 512 parameter sets and stores them directly in the parameter directory of the task *020_simulation*. The simulation module now starts a parallel calculation making use of all available 32 CPU-Cores and thus utilizing the full performance of the computer hardware.

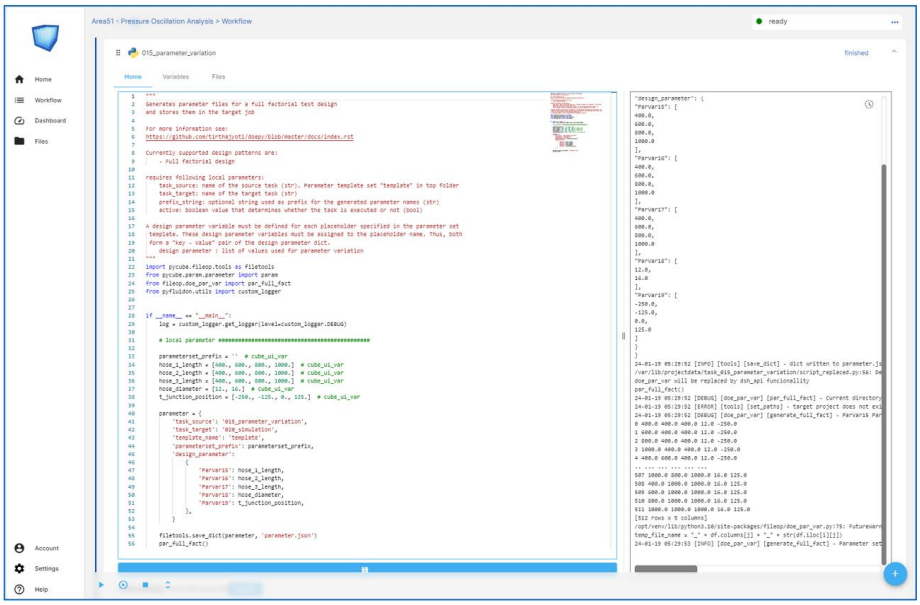


Figure 9: Python script for parameter variation with full factorial experimental design

However, the detailed results of the 512 simulations are initially of secondary interest to the user. First and foremost, the user wants to use the parameter variation to ensure that no disruptive resonance situation occurs in the drive train with the base piping design he has selected and to find out which of the possible line configurations has the lowest pressure pulsation amplitudes. The pressures at the hydrostat flanges serve as criteria for the evaluation.

The extraction of all data relevant for the subsequent analysis from the simulation results is carried out in task *050_fetch_data*. The values are consolidated so that, for example, only the maximum values of the pressures are taken during the simulation run. All values are saved in an Excel file and can therefore be further analysed offline or processed directly in the workflow.
Within the workflow this is done by the Python script of task *060_pivot_table_analysis* (**Figure 10**).

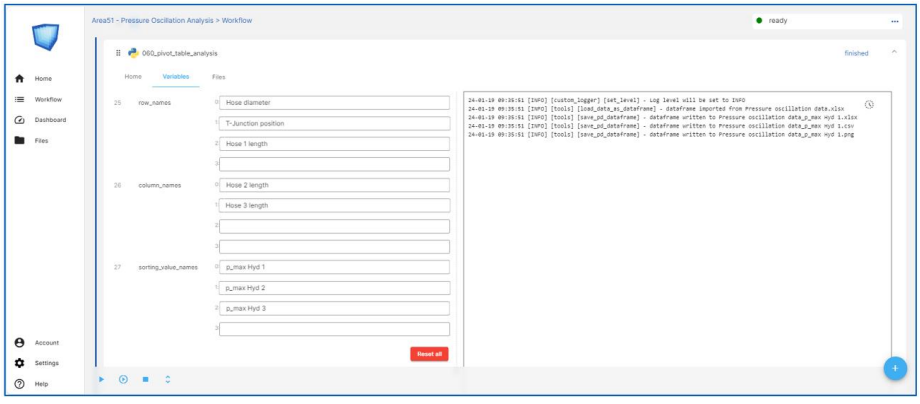


Figure 10: Python script for pivot analysis

Instead of presenting the Python script editor, as done in Figure 9, the left part of the task window now shows the UI (user interface) section, in which the user defines the formatting of the pivot table, which is saved again as an Excel-file for further processing.

Figure 11 presents the resulting pivot table, which is an overview of the pressure amplitudes at hydrostat 3. Colour coding makes it easy to see which pipe system configuration causes higher and which lower pressure values.

		Hose length 1	0,4				0,6				0,8				1			
		Hose length 2	0,4	0,6	0,8	1	0,4	0,6	0,8	1	0,4	0,6	0,8	1	0,4	0,6	0,8	1
Hose diameter	T-Junction pos.	Hose length 3																
0,012	-0,25	0,4	10,79	8,11	7,98	7,05	9,53	8,86	8,06	8,77	11,17	9,61	9,29	9,68	11,45	12,01	11,61	8,43
		0,6	11,25	8,55	7,93	7,26	9,94	8,98	8,28	8,01	11,47	9,88	9,29	9,02	11,04	11,97	11,40	8,93
		0,8	11,62	8,91	8,07	7,75	10,45	9,18	8,33	7,70	11,53	10,04	9,37	8,69	12,22	11,74	11,34	9,22
		1	11,77	9,28	8,18	9,06	10,90	9,49	8,81	8,40	11,52	10,44	9,97	9,31	12,94	12,12	11,87	9,15
	-0,125	0,4	9,29	8,49	7,22	6,84	9,88	8,36	7,43	6,66	10,42	9,08	8,06	7,42	10,87	9,61	8,61	8,45
		0,6	9,55	8,46	7,30	7,15	10,02	8,53	7,83	6,77	10,45	9,67	8,42	7,68	11,25	9,76	8,80	8,30
		0,8	9,83	8,46	7,46	7,29	10,09	8,67	8,12	7,04	10,99	9,89	8,66	7,50	11,46	9,87	8,74	8,34
		1	10,02	8,60	7,73	7,58	10,36	8,89	8,06	7,37	11,44	9,75	8,51	7,88	11,40	10,09	9,08	8,52
	0	0,4	8,61	7,45	6,44	6,34	8,62	7,56	6,77	6,13	9,59	8,19	7,41	6,58	9,68	8,37	7,55	6,97
		0,6	8,78	7,74	6,83	6,57	9,20	7,91	6,96	6,29	9,86	8,56	7,58	6,72	10,28	8,87	7,63	7,27
		0,8	9,08	7,85	6,94	6,96	9,53	8,12	7,12	6,51	10,10	8,66	7,96	7,06	10,45	9,05	7,78	7,18
		1	9,49	8,21	6,93	7,23	9,70	8,30	7,27	6,48	10,45	8,72	7,87	7,06	10,86	9,08	8,17	7,11
	0,125	0,4	7,48	6,38	6,21	6,28	8,08	6,80	6,25	6,06	8,85	7,67	7,53	6,46	8,97	8,02	7,09	6,11
		0,6	7,96	6,72	6,17	6,62	8,17	7,02	6,38	6,32	8,90	7,97	8,10	6,12	9,51	7,92	7,07	6,46
		0,8	8,31	7,18	6,54	6,51	8,79	7,25	6,47	6,43	9,69	8,17	8,09	6,74	9,66	8,19	7,27	6,57
		1	8,33	7,35	6,85	6,85	8,77	7,36	6,58	6,50	9,50	8,09	8,86	6,57	9,93	8,49	7,60	6,74
0,016	-0,25	0,4	8,15	6,28	5,36	6,24	8,92	6,99	5,45	6,02	9,06	7,27	6,02	5,62	9,76	7,93	7,26	5,69
		0,6	8,48	6,88	5,28	6,24	9,36	7,18	5,66	5,93	10,05	7,59	6,20	5,62	10,43	7,93	6,61	5,16
		0,8	8,31	6,96	5,24	5,37	9,71	7,30	5,89	5,26	9,81	7,94	6,29	6,20	10,04	8,45	6,68	5,18
		1	8,95	7,04	5,38	5,22	10,10	7,14	5,74	6,37	10,71	7,82	6,16	6,56	10,59	8,28	7,09	6,01
	-0,125	0,4	7,59	5,80	4,90	6,53	7,93	6,10	5,22	6,23	8,59	6,57	5,91	6,00	9,48	7,31	6,43	7,20
		0,6	8,05	6,00	5,06	6,54	8,20	6,52	5,46	6,40	9,09	7,11	6,17	6,07	9,65	7,93	6,29	6,83
		0,8	8,15	6,37	5,48	6,39	8,44	7,20	5,89	6,31	9,35	7,54	6,92	5,38	10,59	7,75	6,54	7,03
		1	8,41	6,75	5,35	6,81	8,87	7,04	5,72	6,47	9,78	7,47	6,44	5,67	10,63	7,76	6,32	6,53
	0	0,4	7,12	5,39	5,27	6,23	7,50	5,72	5,14	5,95	8,13	6,40	6,17	5,78	8,45	6,61	5,53	5,54
		0,6	7,31	5,39	5,38	5,95	7,80	5,89	4,65	5,86	8,41	6,54	6,61	5,86	9,28	7,23	6,06	6,03
		0,8	7,72	5,85	5,66	5,56	8,19	6,12	5,24	5,05	8,95	6,94	6,95	5,16	9,33	7,35	6,69	5,98
		1	7,85	5,86	5,44	5,88	8,09	6,44	5,22	5,61	9,04	7,24	7,39	5,63	9,32	7,44	6,29	6,34
	0,125	0,4	6,39	4,72	4,82	5,73	7,14	5,61	4,36	5,13	7,83	6,09	6,38	4,92	8,34	6,43	5,14	4,53
		0,6	6,98	4,83	4,39	5,01	7,24	5,77	4,52	4,94	8,04	6,34	5,46	4,97	8,47	6,70	5,62	5,96
		0,8	7,07	5,10	4,71	4,81	7,55	5,86	4,42	4,33	8,12	6,23	5,60	5,76	9,48	7,36	6,12	6,28
		1	7,49	5,21	5,18	5,04	7,76	6,01	4,56	5,00	8,35	6,66	5,31	5,80	8,86	6,87	5,60	6,31

Figure 11: Pivot table of pressure at hydrostat 3

The assignment of the pressure values to a certain piping system configuration is done by the row and column sorting of the table. The pressure values are sorted line by line first by the two hose diameters, followed by sorting by the T-branch position and finally by hose length 1. The columns are sorted first by hose length 2 and then by hose length 3.

The pivot table shows, that the configuration with hose lengths 1 = 1000 mm, 2 = 1000 mm, 3 = 400 mm, diameter 12 mm and T-branch position -250 mm has the highest dynamic pressure pulsation amplitude at almost 13 bar and is therefore rather unsuitable, whereas, for example, the configuration with hose lengths 1 = 600 mm, 2 = 400 mm, 3 = 800 mm, diameter 16 mm and T-branch position +125 mm has a low dynamic pressure pulsation amplitude at approx. 4.4 bar and appears suitable.

However, this finding must still be cross-checked with the pressure pulsation amplitudes at the other two measuring points to finally evaluate whether the configuration provides the best overall result.

5. DETAILED ANALYSIS OF THE RESULTS

Of course, the developer can also take a closer look at the simulation results at any time. For this purpose, additional Python script tasks are added to the workflow as needed. **Figure 12** shows the results of the task *080_spectrogram_analysis*.

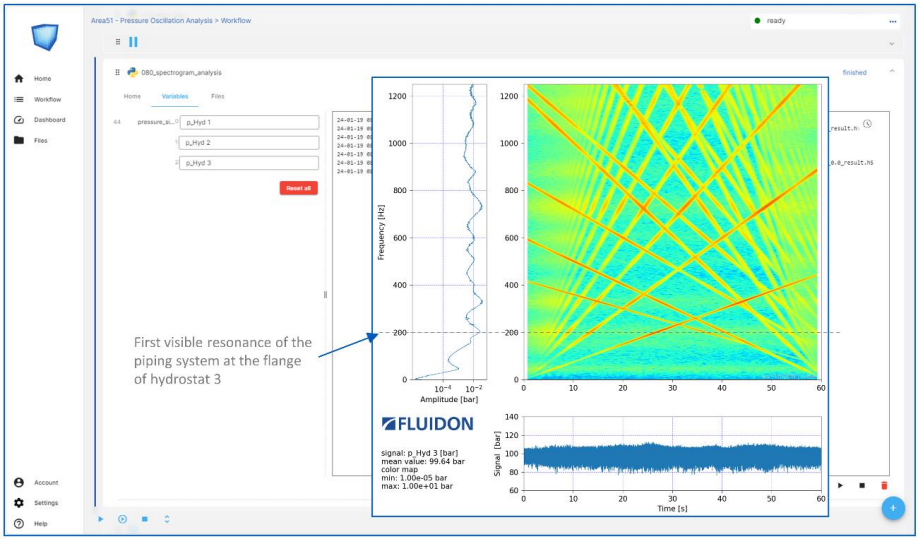


Figure 12: Python script for spectrogram analysis

Data source for the diagram is again the pressure signal at the flange of hydrostat 3. At the yellow lines of the individual excitation orders of the hydrostats it can be clearly seen that all three hydrostats change their speed. Whenever one of the excitation frequencies "hits" a resonance frequency of the piping system, the colour changes from yellowish to reddish. The first piping system resonance is at approximately at 190 Hz. And there are more piping system resonances with even higher pulsation amplitudes that become visible at higher orders.

For now, it is of interest which pressure oscillation mode of the piping system is associated with this frequency. The pressure oscillation mode is identified by task *090_pressure_vector_plot* (**Figure 13**).

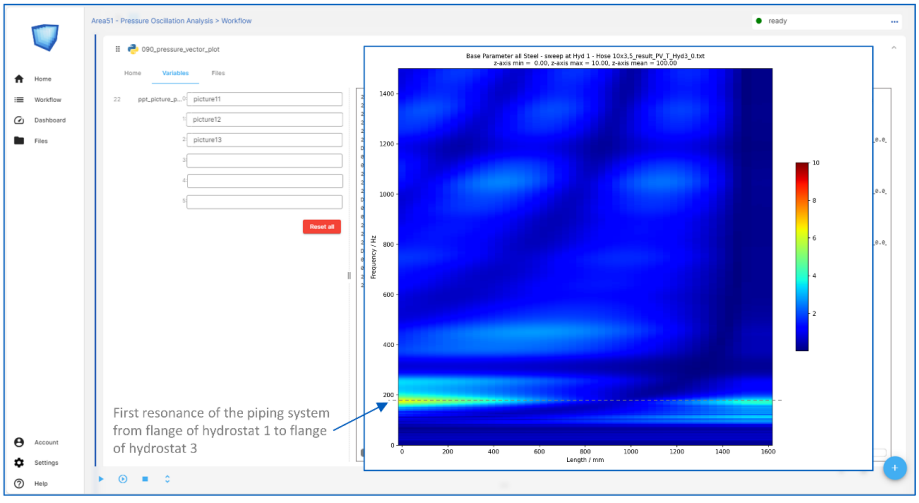


Figure 13: Python script to create pressure vector diagrams

The resulting diagram visualizes the pressure oscillation situation along the piping's middle axis vs. excitation frequency. The visible oscillation mode at 190 Hz, indicated in Figure 13, is the fundamental λ/2 resonance frequency of a piping system with both ends closed. However, an

investigation or detailed explanation would be another story. For a detailed description and interpretation of spectrograms and pressure vector plots, please refer to [2].

The compact exploration of the detailed time series results, along with the corresponding diagrams, demonstrates that the example workflow not only identifies whether the piping system configuration is favourable or unfavourable but also offers valuable insights into why a particular configuration is considered as such. This provides users with information regarding which parameters have the most significant impact on the system's dynamic behaviour.

6. SUMMARY

Using the example of a hydrostatic drive, it was shown how an easy-to-use, yet powerful simulation workflow can be set up within the Cube SaaS development platform. However, Cube is not limited to fluid power issues, but was designed as a platform for solving simulation or design issues from all technical areas. With the FMU plug-in, digital twins from various simulation tools can be integrated into the platform. Thanks to the flexible composition of the workflows from predefined plug-ins and Python scripts, Cube can be used to solve a wide range of problems.

REFERENCES

[1] VEL - a modular virtual commissioning environment for mobile machinery, 11th Colloquium Mobile Hydraulics, September 10, 2020, Karlsruhe, Germany, pp. 133 - 146

[2] Baum H.: Pressure Oscillation Analysis of Hydrostatic Drive Train, O+P, 06-2019, page 36 - 41

[3] DSHplus - Simulation of hydraulic and pneumatic systems in a mechatronic environment, https://www.fluidon.com/, last visited 01.2024

[4] FMI-standard - Functional Mock-up Interface for Model Exchange and Co-Simulation, https://fmi-standard.org, last visited 01.2024

[5] Statistical Design of Experiments, https://de.wikipedia.org/wiki/Statistical_design_of_experiments, last visited 01.2024

CREDIBLE SIMULATION: EVALUATING THE CREDIBILITY OF SIMULATION MODELS AND SIMULATION MODEL LIBRARIES

peer reviewed

Simon Leutz[1]*, Ehsan Khademi[1]

[1]*Bosch Rexroth AG, Lise-Meitner-Straße 2, 89081 Ulm, Germany*

* Corresponding author: Tel.: +49 731 49373-192; E-mail address: Simon.Leutz@boschrexroth.de

ABSTRACT

This paper deals with the evaluation of the credibility of simulation models and simulation model libraries exemplary shown on a model of the Rexroth BRSL Simulation Library.

The topic tackles the questions on how engineers know that they really can trust the results of their simulation models, how exact their simulation models are and if the accuracy of their simulation models is sufficient or not.

In industry, engineers usually validate their simulation model results against real measurement data to examine if the simulation model results match the reality. Usually there is no numerical number describing the accuracy of the simulation model. In the topic of credible simulation, numerical methods of data analysis are used to try to determine a numerical number, the so-called model form error. This then gives an idea of how plausible the simulation model is in terms of uncertainties.

At Bosch Rexroth there is a prototypically developed app that enables users estimating the model form error, performing sensitivity analyses, and simulating variations of parameter uncertainties on key model parameters. Furthermore, there is the possibility to compare the simulation results with real measurement data.

The app may help to certify credible simulation models in the future.

Keywords: Simulation credibility, model form error, sensitivity analysis, p-box, Wasserstein metric

1. INTRODUCTION

Credibility of simulations is becoming increasingly relevant. Especially in times of ever-increasing digitalization of products and workflows in the industry up to purely virtual-based product releases, simulation-based decisions become more and more important. For this purpose, it is crucial that simulation results are given the necessary confidence and that their agreement with reality can be scientifically proven. [1]

Product validation via measurements on the real product is common and state of the art. Measurement data also contain certain inaccuracies and are not always easy to obtain, which is why they are already increasingly being supplemented by computer simulations. In this way, further confidence in the product can be gained and its behavior better understood.

To save time and thus money, and provided that the simulation results can be trusted, real measurements on the product can be completely replaced by purely virtual computer simulations.

The aim of this work at the Bosch Rexroth AG is to develop methods that allow full confidence in simulation results to be obtained by simulation and that this confidence level can be suitably represented by means of a numerical key figure, the so-called model form error.

In a further step, the determination of this credibility level should also be made possible for other models of other areas within the Bosch Rexroth AG. The credibility obtained can be integrated into digital product passports as metadata of the simulation models to certify the quality of the simulation model regarding its trustworthiness. This gained credibility stamp may help answering the following questions. Can you really trust your simulation? Do the simulation results match the reality? How exact is the simulation? Is it sufficiently exact?

2. STATE OF THE ART

2.1. Credibility standards in industry

Currently there are some standards in industry developed by NASA, NAFEMS, ASME and few more [1]. Also, there is a lot of research done within the Robert Bosch GmbH trying to merge these different standards regarding their quality metrics describing simulation and model (S & M) credibility. [1] The paper in hand focusses on the calculation of a so-called model form error as a numerical key indicator describing S & M credibility. There are different methods which will be explained hereafter.

2.2. Probability box (p-box) method

The probability box (p-box) method is an interval-based method to represent mixed uncertainties of a simulation model based on epistemic and aleatory parameter uncertainties, see **Figure 1** [2].

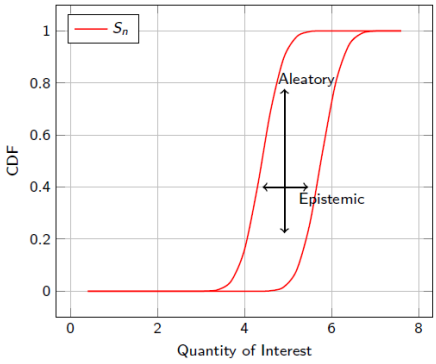


Figure 1: Representation of mixed uncertainties as p-box [2]

Epistemic parameter uncertainties are characterized to have a constant but unknown parameter value due to missing system knowledge or data. They are represented as an interval [2] [3].

Against this, aleatory parameter uncertainties do not have a constant value, e.g., the temperature in a system. Aleatory parameter uncertainties are usually represented as cumulative distribution functions (CDFs) [2] [3].

A p-box is created from the CDFs of the different simulation results of parameter combinations of epistemic and aleatory parameters including numerical uncertainties. The most left and the rightest CDF are the boundaries of the p-box representing the lower and the upper limit of trustworthiness [2], compare **Figure 2**.

In **Figure 2** the CDFs of the interesting simulation result variable, the so-called Quantity of Interest (QoI), for the different simulated parameter combinations of epistemic and aleatory parameters, named here with m_1, …, m_5, are shown for the evaluated point in time while m_1 and m_5 form the lower (m_1) and the upper (m_5) boundaries of the p-box, defining the interval of trustworthiness for the simulated setup and the specific point in time.

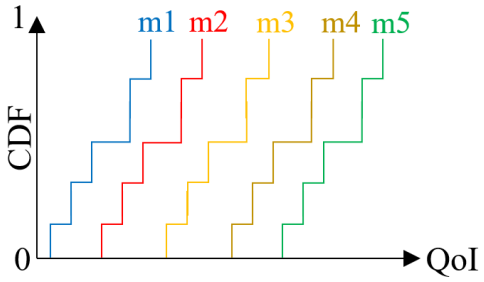


Figure 2: Exemplary CDFs for p-box generation

With the help of a p-box of a simulation model, a so-called model form error can be calculated in comparison with real measurement data, which quantifies the quality of the simulation model and its consistency with reality depending on how far the real measurement data are within or outside the p-box limits for a single point in time [2].

Therefore, the CDF of the measurement data is compared with the generated p-box of the simulation model and the model form error is calculated as area outside the p-box between the CDF of the measurement data and the lower and the upper p-box limit, as exemplary shown in **Figure 3**.

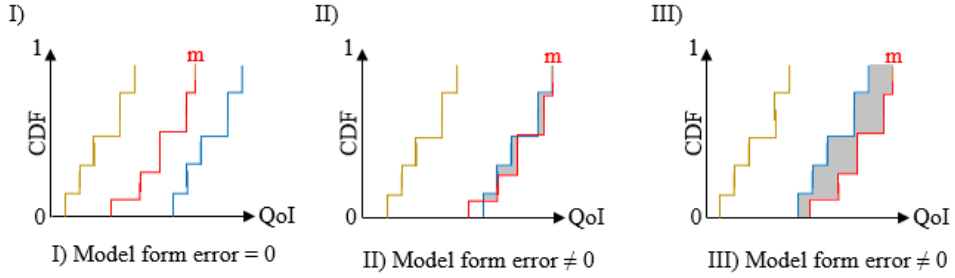


Figure 3: Calculation and graphical representation of the model form error

2.3. Wasserstein metric

The Wasserstein metric, also known as Wasserstein distance, Kantorovich-Rubinstein metric, optimal transport, or earth mover's distance, is used to calculate the numerical cost of an optimal transport problem [3]. It is used for the (statistical) time series analysis of dynamic systems and considers temporal progressions as well as spatial distances [3][4].

Since the Wasserstein metric considers not only a purely vertical point-to-point comparison between two time series, but also the horizontal distances between the curves to be compared, it is a very powerful method for time series analysis, see **Figure 4** [4].

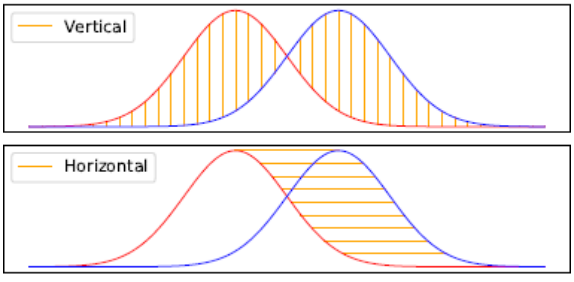


Figure 4: Comparison of two time series using the Wasserstein metric

A time series $y(t,\zeta) = x,\ x\epsilon\mathbb{R}^m$ is modeled as a stochastic process that follows a probability measure. Calculating the empirical distribution function for the x-values that occur, one obtains a point set p with a weighting ω_p [5].

Between two time series $p,\ \omega_p\epsilon\mathbb{R}^m$ and $q,\ \omega_q\epsilon\mathbb{R}^n$ with probability measures γ, ν the following optimization problem, shown in (1), applies [6]:

$$W_\alpha(\gamma,\nu) = \left(\min_{(f_{ij})} \sum_{i=1}^{m}\sum_{j=1}^{n} d_{ij}f_{ij}\right)^{\frac{1}{\alpha}} \tag{1}$$

with the following properties:

- $d_{ij} = \|p_i - q_j\|^\alpha$
- $\sum_{j=1}^{n} f_{ij} \leq \omega_{p_i},\ \forall i \in \{1,..,m\}$
- $\sum_{i=1}^{m} f_{ij} \leq \omega_{q_j},\ \forall j \in \{1,..,n\}$
- $f_{ij} \geq 0,\ \forall i \in \{1,..,m\},\ \forall j \in \{1,..,n\}$
- $\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} = min\{\sum_{i=1}^{m}\omega_{p_i}, \sum_{j=1}^{n}\omega_{q_j}\}$

This problem is then solved as a linear optimization problem.

In the Wasserstein metric, as with the p-box method, the most influential aleatory and epistemic parameters are varied, and the time series of the output variables are calculated. A Wasserstein distance is then calculated between the time series of the simulation and that of the measurement and is expressed as the model form error.

While the p-box method can be used for calculating the model form error for single points in time, the Wasserstein metric can be used to calculate the model form error for whole time series.

2.4. Sensitivity analysis

There are already numerous, different sensitivity analysis methods for different use cases as well as (commercial) sensitivity analysis tools available on the market. However, since we do not need the complete, comprehensive range of available sensitivity analysis methods (of a tool), but rather a tailored, custom-fit solution for our own investigations, we have decided to implement the required functionalities for the Rexroth BRSL Credible Simulation app ourselves.

Just as in chapter 2.3, it should be referred here mainly to the bachelor thesis of Ehsan Khademi [3] in which he has examined different sensitivity analysis methods to determine the main influencing epistemic and aleatory parameter combinations for the model form error calculation of simulation models for single points in time as well as for whole time series.

The sensitivity index of the input parameters ζ_i on the output variable $y(t_k,\zeta)$ for a point in time $t_k\epsilon[0,T]$ is defined as $\frac{\partial}{\partial\zeta_i}y(t_k,\zeta)$. It is approximated by a finite difference method called elementary effects (EE) as follows (2):

$$EE_i = \frac{y(t_k,\zeta + \Delta_i\cdot e_i) - y(t_k,\zeta)}{\Delta_i} \tag{2}$$

Using Morris' method, the sensitivity indices are calculated at a point in time for various combinations of input values [7][8].

The result is a ranking of the parameters according to their influence on the output variable as shown in **Figure 5** [3] on the example of the well-known predator-prey model.

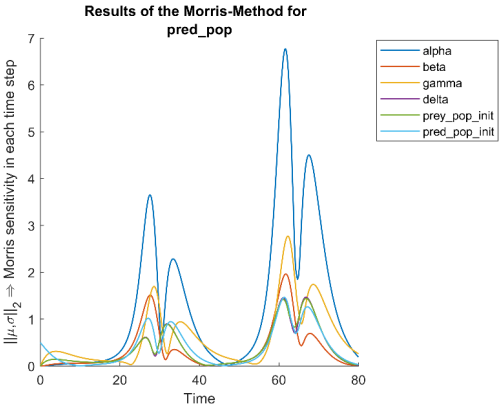


Figure 5: Sensitivity indices of the input parameters to the output variable at each time step

To check the sensitivity indices for the entire time series, the L_2 norm is calculated for the sensitivity of each individual parameter ζ_i over the entire period for all k time steps after each variation of the input values, see (3).

$$\int_0^T \frac{\partial}{\partial \zeta_i} y(t,\zeta)^2 \, dt \approx \sqrt{\sum_{k=0}^{T} \left(EE_i^{(k)}\right)^2} \tag{3}$$

These values are then displayed as box plots, which provide an overview of the influence of the parameters on the time series of the output variable, exemplary see **Figure 6**.

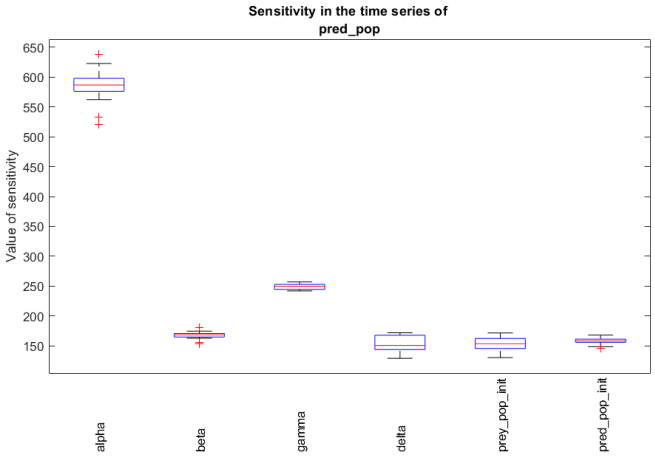


Figure 6: Sensitivity indices of the input parameters to the entire time series of the output variable

In [3], we have also investigated the use of the Wasserstein metric for a more detailed estimation of the sensitivity indices.

For calculating the model form error uncertain epistemic and aleatory parameters must be chosen by the user. To know which epistemic and aleatory parameters make sense to vary and therefore have an influence on the model result of interest, the QoI, and which do not, performing a sensitivity study

on the simulation model parameters using suitable methods before can be useful. This makes the calculation of the model form error of the simulation model for relevant QoI and therefore the credibility evaluation of the simulation model more efficient.

3. REXROTH BRSL SIMULATION LIBRARY

The Rexroth BRSL Simulation Library, shown in **Figure 7**, is a simulation model library containing simulation models, so-called digital twins, which are designed for the usage of dynamic 1D system simulations. It contains many generic components for modelling physical systems as well as specific components of Bosch Rexroth products.

The models are implemented in Modelica. Modelica is considered as a tool-independent programming language [8]. This means that the Rexroth BRSL Simulation Library models can be used in various simulation tools. The use as functional mock-up unit (FMU), a standardized exchange format, is also conceivable.

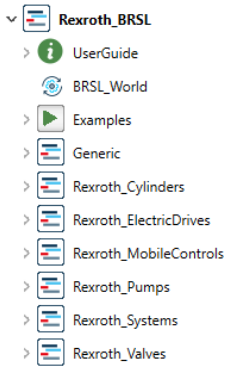


Figure 7: Rexroth BRSL Simulation Library loaded in Dymola 2023x Refresh 1

All simulation models from the Rexroth BRSL Simulation Library have been tested and validated together with experts against real measurement data as far as available or at least compared with more detailed and more complex simulation models if no measurement data have been available.

The simulation models of the Rexroth BRSL Simulation Library are pre-parameterized and designed for the generic use case of dynamic system simulations. Users can focus on system modelling using the simulation models out of the rack without the need to model single system components before on their own.

Furthermore, the generic components can be completely parameterized by the users according to their needs while the Bosch Rexroth specific components are already pre-parameterized offering the option to choose between specific products variants via their specific product type keys in a dropdown menu, see **Figure 8**.

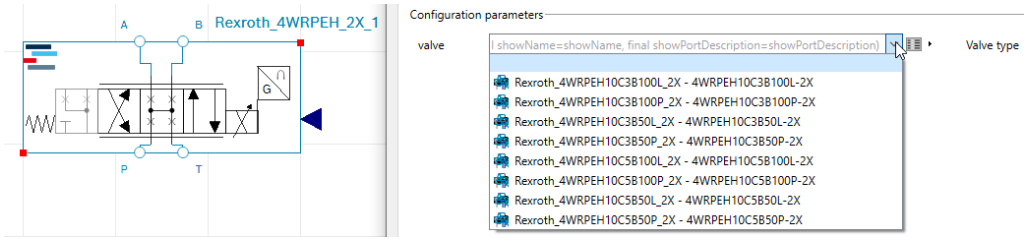


Figure 8: Pre-parameterized Bosch Rexroth_4WRPEH_2X_1 type key choices

This makes the Rexroth BRSL Simulation Library easy to use, reduces parameterization errors as well as the time needed for modelling.

The Rexroth BRSL Simulation Library can help users validating the dynamic system behavior of their systems using computer simulation models. Furthermore, it can help answering future relevant questions regarding system energy consumptions and CO_2 footprints.

Nevertheless, due to simplified modelling assumptions, the simulation models can be used differently well for different use cases. How well a simulation model fit the specific use case can be estimated using credible simulation methods.

4. REXROTH BRSL CREDIBLE SIMULATION

The Rexroth BRSL Credible Simulation app written in MATLAB can be used to calculate the model form error of an uploaded simulation model using either the p-box method or the Wasserstein metric. It is currently limited to Modelica simulation models (not only Rexroth BRSL Simulation Library models but Modelica models in general) but is to be extended for other simulation models and FMU as well.

The general workflow and the communication between the Rexroth BRSL Credible Simulation app and the simulation tool chosen by the user, pictured in **Figure 9**, is exemplary shown with Dymola as simulation tool and optional usage of the Rexroth BRSL Simulation Library.

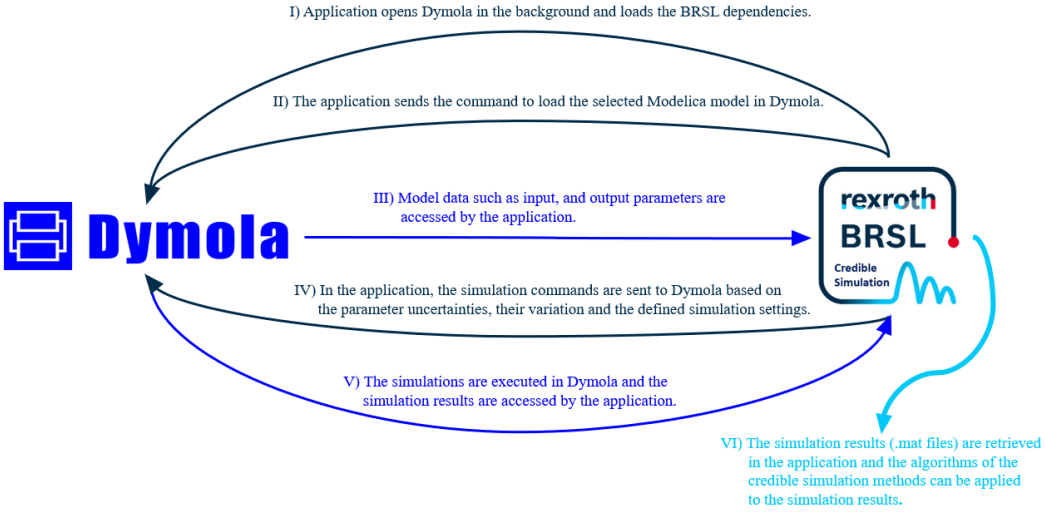


Figure 9: Communication of the Rexroth BRSL Credible Simulation app with commercial simulation tools using the example of Dymola

The users can define epistemic and aleatory parameter uncertainties on specific model parameters which can be chosen via dropdown menus (due to step III) in **Figure 9**). If only one epistemic or one aleatory parameter should be varied, for the other parameter can be set a fixed value using the manual input checkbox shown in **Figure 10**.

Moreover, the QoI can be chosen via dropdown menu as well, see **Figure 10**, and the simulation settings (‘*Start time*’, ‘*Stop time*’, ‘*Solver method*’, ‘*Output interval*’, ‘*Tolerance*’) for simulating the uploaded simulation model with varying parameter combinations of epistemic and aleatory parameters (simulating the parameter uncertainties) can be set in the app (compare step IV) in **Figure 9**).

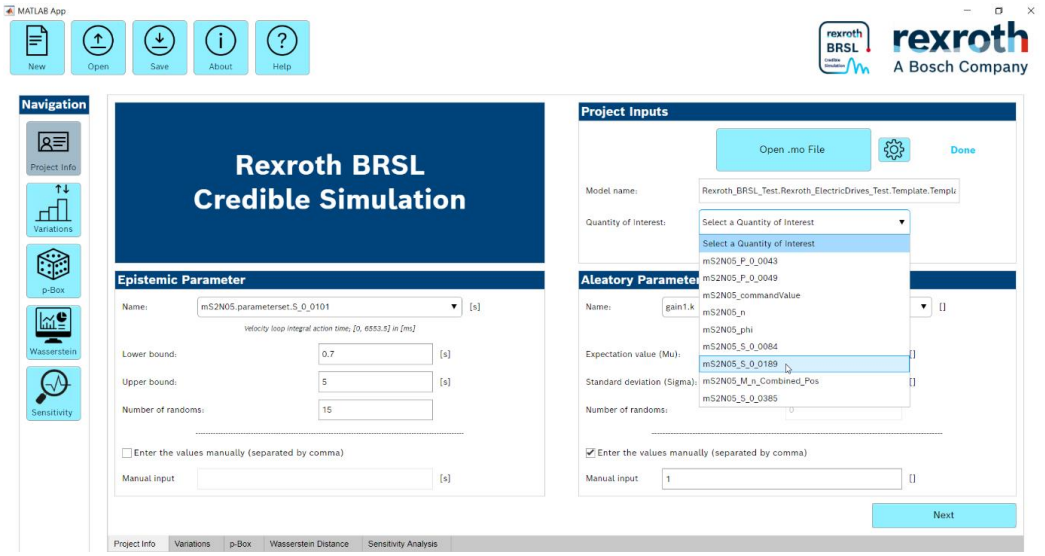


Figure 10: '*Project Info*' tab for model and parameter settings

A tool for simulating the simulation model with each parameter combination needs to be installed locally on the user's personal computer. Which tool should be used (currently limited to Dymola but is to be extended for other simulation tools as well) can be specified in the '*Project info*' tab using the settings button right next to the '*Open .mo File*' button pictured in **Figure 10**.

Starting the simulations of all variations, a pop-up window opens showing the progress of simulating all parameter combinations and plotting the result of each parameter combination in the '*Variations*' tab. The result plot window can be maximized and curves of single (of all) parameter combinations (plotted in different colours) can be de-/selected in the results plot for better postprocessing.

The QoI can be changed without recompilation of the simulation model choosing another QoI from the '*Changing QoI*' dropdown menu and clicking on the '*Change*' button. Thereby, the question mark button right next to the '*Change*' button offers further information of the chosen new QoI.

To find out which model parameters make most sense to be varied and which model parameters do not need to be varied, if they are not already known yet, there is the possibility to run a sensitivity analysis within the Rexroth BRSL Credible Simulation app ('*Sensitivity*' tab) to gain optimal simulation results.

Having defined the best suitable setup for varying epistemic and aleatory model parameters to simulate parameter uncertainties of these parameters and watching the effect on the simulation results of the specified QoI, the model form error of the uploaded user's simulation model regarding the defined setup can be calculated using either the p-box method ('*p-Box*' tab) or the Wasserstein metric ('*Wasserstein*' tab), see step VI) in **Figure 9**.

While the Wasserstein metric exclusively provides a numerical calculation of the model form error of the whole time series, the p-box method provides a numerical calculation of the model form error as well as a graphical interpretation of specific points in time.

The points in time that are of interest for generating a p-box plot can be chosen by the user from the '*Result variations*' plot of the specified QoI in the '*Variations*' tab. The data can then be entered into the '*p-Box Options*' within the '*p-Box*' tab, see **Figure 11**.



Figure 11: '*p-Box*' plot without measurement data

Measurement data can be integrated into the p-box plot for the calculation of the model form error via the '*Table of Measurements*' tab inside the '*p-Box*' tab. This tab opens automatically unchecking the '*Plot p-Box without measurements*' checkbox in the '*p-Box Options*'.

Within the '*Table of Measurements*' tab there is a slider choosing between manual input of measurement data or upload measurement data as .csv data into the Rexroth BRSL Credible Simulation app. The unused option is greyed out as shown in **Figure 12**.

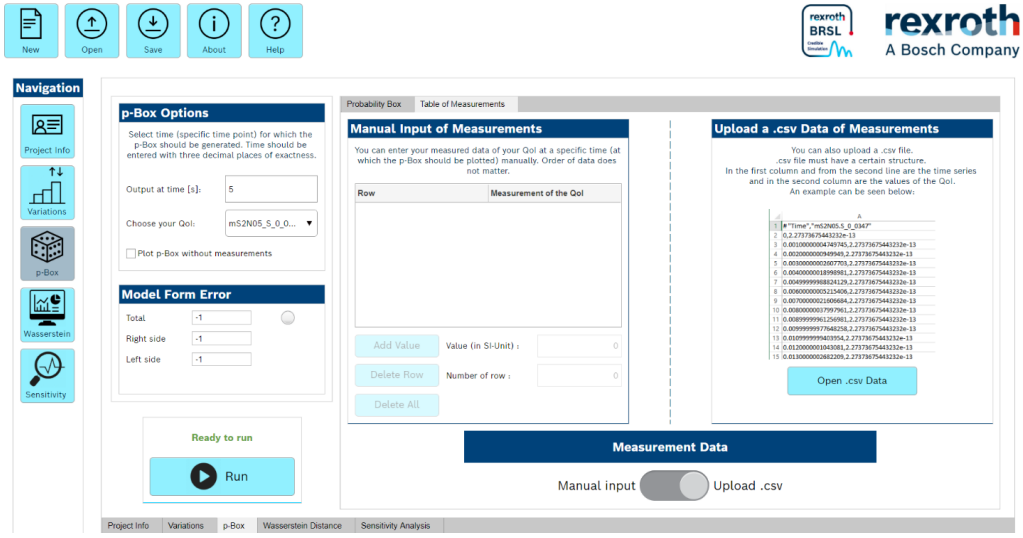


Figure 12: '*Table of Measurements*' tab for the p-box

The measurement setup must match the simulation model under investigation, which is why real measurement data are often not available. Therefore, the model form error of the simulation model can not be calculated using the p-box method. Nevertheless, sensitivity studies on uncertain model parameters can be performed for a better system understanding.

As far as available, having inserted measurement data for comparison, the p-box can be plot again with the measurement data included to calculate the model form error of the uploaded simulation model for the chosen point in time.

If the uploaded simulation model totally matches the uploaded measurement data, the CDF of the measurement data will be totally inside the p-box plot and the model form error will be zero (compare **Figure 3** I)).

Otherwise, if the CDF of the measurement data is outside the p-box plot (on the left and/or on the right side), the model form error on the left side, on the right side and in total (left side + right side) will be calculated as area between the CDF of the measurement data and the left and right boarders of the p-box plot of the simulation model for the specified point in time.

The circle to the right of the total model form error colours in traffic light colours depending on the calculated model form error. If the calculated model form error is zero, which means that the CDF of the measurement data is completely within the p-box limits, the circle is coloured in green. If parts of the CDF of the measurement data are outside the p-box limits the circle is coloured in orange (compare **Figure 3** II)) and if the CDF of the measurement data is completely outside the p-box limits (compare **Figure 3** III)), the circle is coloured in red.

A green circle (model form error = 0) means, that the user can trust the simulation model under the simulated conditions and for the samples used. A red circle means, that the user cannot trust the simulation model under the simulated conditions and an orange circle means, that the user can partially trust the simulation results and carefully needs to check the conditions for which the simulation model is trustworthy, and for which it is not.

The calculated model form error is use case specific. That means that the user always must check the simulation results and the calculated model form error for the specific use case the simulation model is designed for. The Rexroth BRSL Credible Simulation app therefore provides an indicator if the simulation model is (partially) credible or not.

Moreover, error handling routines are implemented for a better user experience dealing with invalid user inputs as well as errors in the used simulation tool or within the Rexroth BRSL Credible Simulation app and providing suggestions on how to solve the occurred issue.

5. USE CASE EXAMPLE

The Rexroth BRSL Credible Simulation app can be used to evaluate the simulation model credibility of simulation models such as those of the Rexroth BRSL Simulation Library, shown on a small exemplary test model of a MS2N05 electric drive of the Bosch Rexroth AG, shown in **Figure 13**.

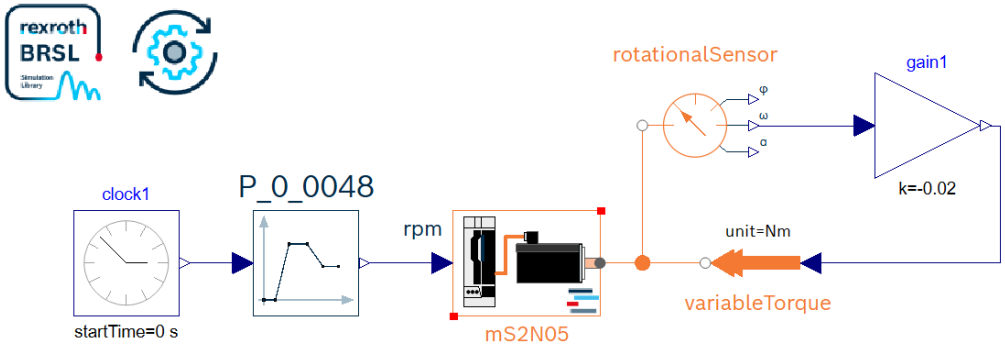


Figure 13: Simulation model of an electric drive of the Bosch Rexroth AG

In this small example a Bosch Rexroth MS2N05 electric drive is simulated in position control. It is controlled by the tabular setpoint specification of parameter '*P_0_0048*' over time and runs load-free against an internal system friction modelled at the motor flange to test the idling behaviour of the electric drive.

Having talked with experts and having varied couple of different parameter combinations of epistemic and aleatory model parameters that seemed to be uncertain, it came out that for the example pictured above in **Figure 13**, the internal friction 'gain1.k' as uncertain aleatory and the motor current at standstill 'mS2N05.S_0_0111' as uncertain epistemic parameter are most suitable.

For the motor current at standstill, 'mS2N05.S_0_0111', five random epistemic parameter samples were generated within the interval of $0 - 10$ Ampere, while for the internal friction 'gain1.k' twenty random aleatory parameter samples were generated with an expectation value $\mu = -0.5$ and a standard deviation of $\sigma = 0.2$ what leads to 100 simulations since each epistemic parameter value is combined with each aleatory parameter value.

The QoI are the motor speed '*mS2N05.n*', shown in **Figure 14**, the torque/force feedback value '*mS2N05.S_0_0084*', shown in **Figure 15**, and the velocity error '*mS2N05.S_0_0347*', shown in **Figure 16**, which are interesting to look at and for which a p-box plot for the interesting point in time should be created.

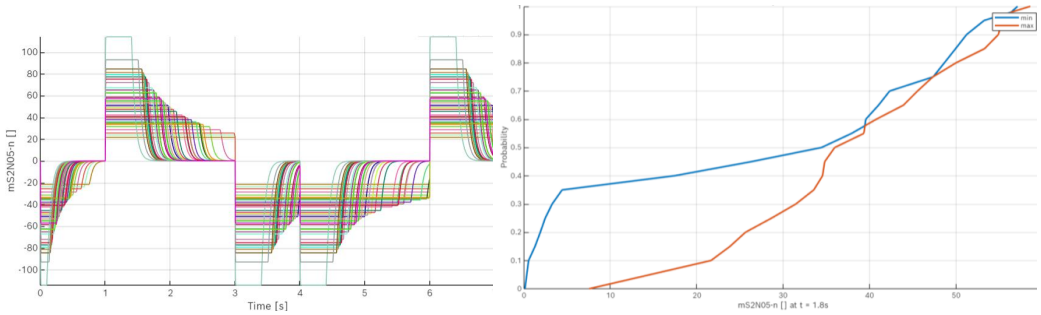


Figure 14: Variation and p-box plots of the motor speed of the MS2N05 simulation model

Figure 14 shows varying simulation results for the motor speed for the 100 simulation runs of all different parameter combinations of aleatory and epistemic parameters. The epistemic and aleatory influences on the QoI are also visible in the p-box plot. The more the CDF of the left and the right boarder of the p-box plot vary in x direction (along the time axis) the greater is the epistemic parameter influence on the QoI. The more the CDF of the left and the right boarder of the p-box plot vary in y direction (along the axis of the QoI) the greater is the aleatory parameter influence on the QoI (compare **Figure 1**).

Moreover, the permissible validity range of the real measurement data (min/max values) can be read directly from the x-axis values of the p-box limits. For example, the p-box plot of the motor speed (**Figure 14** on the right) shows that about 58% of the measurement values need to be less than 39 rpm, otherwise the measurement data will indicate a model form error. In addition, if most of the measured values are greater than 35 rpm, a model form error (outside the p-box limits) can be assumed, as most measured values are expected to be smaller than 35 rpm according to the p-box plot of the motor speed.

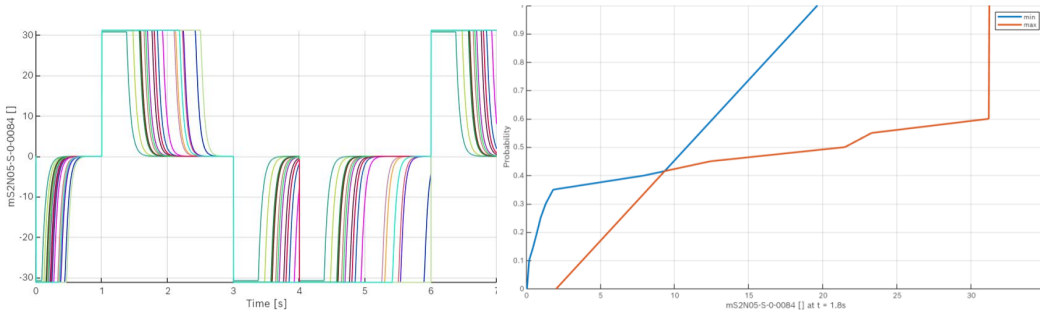


Figure 15: Variation and p-box plots of the torque/force feedback value of the MS2N05 simulation model

Figure 15 shows varying simulation results for the torque/force feedback value for the 100 simulation runs of all different parameter combinations of aleatory and epistemic parameters.

The p-box plot at t = 1.8 s on the right shows that measurement values greater than 31.3 N will directly lead to a model form error. Also, all values need to be greater than zero and between 19.8 N and 31.3 N. Furthermore, 41 % of the measurement values are expected to be smaller than 9 N.

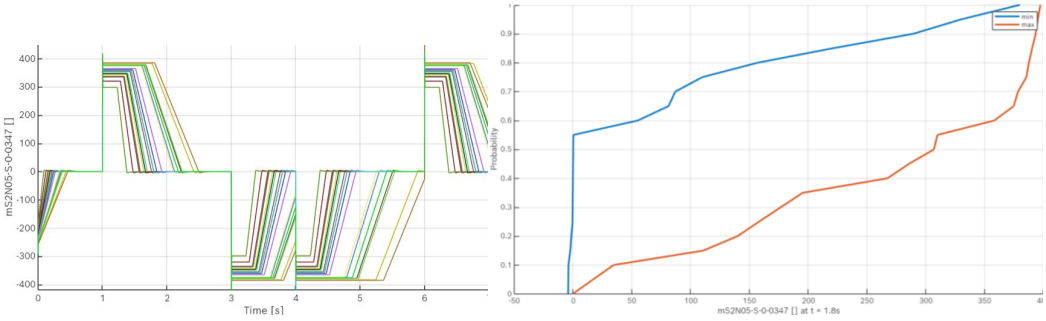


Figure 16: Variation and p-box plots of the velocity error of the MS2N05 simulation model

Figure 16 shows varying simulation results for the velocity error for the 100 simulation runs of all different parameter combinations of aleatory and epistemic parameters. The p-box shows big epistemic and aleatory parameter influences. Compared to the other two QoI, the p-box plot for the velocity error shows the biggest scope of validity.

The plotted p-boxes of the three QoI show the ranges in which the user can trust the simulation results obtained according to the aleatory and epistemic parameter combinations (values stored in '*Variations*' plot) which is also kind of a sensitivity study for the chosen parameter combination.

Unfortunately, there are no real measurement data available for this example. If there would be measurement data available, they could be inserted in the p-box plots, and if the CDF of the measurement data would be totally within the p-box boarders, the user could totally trust the simulation model for the samples used. Since there are no real measurement data available the model form error cannot be calculated via the Rexroth BRSL Credible Simulation app.

Nevertheless, there is an even more detailed simulation model for this example available which was used during development as a reference model for the validation of the MS2N05 simulation model of the Rexroth BRSL Simulation Library, so the user can trust the simulation results of the simulation model which matches reality. Only the model form error as numerical key figure describing how exact this simulation model matches the reality is missing.

6. CONCLUSION AND OUTLOOK

The paper in hand has shown how to evaluate the simulation credibility of simulation models and simulation model libraries exemplary shown on the Rexroth BRSL Simulation Library and how it is done at the Bosch Rexroth AG. It has presented useful methods such as the interval-based p-box method including a graphical interpretation as well as the Wasserstein metric for calculating the model form error of simulation models. The paper in hand has presented the inhouse Rexroth BRSL Credible Simulation app and its usage on an example simulation model of the Rexroth BRSL Simulation Library.

Nevertheless, the model form error calculation for the estimation of the simulation model credibility is use case dependent and depends on the availability and quality of measurement data for the specific use case and the epistemic and aleatory parameter samples used. The more data the better.

In the future it is planned to consecutively estimate the simulation model credibility for all simulation models of the Rexroth BRSL Simulation Library using the in chapter 2.1 mentioned framework [1] of the Robert Bosch GmbH. The result will be a simulation model credibility estimation like the one exemplary shown for the '*Rexroth_ElectricDrives*' (see **Figure 7**) in **Figure 17**.

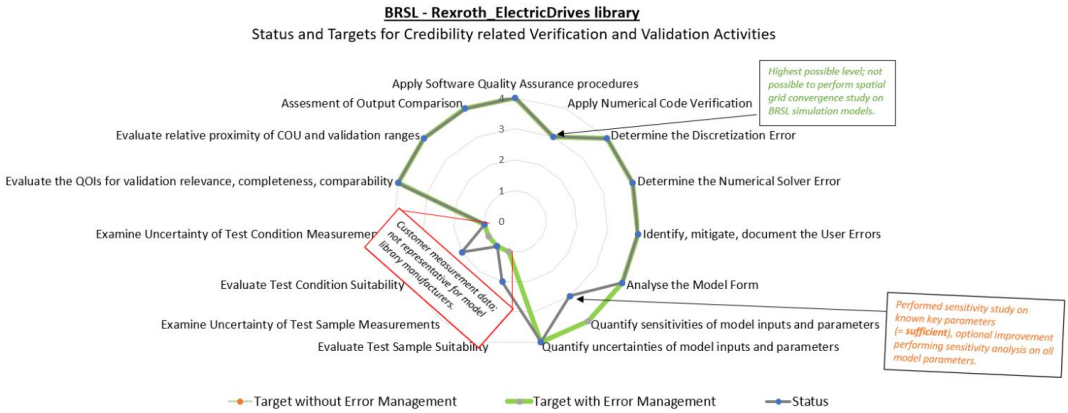


Figure 17: '*Rexroth_ElectricDrives*' credibility estimation

In a next step, the credibility attributes shown in **Figure 17** can be integrated as metadata to the simulation models and thus also into the Asset Administration Shell (AAS) [9].

In the future, business models regarding certification services of the credibility of simulation models based on simulation credibility levels are also conceivable.

The further app development such as supporting further simulation tools and simulation model types or FMU in the Rexroth BRSL Credible Simulation app is constantly progressing at the Bosch Rexroth AG. Furthermore, final test validation and release for productive usage of the Rexroth BRSL Credible Simulation app is pending.

NOMENCLATURE

Rexroth BRSL Simulation Library	Simulation model library of the Bosch Rexroth AG containing generic 1D simulation models as well as simulation models of Bosch Rexroth specific products for system simulation
Rexroth_ElectricDrives	Simulation model library of the Rexroth BRSL Simulation Library containing electric drive simulation models of the Bosch Rexroth AG

Rexroth BRSL Credible Simulation	App developed by the Bosch Rexroth AG to estimate the credibility of simulation models	
V & V	Verification & Validation	
NASA	National Aeronautics and Space Administration	
NAFEMS	International Association for the Engineering Modelling, Analysis and Simulation Community	
ASME	American Society of Mechanical Engineers	
S & M credibility	Simulation & Model credibility	
p-box	probability box	
e.g.	for example	
CDF	Cumulative Distribution Function	
QoI	Quantity of Interest / Quantity of Interest	
$m_1, \ldots, m_5$	Exemplary CDFs	
$y(t, \zeta)$	Time series as a function of time and model input parameters	
t	Time	s
ζ	Vector of model input parameters	
γ, ν	Probability measures of the time series	
x	Representation of output values of an output variable	
m, n	Dimension of the output variable	
p, q	The occurred values of the output variable	
ω_p, ω_q	Cumulative probability of p, q values	
W_α	Wasserstein distance of order alpha	
d	Cost matrix between two-point sets (Euclidean distance)	
f	Flow matrix between two-point sets	
T	Simulation time limit	s
EE	Elementary effects	
e	Unit vector	
alpha	Reproduction rate of the prey (Lotka-Volterra model)	
beta	Mortality rate of the predator per prey (Lotka-Volterra model)	
gamma	Mortality rate of the predator (Lotka-Volterra model)	
delta	Reproduction rate of the predator (Lotka-Volterra model)	
prey_pop_init	Initial prey population (Lotka-Volterra model)	
pred_pop_init	Original predator population (Lotka-Volterra model)	
pred_pop	Predator population (Lotka-Volterra model)	
L_2 *norm*	Euclidean norm	
k	Number of time steps	
FMU	Functional Mock-up Unit	
Dymola 2023x Refresh 1	Dynamic Modeling Laboratory, commercial Modelica simulation tool	
Rexroth_4WRPEH_2X_1	4WRPEH valve of the Bosch Rexroth AG	
CO_2	Carbon dioxide	
MS2N05	MS2N servo motor of the Bosch Rexroth AG	
P_0_0048	Velocity command value of controller	
gain1.k	Factor for friction modelling	
mS2N05.S_0_0111	Motor current at standstill	A
mS2N05.n	Motor speed	rpm
mS2N05.S_0_0084	Torque/force feedback value	
mS2N05.S_0_0347	Velocity error	rpm
μ	Expectation value	
σ	Standard deviation	
AAS	Asset Administration Shell	

REFERENCES

[1] Atak M, Koehler R, Goeppel T, Heinkel H-M, Karl A, von Keler J, Kerst A (2023) On the Credibility of Modeling and Simulation Results in Cross-Domain and Cross-Company Product Engineering. Paper presented at the NAFEMS World Congress 2023 (NWC23), Robert Bosch GmbH, Tampa, Florida, USA, 15-18 May 2023

[2] White L, West T (2019) Area Validation Metric for Applications with Mixed Uncertainty. Vehicle Analysis and Concepts Directorate, NASA Langley Research Center, DATAWorks 2019, 11 April 2019

[3] Khademi E (2023) Simulative Untersuchung von Parameterunsicherheiten mittels Wasserstein-Metrik. Bachelor thesis, Ulm University

[4] Cazelles E, Robert A, Tobar F (2019) The Wasserstein-Fourier Distance for Stationary Time Series. eprint arXiv:1912.05509v2 [stat.ML], last revised 11 Dec 2020. https://doi.org/10.48550/arXiv.1912.05509. Accessed 19 Dec 2023

[5] Hartmann V, Schuhmacher D (2017) Semi-discrete optimal transport — the case p = 1. arXiv:1706.07650 [math.NA], last revised 5 Oct 2018. https://doi.org/10.48550/arXiv.1706.07650. Accessed 19 Dec 2023

[6] Feng K, Lu Z, Yang C (2018) Enhanced Morris method for global sensitivity analysis: good proxy of Sobol' index. Research Paper. In: Structural and Multidisciplinary Optimization (2019), vol 59. Springer, Heidelberg, pp 373–387. Available via SPRINGER LINK. https://doi.org/10.1007/s00158-018-2071-7. Accessed 19 Dec 2023

[7] Dellino G, Meloni C (2021) Uncertainty Management in Simulation-Optimization of Complex Systems. Springer, Heidelberg, pp 100–115, doi: 10.1007/978-1- 4899-7547-8

[8] Modelica Association (2023) Modelica Tools. https://modelica.org/tools/. Accessed 18 Oct 2023

[9] Fraunhofer Institute for Manufacturing Engineering and Automation IPA (2023) Asset Administration Shell (AAS). https://www.ipa.fraunhofer.de/de/Kompetenzen/kompetenzzentrum-digitale-werkzeuge-in-der-produktion/digital-twin/asset-administration-shell.html. Accessed 18 Oct 2023

HAZARD-FREE STEER-BY-WIRE IN ARTICULATED HEAVY EARTH MOVING MACHINERY USING CO-SIMULATION MODEL

peer reviewed

Vinay Partap Singh[1]*, Ville Raunio[1], Jesper Niemelä[1], Tatiana Minav[1]

[1]*Innovative Hydraulics and Automation, Tampere University, Tampere, Finland.*

* Corresponding author: Tel.: +358 465764198; E-mail address: vinaypartapsingh@tuni.fi

ABSTRACT

The articulated Heavy Earth Moving Machinery predominantly use hydrostatic steering, because of its reliability and redundancy. In earlier studies an energy efficient Electro-Hydrostatic Steering System was proposed, which works on Steer-by-Wire principle and comply with the safety standards. This paper presents co-simulation of a wheel loader model with hazard-free Steer-by-Wire. A co-simulation model using three software platforms; Mevea for multibody dynamics and mechanics, Simcenter AMESim for hydraulics, and MATLAB/Simulink for hydraulics, control, and data analysis, is created to analyse the hazard-free functionality of the steering. The simulation model of primary steering, which is an electric motor controlled Electro-Hydrostatic Actuator, is validated experimentally. In these heavy machines, as required by the standards there shall always be a secondary steering system for redundancy. The secondary steering, which is through a proportional control valve is modelled using the characteristics of the commercial product. There are five possible hazard scenarios in steering application of such machinery, which have been identified by the authors in earlier publication. These five hazard scenarios are realised in co-simulation model by modelling the respective faults in the primary steering, and the effectiveness of the hazard-free functionality in the steering is analysed. The study demonstrates that the novel Steer-by-Wire for articulated steering can effectively mitigate the potential hazards associated with steering in Heavy Earth Moving Machinery, moreover, co-simulation model provides an effectual mean to analyse the novel solutions.

Keywords: Articulated Steering, Steer-by-Wire, Hazard-free steering, Co-simulation model, Heavy Earth Moving Machinery

1. INTRODUCTION

The articulated steering is used commonly in Heavy Earth Moving Machinery (HEMM) for their superior performance in heavy loads and good manoeuvrability on rough terrain. The conventional articulated steering is powered by centralised hydraulic pump via priority valve and controlled by an orbital steering unit [1], [2]. Being a safety control element, steering shall be redundant and safe for operation under all operating conditions [3]. The redundancy in conventional articulated steering is achieved by the gerotor unit of the orbital steering unit, where the operator uses muscle power to steer the vehicle to safety. However, the conventional steering also has its limitations when it comes to safety as the response time and torque needed to control the vehicle in hazardous scenario depends on external factors including but not limited to vehicle speed. The human operator has limitations in terms of response time and muscle power, as a result of that, these machines are subjected to corresponding limitations specially in maximum allowed speed. Nevertheless, the electrification and autonomous trend in HEMM is changing the entire powertrain architecture and related safety requirements. The steering however because of its safety is unchanged up to a large extent in HEMM. The Steer-by-Wire (SbW) is one promising solution in this regard and has been commercially

accepted in passenger cars, but because of the power level and operating conditions, it has not been the same success in HEMM.

Some studies in past have focused on the subject, but they are limited to having an energy efficient digitally controlled steering, and safety aspect is not taken into account. A pump-controlled steering is proposed by Daher et al. [4] where the volumetric displacement of the hydraulic pump is controlled in order to control the steering, the study does not consider the redundancy. Similarly, Wang et al. [5] proposed a steering with dual- mode which has two steering control elements, hence not according to the ISO standards [6] which does not allow to have separate steering control elements for redundant steering. A Steer-by-Wire (SbW) for HEMM has been proposed by authors that has potential to fulfil the required safety standards in a previous article [1]. The said study has also demonstrated that the proposed SbW is significantly energy efficient as compared to the conventional steering. The functional safety of the proposed solution has been analysed in a later article [7] and it has been found that with even low diagnostic coverage, the proposed system satisfies the standards for safety critical parts of control system. In the same article, the effectiveness of the steering has been demonstrated using a co-simulation model for one major hazard scenario.

In this article, a co-simulation model of the articulated wheel loader is made with a hazard-free Steer-by-Wire. The co-simulation model is made using three different software for their different use and advantages. The steering system consist of a primary steering which is an electric motor-controlled electrohydraulic actuator (EHA), and secondary steering is realised by proportional valve through a priority valve as shown in **Figure 1**. The electric motor-controlled EHA model is validated on a test rig and used in the steering model except the hydraulic cylinder part. There are five hazard scenarios related to steering in a wheel loader which are identified in [1]. Those hazard scenarios are modelled in present article and the effectiveness of the proposed hazard-free SbW is analysed using the wheel loader co-simulation model.

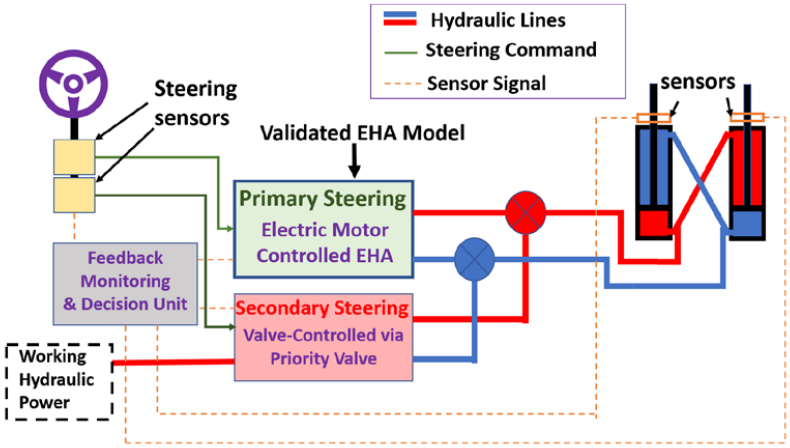


Figure 1: Topology of the novel hazard-free steering system

The next section describes the experimental setup of electric motor-controlled EHA and it's modelling and validation, as it has been used as the primary steering in the co-simulation model. The third section describes the co-simulation model in details and the steering architecture and operation of steering in hazard-free mode. The results of the steering effectiveness in different hazard scenarios are presented in fourth section where the results of the normal operation and in event of hazard related to steering are presented. Finally, the future aspects of the study are discussed with conclusion.

2. ELECTRO-HYDRAULIC ACTUATOR TEST-RIG AND MODEL VALIDATION

2.1. Experimental set-up

An electric motor-controlled EHA set-up is used to validate the model for the primary steering. The setup consists of a crane with double acting hydraulic actuator, where load on the tip of the crane can produce resistive and overrunning external load, i.e., 1st and 4th quadrant operations. The same test rig can be operated with three different EHA configurations with single or double pumps, the description of those is out of scope of this study. For this paper, the single pump EHA with pilot-operated check valves is used. The test rig and the schematics are shown in **Figure 2**, whereas the main components of the test rig are listed in **Table 1**.

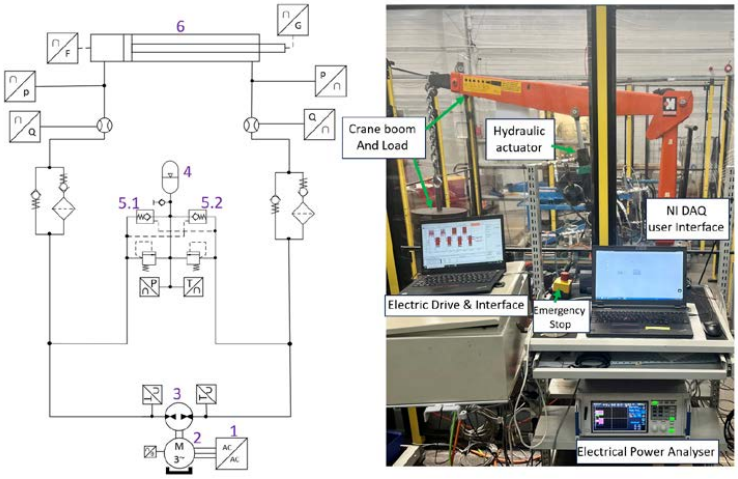


Figure 2: Hydraulic schematics and test-rig overview of electro-hydraulic actuator

Table 1: Main components of EHA and their specification

Component	Manufacturer and code	Main Parameters
Electric Drive	Unidrive SP1406	Nominal Power 5.5kW, Max Continuous output current 11A
Electric Motor	Emerson 115U2C300	Rated power 2.54 kW, Rated speed 3000 rpm
Hydraulic Pump	Casappa PLM20.14R0	14.53 cc/rev
Accumulator	Parker AD100A20T9A1	Volume 1L, max pressure 200 bar
Pilot Check Valve	Sun hydraulics CKCB	Cracking pr. 0.3bar, Pilot ratio 3:1
Hydraulic Cylinder	Miro C-10	60/30-400

The hydraulic system schematic shown on the left of **Figure 2** works in closed-circuit configuration. The Hydraulic pump (3) is driven by Electric Motor (2) which in turn is controlled by the electric drive (1). The differential volume of the hydraulic cylinder (6) is compensated by an accumulator (4)

through pilot-operated check valves (5.1) and (5.2). When the cylinder is in extension stroke and have a resistive force, the pressure in piston side of the cylinder is high which opens pilot check valve (5.2). The differential volume of cylinder means the flow going in the piston side of cylinder is higher than the flow coming out of the cylinder from rod side, in this extension stroke. The opening of pilot check valve (5.2) ensures the low-pressure side which is rod side of the cylinder in this case, is connected to accumulator which then acts as a pressurized reservoir and makes up the flow difference and pump supplies required flow to the piston side of the cylinder. The pilot check valves also ensure the minimum pressure in low pressure side of the hydraulic circuit during assistive load operation cycle. The passive hydraulic safety is ensured by the two pressure relief valves and the filters on both sides of the circuit entraps any contamination in hydraulic fluid.

The National Instruments' CompactRIO modules with LabVIEW Field-Programmable Gateway Arrays (FPGA) is used for data acquisition. The pressure is recorded on both sides of the cylinder and also for the accumulator, whereas the flow is only recorded for piston side of the cylinder, Trafag 8891 pressure sensors and gear type flow meter from Kracht are used for respective purpose. The position of the cylinder is recorded by draw-wire displacement sensor from Micro-Epsilon, whereas electric motor rotational speed is recorded in drive. There are other sensors in test-rig like temperature and force sensors, but their measurements have not been used for present article.

2.2. Modelling and validation

A simulation model of the test set-up is made using physics-based modelling in MATLAB/Simulink environment. The electric motor is controlled using field-oriented control strategy and the details about the electric motor model and control can be found in [8]. The Schlösser mathematical loss coefficient model [9] is used for hydraulic pump where the loss coefficients has been adjusted using experimental data. The modelling of rest of the hydraulic system has been done in similar way as [10] and [11], whereas the friction in hydraulic cylinder is realised using LuGre friction model. As, the load on the tip of crane is provided using the circular metal disc weights of 25 kg each, in simulation model the mass of other parts of the crane boom and load support along with mechanical advantage has been taken into account and the effect of extension of cylinder rod is accommodated using varying joint angles. The overall mechanical advantage including the mass of crane boom and load support varies from '1.594' when cylinder is fully retracted, to '1.686' when it is fully extended.

The results of experimental and simulation model are shown in **Figure 3** and **Figure 4** for two different speed and load profiles. Because of certain temporary physical limitation on test-rig, the operation is subjected to limitations on power of electric motor and stroke of hydraulic cylinder.

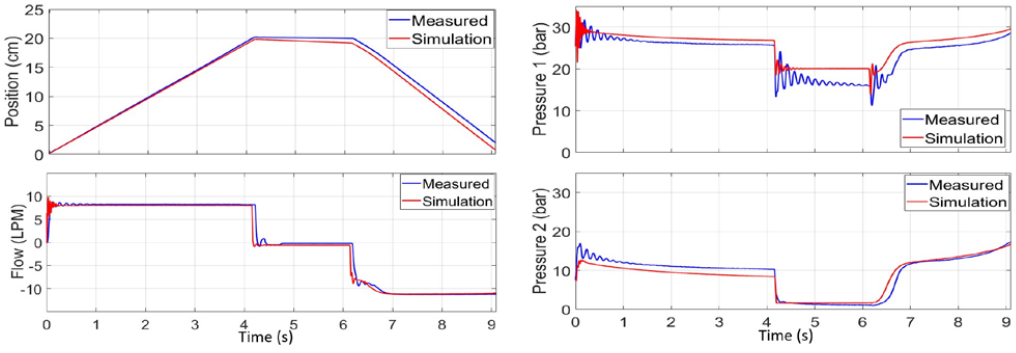


Figure 3: Measurement and simulation results for 600 rpm and 175 kg load

Figure 3 shows the results for 600 rpm of electric motor speed and load mass of 175 kg on the tip of the crane. It is worth noting that the actual load acting on the hydraulic cylinder will be higher because of mass of crane boom, load support, and mechanical advantage of the boom. The figure shows position of cylinder, flow in the piston side of cylinder, and pressure on piston and rod side of cylinder termed as 'Pressure 1' and 'Pressure 2', respectively.

Figure 4 shows the results for operation at 400 rpm of electric motor speed and 100 kg of load mass on the tip of the boom. As the hydraulic accumulator is an important part in this kind of closed-circuit configuration, the pressure is plotted for experiment and simulation model in **Figure 5.** The model performs satisfactorily and agrees with experimental data for both operating conditions.

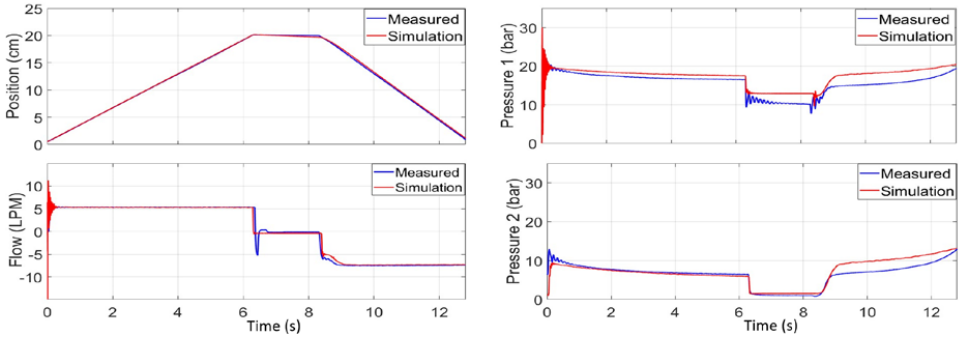


Figure 4: Measurement and simulation results for 400 rpm and 100 kg load

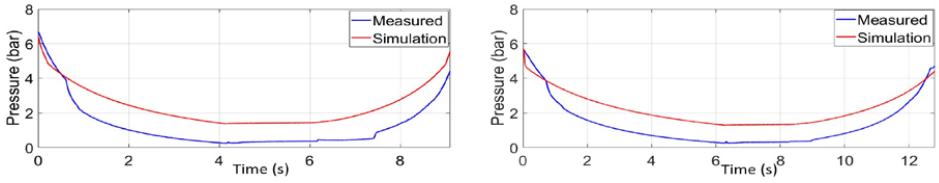


Figure 5: Accumulator pressure at 600 rpm (left) and 400 rpm (right)

3. CO-SIMULATION MODEL OF WHEEL LOADER WITH HAZARD-FREE STEER-BY-WIRE

3.1. Co-simulation model of Wheel Loader

The co-simulation model of a 16-ton wheel loader is made using three software MATLAB/Simulink, Simcenter AMESim, and Mevea [12] for their respective advantages. Mevea is a digital twin technology and simulation platform, its own physics engine simulates mechanics, hydraulics, power transmission, and operating environment of the machine. Siemens' AMESim has advantage of object-oriented programming with its inbuilt hydraulics, mechanics, and interface libraries. Whereas MATLAB/Simulink has the advantage of better data handling, controller design, and ability to interact with multiple software simultaneously. The wheel loader multibody dynamics model which interacts with real world like environment is modelled in Mevea, the steering actuators along with secondary steering hydraulics and connection of primary and secondary steering are done in AMESim, whereas the previously described electric motor-controlled EHA model with overall control, steering command, and signal monitoring is in MATLAB/Simulink. The operation of the wheel loader can be controlled via a physical joystick, keyboard control, or a script of commands. **Figure 6** shows the overall structure of the co-simulation model. The blocks in **Figure 6** represents

the part of the entire co-simulation model realized in respective software, whereas arrow shows the direction of data flow between them. It is worth noting that only steering is powered and controlled by MATLAB/Simulink and AMESim, the rest of the operation including driving the wheel loader and working implements is solely modelled and controlled in Mevea. The steering operation can also be controlled using an external joystick or keyboard control, but to keep the uniformity of command, steering reference is given in MATLAB/Simulink. The data flows in real time between different software which is handled in MATLAB/Simulink. The movement of the wheel loader is observed visually in Mevea while all the data and control are monitored in MATLAB/Simulink.

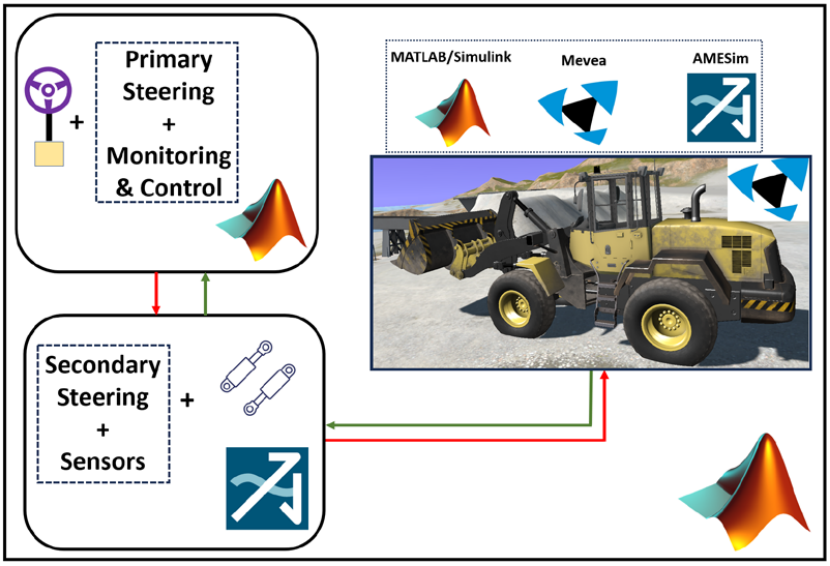


Figure 6: Co-simulation model overall structure

3.2. Hazard-free Steer-by-Wire

Figure 7 shows the schematics of the steering system where primary steering is realised via previously described validated model of electric motor-controlled EHA whereas secondary steering is using proportional valve through priority valve. More details of the steering system, its operation, and it's potential to comply with required standards can be found in [1], [7]. In normal conditions, the main steering operation is performed by controlling the electric motor in EHA, the operation of which is described in **section 2**. In case of a hazard scenario, the locking valves (4.1, 4.2) mechanically isolates the primary steering and operation can be continued with secondary steering. In secondary steering, the priority valve (5) sends the flow to the other hydraulic function when there is no demand for steering application. When the secondary steering is activated, and the proportional valve (6) have flow demand the flow is directed to steering operation. The control block diagram and signal flow for the above-mentioned hazard-free steering operation of wheel loader co-simulation model are presented in **Figure 8**. The main values and parameters of the components used in the steering are presented in **Table 2**. It shall be noted that the parameters of the electric motor-controlled EHA except hydraulic cylinder, which is used as primary steering are same as described **Table 1** in **Section 2.1**.

For this paper to keep the uniformity of the operation, the steering actuator position command (X_{ref}) is provided in MATLAB/Simulink, whereas the rest of the operation of the wheel loader is controlled in Mevea via keyboard control. The same command is provided separately for primary and secondary

steering to make model partially in accordance with category 3 architecture of ISO 13849-1 [13], however the fault diagnosis does not fully comply with the said category, so the overall steering co-simulation model is not entirely in compliance with category 3 architecture. Nevertheless, it is worth noting that the overall architecture has potential to fully comply with category 3 architecture, this has been described in detail in [7].

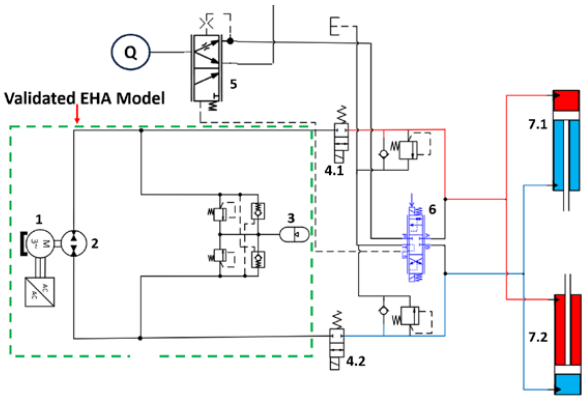


Figure 7: Schematics of steering system

Table 2: Parameters of steering system components

Component	Main Parameters
Steering actuator	80/50 stroke – 340mm
Proportional valve	Nominal flow – 50 LPM, corresponding pressure drop - 15 bar, LS port.
Priority valve	Characteristic flow – 70LPM, corresponding pressure drop- 20 bar, spring pressure – 10 bar.
Constant flow source	68 LPM
Locking valves	Nominal flow – 70LPM, corresponding pressure drop – 5 bar.
Pressure relief valves	210 bar with anti-cavitation.

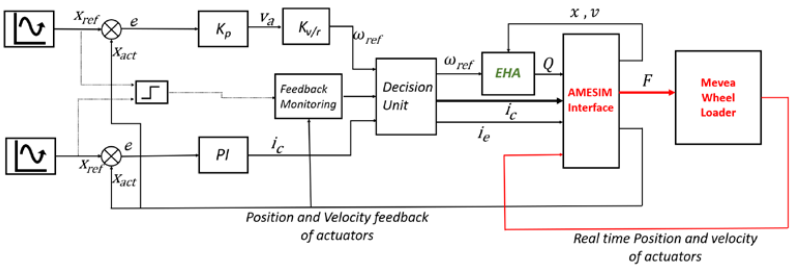


Figure 8: Control diagram and signal flow

For the primary steering, the actual position of the steering actuator (X_{act}) is taken from AMESim where virtual sensors are used. In primary steering, the rotational speed command to the electric motor (ω_{ref}) is provided by simple P controller (K_p). The velocity command of the linear steering actuator (V_a) is converted in rotational speed command of the electric motor by coefficient ($K_{v/r}$) [2]. The validated EHA model takes the electric motor speed reference, and position and velocity of the steering actuators (x,v) from AMESim, and provides the hydraulic flow (Q) to AMESim where steering actuators are present. For the secondary steering, a PI controller is used, which provides the valve command signal (i_c) to the load sensing proportional valve in AMESim. The integral element (I) is used here because of the additional non-linearity introduced by the dead zone, which further is used because of the characteristics of proportional valve. The feedback monitoring unit receives the position and velocity of the actuators along with the steering signal in form of 1-, 0, or 1, according to the presence of steering signal and its direction. The decision unit takes the command signal from both steering channel and signal from feedback monitoring unit. In normal operation, the decision unit directly sends the electric motor speed reference command to EHA and locking valve signal (i_e) accordingly. If it detects any anomaly in the steering command and the actual steering signals, it isolates the primary steering by stopping the EM and deenergising the locking valves for rest of the operation and send the secondary steering command signal onwards.

The steering force command (F) goes to the wheel loader in Mevea, and based on the Mevea solver, the real time position and velocity of both the steering actuators is received back to AMESim, where it converts to corresponding force and pressure requirement. The MATLAB/Simulink environment connects, monitors and handles the data flow between different software in real time.

4. HAZARD-FREE STEERING OPERATION OF WHEEL LOADER

To demonstrate the behaviour of wheel loader in case of a hazard scenario in steering a normal cycle of the steering is chosen so that the steering actuators complete one full cycle of movement, whereas the wheel loader was moved on similar path. The different hazard scenarios are injected by altering the speed command of the electric motor in primary steering accordingly. The severity, possible cause, and consequences of these hazard scenario are described in detail in [1].

First, the wheel loader is operated with primary steering only without any faults, to compare the results further with the hazard scenario. **Figure 9 (left)** shows the steering actuator position, electric motor speed, and pressure in both chambers of the steering actuator. The position of only one actuator is shown as command is given to control the single actuator position and the other moves in mirror action. As the wheel loader in Mevea interacts with real-world like environment which includes the sand, the gravel and other forces on a non-uniform ground, the effects of these uncertain ground forces on steering can be seen in the position and pressure response. The dynamic response of electric motor-controlled EHA can be improved with better control strategies as demonstrated in [14] [15] but considered to be out of scope of current study. **Section 4.1** to **Section 4.5** shows the results for different hazard scenarios while **Table 3** summarise the results briefly.

The faults are considered only in primary steering channel for the study i.e., EHA, as purpose of this part of the study is to analyse the overall behaviour and effectiveness of redundant steering channel in case of different hazard scenarios. However, there is always a possibility for faults in other parts of the system including secondary steering channel, and sensors, for which the relevant safety guidelines shall be followed for redundancy in secondary channel.

4.1. Hazard 1: - Loss of power in primary steering

The loss of power in primary steering is a major hazard and can become even more dangerous if

machine is working at a place where co-workers or bystanders are present. This is modelled by forcefully stopping the electric motor of EHA at a time when steering command is present. The electric motor is stopped at t= 7s, and the secondary steering became effective in 600ms, and vehicle is controlled. It is worth noting that this time includes time in detections of fault, isolating primary steering and secondary steering controlling the vehicle. As can be seen by electric motor speed and valve command plot, the electric motor stops at 7s and after 600ms, the valve fully takes the control of the operation. During the loss of steering power, the pressure in steering actuators jumped to a high peak because of effectively absence of the control of steering and presence of external forces. Figure 9 (right) shows the steering actuator position, the excerpted view at the time of hazard, electric motor speed and valve command for respective modes of steering, and pressure in both chambers of the steering actuators. The subsequent hazards scenario sections show the similar plots for respective hazards.

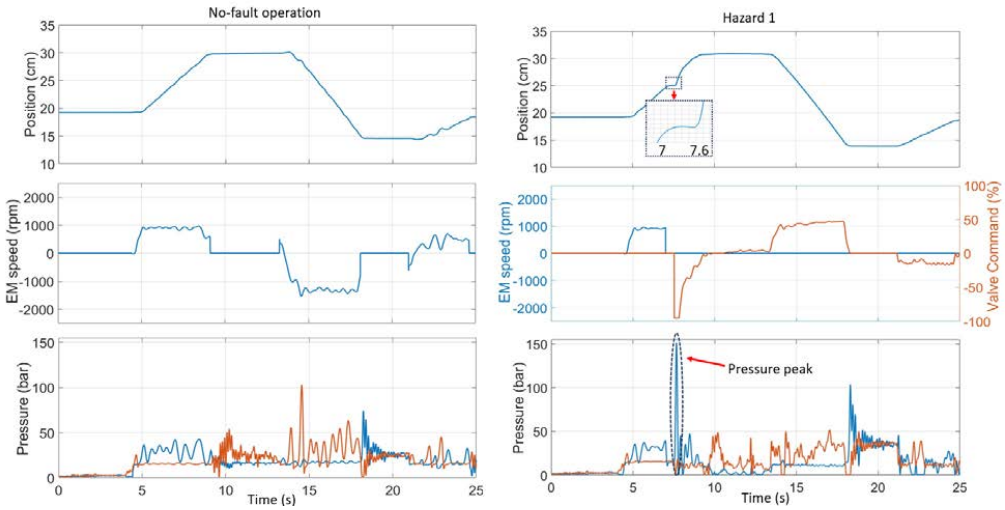


Figure 9: Operation with EHA steering (left), loss of power in primary steering (right)

4.2. Hazard 2: - Unintended vehicle steer

Unintended vehicle hazard is modelled by starting the electric motor in primary steering at t= 3s, when the steering command is not present. The unintended movement of the steering is detected, and the primary steering is isolated for rest of the operation and operation is completed with secondary steering. **Figure 10 (left)** shows the similar plots as in Hazard 1 case, the total time to control the vehicle by secondary steering is around 400ms. The pressure peaks are also observed in this case even if there is a control element present, as the electric motor speed is increased suddenly and that tries to move steering actuators with a very high speed, resulting in high forces. Further the switching between primary and secondary steering created the sudden change in pressures of the steering, hence the peak can be observed in both sides of the chamber.

4.3. Hazard 3: - Steer in opposite direction as commanded

The rotational speed command of the electric motor of EHA is inverted to create the scenario when the steering operation starts in opposite direction as commanded. As a result of that, when the steering command starts at t=4.6s, the steering actuator moves in opposite direction as commanded, the results are shown in **Figure 10 (right).** It took overall around 500ms to detect the fault and start the operation by secondary steering, while it took more time to catch up with the actual command. The pressure

peaks in this case can be attributed to the switching between primary and secondary steering, as the locking valves close and the proportional valve starts operation. The proportional valve opens to its maximum capacity as it tries to bring the actuator back to the normal position quickly, resulting in sudden pressure rise, the other side of pressure meanwhile takes a dip.

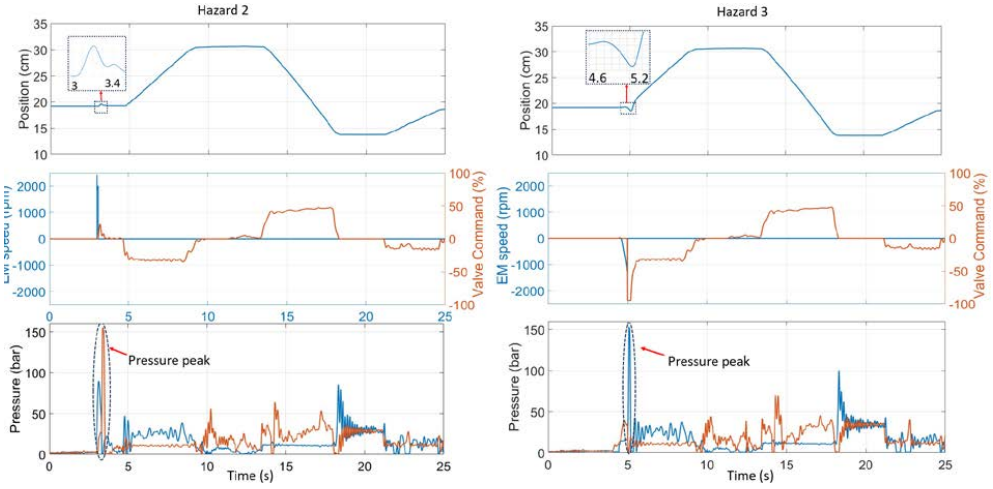


Figure 10: Unintended vehicle steer (left), steer in opposite direction as commanded (right)

4.4. Hazard 4: - Oversteer

Oversteer can be termed as a condition when the steering actuators move ah higher speed than commanded. The speed of the electric motor in EHA raised to a value higher than actual command at t=16s to model this which results into faster movement of the actuator. It took around 600ms for the system to detect the fault and start the operation with secondary steering, as shown in **Figure 11 (left)**. The pressure peaks in this case have similar behaviour and can be explained by same phenomena, as in the case of opposite direction steering in previous section. It can be observed that the peaks occur not at the time of hazard, i.e., t= 16s, but when the secondary steering is activated.

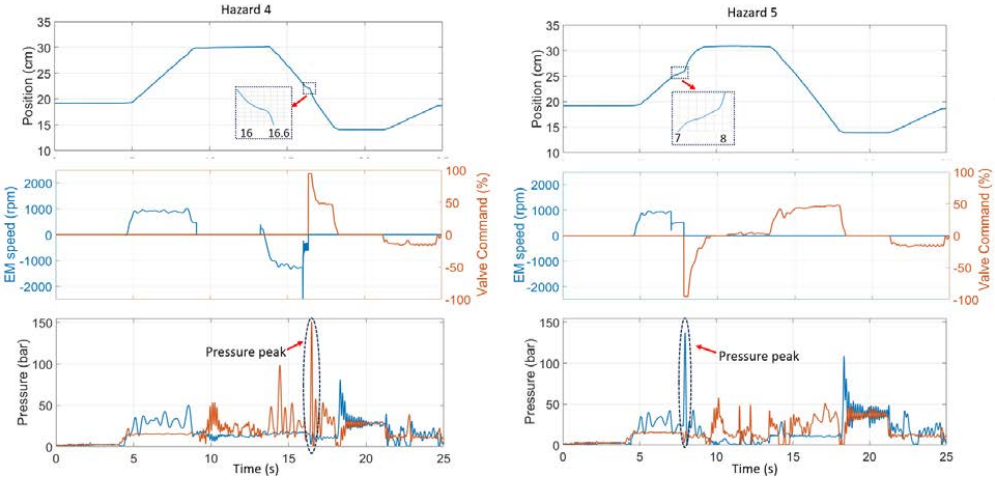


Figure 11: Oversteer (left), understeer (right)

4.5. Hazard 5: - Understeer

Understeer can be termed as when the steering actuator moves at less velocity then commanded. Similar to the oversteer case, the electric motor speed in primary steering is reduced to an arbitrary value at t=7s, which resulted in slower movement of the actuator. Although there is a fault in the primary steering, it is still moving, hence it took almost 1s to detect it and start operation with secondary steering. Once more the pressure peaks in this case have similar behaviour as previous two hazard scenarios and have the similar reason behind them.

Table 3: Summary of the Hazard scenarios

Hazard Scenario	Fault in Primary steering simulation method	Reaction time
Loss of power in primary steering	EM stopped when steering command is present	600ms
Unintended steer	EM started without steering command	400ms
Steer opposite then commanded	EM steed inverted as commanded	500ms
Oversteer	EM speed raised higher than required	600ms
Understeer	EM speed decreased than required	1s

5. CONCLUSION AND OUTLOOK

In this study a co-simulation model of a wheel loader with hazard-free steer-by-wire is created. The hydrostatic steering consists of a primary steering which is an electric motor controlled electro-hydraulic actuator, is experimentally validated on a test bench. The five possible hazard scenarios related to steering in a wheel loader operation are modelled, and the effectiveness of the novel steering system is analysed. It has been demonstrated that the novel steering system is effective in any hazard scenario where primary steering have some faults. The time of reaction to activate secondary steering in event of a hazard varies for different failure. However, with suitable control strategies and effective fault diagnosis in the steering, the reaction time can be improved. The study demonstrates the novel steer-by-wire for heavy earth moving machinery is effective in a hazardous scenario in steering, also the co-simulation model can be an effective solution to study novel solutions.

The study of steering and the analysis of intended functionality and behaviour is strictly within the domain of simulation using partially validated model, in this article. There is no hardware in control or sensor elements is involved, which will contribute to additional complexity in prototyping. The future work in the direction includes advanced fault detection techniques using machine learning and deep learning, followed by prototyping and proof of concept.

ACKNOWLEDGEMENTS

This research is supported by Finnish Foundation for Technology Promotion (Tekniikan Edistämissäätiö) grant no. 8917 and Business Finland project EMMA-2 (2471/31/2021).

REFERENCES

[1] Singh, V. P., Huova, M., & Minav, T. (2023). Energy Efficient Steer-By-Wire in articulated non-road mobile machines: Analysis and proposal. In Proceedings of the 18th Scandinavian International Conference on Fluid Power, SICFP (Vol. 23, pp. 1-15).

[2] Danfoss. (2015). EHPS Steering Valve PVE Actuation Module OSPCX CN Steering Unit. May 2015 BC152886484652en-000502.pdf (danfoss.com) Accessed Aug 2023.

[3] SFS-EN ISO 19014-1:2018:en Earth-moving machinery. Functional safety. Part 1: Methodology to determine safety-related parts of the control system and performance requirements (ISO 19014-1:2018, Corrected version 2019-02).

[4] Daher, N., Wang, C., & Ivantysynova, M. (2013). Novel energy-saving steer-by-wire system for articulated steering vehicles: A compact wheel loader case study. In 13th Scandinavian International Conference on Fluid Power; June 3-5; 2013; Linköping; Sweden (No. 92, pp. 541-552). Linköping University Electronic Press.

[5] Wang, Y., Liu, X., Chen, J., Chen, W., Li, C., & Huo, D. (2022). Design and control performance optimization of dual-mode hydraulic steering system for wheel loader. Automation in Construction, 143, 104539. https://doi.org/10.1016/j.autcon.2022.104539

[6] SFS-EN ISO 5010:en (2019) Earth-moving machinery. Wheeled machines. Steering requirements (ISO 5010:2019), International Organization for Standardization.

[7] Singh, V. P., Huova, M., & Minav, T. (2023, October). Simulation Study of a Fail-Safe Steer-by-Wire for Heavy Earth Moving Machinery. In Fluid Power Systems Technology (Vol. 87431, p. V001T01A073). American Society of Mechanical Engineers. https://doi.org/10.1115/FPMC2023-113624

[8] Zakharov, V., & Minav, T. (2020). Analysis of field oriented control of permanent magnet synchronous motor for a valveless pump-controlled actuator. In Proceedings (Vol. 64, No. 1, p. 19). MDPI. https://doi.org/10.3390/IeCAT2020-08491

[9] Schlosser, W. M. J. (1961). Mathematical model for displacement pumps and motors. Hydraulic power transmission, 7, 252-328.

[10] Minav, T. A., Sainio, P., & Pietola, M. (2014). Direct-driven hydraulic drive without conventional oil tank. In Fluid Power Systems Technology (Vol. 45974, p. V001T01A022). https://doi.org/10.1115/FPMC2014-7834

[11] Qu, S., Fassbender, D., Vacca, A., & Busquets, E. (2021). A high-efficient solution for electro-hydraulic actuators with energy regeneration capability. Energy, 216, 119291. https://doi.org/10.1016/j.energy.2020.119291

[12] Mevea Ltd. (2022). Mevea Real-Time Simulation Software | Digital Twins Techology. Mevea. https://mevea.com/solutions/software/ Accessed Aug 2023

[13] SFS-EN ISO 13849-1:2015 (2015) Safety of machinery - Safety-related parts of control systems - Part 1: General principles for design. (ISO 13849-1:2015) International Organization for Standardization.

[14] Hagen, D., Padovani, D., & Choux, M. (2019). A comparison study of a novel self-contained electro-hydraulic cylinder versus a conventional valve-controlled actuator—part 1: motion control. In Actuators (Vol. 8, No. 4, p. 79). MDPI. https://doi.org/10.3390/act8040079

[15] Zhang, S., Li, S., & Minav, T. (2020). Control and performance analysis of variable speed pump-controlled asymmetric cylinder systems under four-quadrant operation. In Actuators (Vol. 9, No. 4, p. 123). MDPI. https://doi.org/10.3390/act9040123

Chapter 18

System Design and Architecture-II

SOLUTIONS FOR ENERGY-EFFICIENT AND EASY IMPLEMENTABLE ELECTRIFIED VARIABLE-SPEED PUMP DRIVES IN MOBILE APPLICATIONS

Dr. Steffen Rose[1]*, Dominik Hoffmann[1], Fabian Wiedmer[1], Viktor Rill[2]

[1]*Bosch Rexroth AG, Lise-Meitner-Strasse 2, 89081 Ulm, Germany*
[2]*Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany*
* Corresponding author: Tel.: +49 731 49373-550; E-mail address: steffen.rose@boschrexroth.de

ABSTRACT

Mobile machinery is currently undergoing a transformation towards an increasingly electrification of its powertrain systems, which is subject to ever-increasing demands in terms of performance, functionality, noise behaviour, integration and – above all – energy efficiency. Especially in the latter case, solutions on a system level appear more and more attractive as the advantages of electric drives can be fully revealed this way – most of all the high torque build-up dynamics and speed variability across a much wider speed range compared to combustion engines that also involves standstill as well as negative speeds. Following this idea, the contribution introduces a model-based and holistic approach to achieve a customer-specific system design that is targeted at a given individual load specification for the energy-efficient supply of the hydraulic implement systems with a demand-oriented volume flow at an optimal speed. Both, the application-optimized dimensioning of the drive components and the use of an operating strategy, lead in combination to significant energy savings during operation. These improvements are presented in this work, which is based on simulations and test bench measurements. Furthermore, this paper concludes with an investigation of the thermal behaviour of the electric drive using different speed operation strategies.

Keywords: Electrification, Mobile Machinery, Optimal Speed Strategy, Energy Efficiency, Model-Based Design

1. INTRODUCTION

1.1. Motivation and system solutions

Next to zero local emissions, the potential for much higher energy efficiency is a main motivation and driver for the increase in numbers of electric drives in mobile machines. Both aspects can be seen as an important basis for future CO_2 neutrality [3]. For battery-electric systems, saving energy plays a special role as well, because battery capacities are still rather limited and the amount of energy carried is currently still significantly smaller than in diesel-driven machines [2]. For the supply of hydraulic implement systems in these electrified machines, electric motor-driven pumps become more and more available. In addition to the already high efficiency of the electrical components, they allow a more consistent speed adjustment in a much larger range compared to diesel-driven systems due to their advantages in dynamics, like providing maximum torque from zero speed [7]. This results in further potentials for energy savings at system level.
The adjustable speed of an electrified pump system combined with commonly used variable displacement pumps (demand-based volume flow supply) provide a degree of freedom which allows a load and system state depending optimal utilization of all drivetrain components involved during the operation [1]. In order to achieve these goals and at the same time to take into account the very

high variance of mobile applications themselves as well as their different application profiles, this contribution introduces a model-based and holistic approach to achieve a customer specific system design that is targeted at the customers' load specifications for the supply of their hydraulic implement systems with a demand-oriented volume flow at an optimal speed.

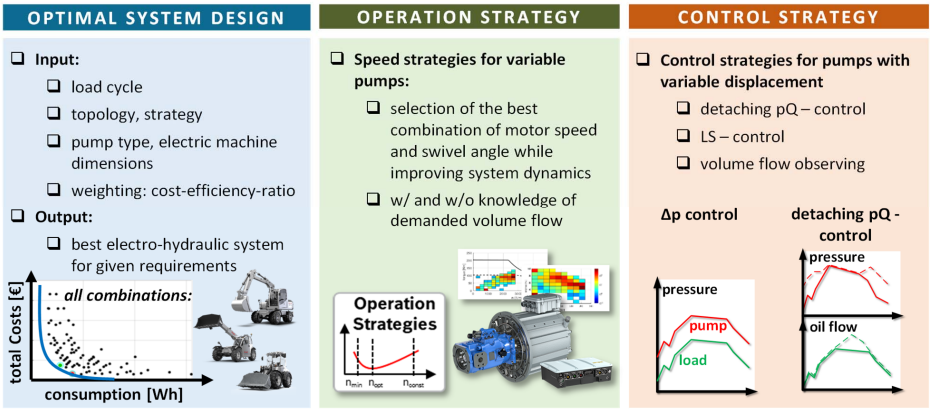


Figure 1: System solution approach for electrified pump drives

For the ideal dimensioning as well as the subsequent operation of an electrified pump drive, a model-based design tool is used, as shown in the left field in **Figure 1**. With the knowledge of the application, the considered load cycles, the relevant component data as well as the preferred operating strategy, a picture of all feasible combinations can be generated in a very efficient way.
Furthermore, real-time capable operating and control strategies adapted to the system can be implemented as software functions, as depicted in the middle field of **Figure 1**. These strategies allow the operation of variable-speed pump drivetrains to supply any common valve systems, such as for example load sensing (LS) and electronic positive control (ePC) systems, considering the efficiency maps and fundamental dynamic effects of all involved drive components.
In addition, a pump control system that is optimally selected for the respective application ensures the supply of a demanded volume flow and – if applicable – also a specified pump pressure limitation. Such control strategies can be, for example, a pressure difference control for LS valve systems or an alternating pressure or flow control for electronic flow matching (EFM) or ePC systems, like shown exemplarily in the right field of **Figure 1**. However, both control strategies can result in situations where the volume flow demand does not mandatorily correspond to the actual volume flow of the pump (pressure or power limitation in intervention and/or volume flow demands generally not known). For these cases, developed mechanisms ensure a stable behaviour and energetically efficient operation at the same time. All of these measures enable the best possible operation of the mobile machine in terms of performance, functionality and energy efficiency and thus also makes a notable contribution to energy, CO_2 and cost savings.

1.2. Related work and contribution of this paper

Electro-hydraulic pump systems which are using the inherent two degrees of freedom – variable speed and variable displacement – have been in focus especially for stationary industrial applications for a couple of years already. An example is hydraulic presses, whose energetic behaviour has been experimentally studied by Huang et.al. [1], varying speeds and loads. Furthermore, Zheng [9] proves the feasibility of variable speed drives also for pressure-controlled pumps in the same context.
Reidl et.al. [5] show that energy savings of up to 29% can be achieved for electro-hydrostatic axes.

Therefore, a loss optimal displacement trajectory is determined in this study, which presupposes the knowledge of the load cycle. Complex models of the pump, in which neural networks are used, form the basis for the dynamic optimization algorithm. The specific application with all its drive and transmission components is known and all their characteristics are considered within the optimization model. Moreover, neural networks, here in backpropagation form, can be found in Rui et al. [6] to set the optimal speed and swivel angle based on the measured pressure and flow rate of the system.
Furthermore, many solutions provide an offline determination of the trajectories for optimal swivel angle and optimal speed – see for example [4] and [5]. In [4], two different algorithms are investigated by Reichert et.al. One algorithm looks purely at the quasi-stationary states during operation and the second algorithm is a model-based approach that also includes dynamic effects. With Willkomm [8] the consideration of very dynamic processes is in the foreground, whereby an approach is pursued with the help of model predictive methods, to achieve – above all – very high process dynamics with simultaneous energy efficiency.
In contrast to the related work mentioned above, this contribution will take up the aspect of the speed variability of a pump drive and examine especially its energetic advantages in the context of mobile machines. The focus is on a model-based approach, which is used both in the already mentioned speed strategies but also in a dimensioning tool. With this tool, dimensioning becomes automated and much easier, as a variety of different component size combinations can be compared with each other regarding several criteria that go beyond energy efficiency. In addition, both, an optimal design and the use of an optimal speed strategy have beneficial effects on the thermal behaviour mainly of the electrical components. In this contribution, the results of an exemplary design for the application mini excavator are presented and validated by measurements on a test bench. In the referenced works, the corresponding load cycle is often very specific and well known. For this cases, offline calculated trajectories are often used. This is generally not possible with mobile machines, as the loads and speeds typically depend on the machine type, work task, operator, and potentially many other influences. Therefore, online strategies or approximations can be used.

2. MODEL BASED APPROACH FOR OPTIMAL ELECTRIFIED PUMP DRIVES

2.1. Implementation of a speed-variable electrified pump drive system

Figure 2 shows the schematic representation of the relevant assembly and top-level functional architecture of a variable-speed variable-displacement pump drive for electrified mobile applications, which can be seen here in the implementation for a typical digitized control structure of a todays hydraulic implement system. The essential relationships of the assembly and the function groups with each other but also in the context of the implement system are indicated in the form of physical interfaces and signal flows. Here, an implement control unit evaluates the operator demands, for example joystick signals, and determines the control signals for the individual valve units of the main control valve (MCV). In the following, these valve units measure the desired flow rates for the hydraulic loads and the available pump flow rate is distributed according to the operator demands and the valve system characteristic. The exemplary valve system shown here is a LS system, but other valve systems such as ePC or EFM systems are also possible, in which the nominal volume flow of the pump can be determined from the operator demands.
The task of the electrified pump drive system is now to provide the required pump volume flow according to the operator demands and with regard to further criteria in an optimal way. For this purpose, the volume flow request derived from the operator must be made available as a signal. Based on the selection of one of the operating strategies, which are implemented as software on a control unit, this signal is used for the calculation of the desired speed of the electric drive unit driving the pump. This generated speed setpoint signal is simultaneously transferred - alternatively also the signal

of the actual speed - to the pump controller, which, depending on the selected control type (e.g., volume flow control, LS control, etc.) adjusts the corresponding swivel angle so that the pump can provide the requested volume flow.

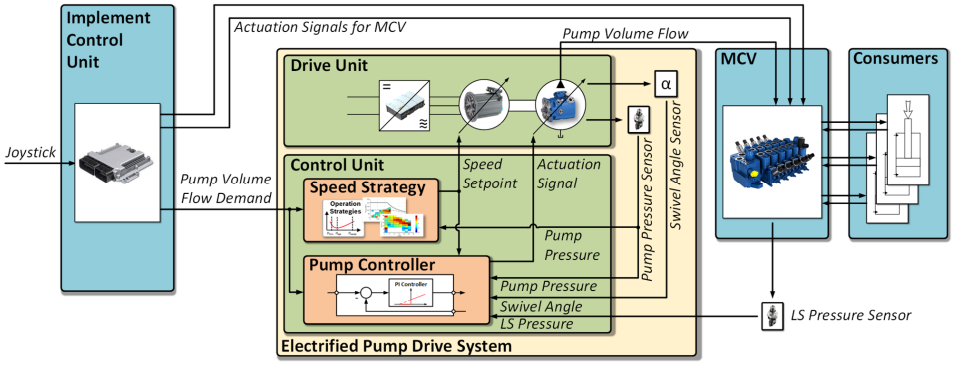


Figure 2: Schematic representation of the electrified pump drive system

By considering the different characteristics related to the energetic behaviour of the electric drive (inverter and electric machine) and the hydraulic pump, the different strategies for speed adjustment can be derived. For this purpose the respective input and output powers of the components involved are to be considered. The following relationship applies to the hydraulic pump:

$$P_{hyd} = Q_{eff} \cdot \Delta p = \eta_V \cdot \frac{\omega}{2\pi} \cdot V_g \cdot \Delta p = \eta_V \cdot \omega \cdot T_{ideal} = \eta_V \cdot \eta_{mech} \cdot \omega \cdot T_{eff} \quad (1)$$

$$P_{hyd} = \eta_V(n, \alpha, \Delta p) \cdot \eta_{mech}(n, \alpha, \Delta p) \cdot P_{mech} = \eta_{Pmp}(n, \alpha, \Delta p) \cdot P_{mech} \quad (2)$$

According to Equation (2), the hydraulic power at the pump outlet results from the mechanical power at the pump shaft multiplied by the volumetric efficiency η_V and the mechanical efficiency η_{mech}. These efficiency values are very often available as measured lookup tables for the leakage (η_V) as well as for the friction (η_{mech}). The input variables of both tables are the speed n and the swivel angle α of the pump as well as the pressure difference Δp between suction and pressure ports. The mechanical power at the input of the pump can be assumed to be equal to the mechanical output power of the electric machine, so no equally lossy transmission is arranged between the electric machine and the pump. The output power P_{mech} of the electric machine can also be determined by an efficiency ratio as shown below:

$$P_{mech} = P_{EmOut} = \eta_{Em}(i_{Ph}, u_{DC}, n) \cdot P_{EmIn} \quad (3)$$

The efficiency map covers both, the occurring mechanical losses as well as the electrical iron and copper losses of the permanent magnet synchronous machine (PMSM) underlying these investigations. To drive the synchronous machine, an electric rotating field is required, which is generated in its three windings. The necessary three-phase current is provided by the electrical inverter, wherein the variable frequency of the for this generated three-phase voltage is responsible for the variability of the speed of the motor and thus also of the pump. Like the pump and electrical machine, a relationship between the input power and the output power of the inverter can also be find out via the determination of a characteristic efficiency map:

$$P_{InvOut} = P_{EmIn} = \eta_{Inv}(i_{Ph}, u_{DC}, n) \cdot P_{ElecIn} \quad (4)$$

The loss behaviour of the semiconductor power elements used in the inverter is mainly described by the phase current i_{Pha} and by the active power factor $cos\ \varphi$, which depends on the utilized electric machine and is mainly characterized by its inductive behaviour. Since $cos\ \varphi$ as well as the efficiency of the electric machine can also be determined by the physical quantities phase current i_{Phi}, DC link voltage u_{DC} and speed n, the equations (3) and (4) can be combined and the corresponding powers can be described by the overall efficiency characteristic η_{eDrv} of the electric drive. For the overall drive system including the hydraulic pump, the following equation results:

$$P_{hyd} = \eta_{eDrv}(i_{Ph}, u_{DC}, n) \cdot \eta_{Pmp}(n, \alpha, \Delta p) \cdot P_{ElecIn} \tag{5}$$

These relationships are depicted in the following **Figure 3**. Shown are the corresponding losses of the pump, the electric machine and the electric inverter. The latter can be combined into an electric drive and be handled in the further course of the explanations in this contribution.

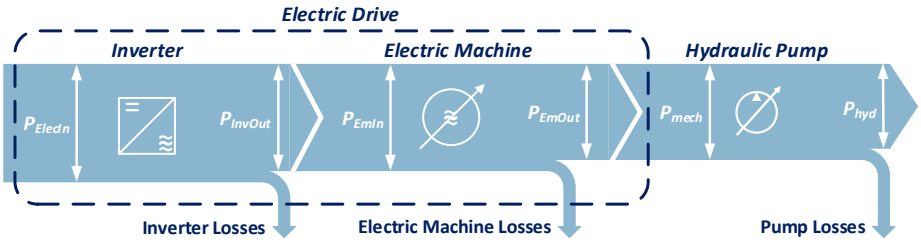


Figure 3: Schematic representation of the power flow within the electrified pump system

2.2. Speed-variable strategies to increase energy efficiency

The following **Figure 4** shows exemplarily the characteristic efficiency behaviour of both the electric drive and the hydraulic pump:

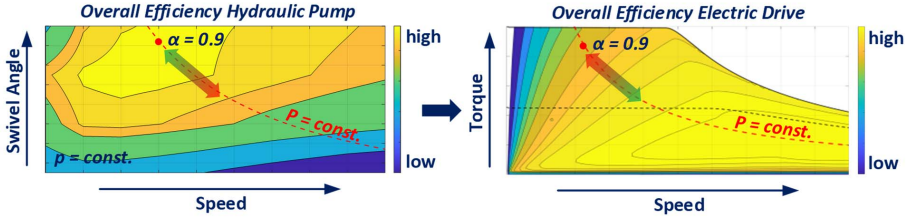


Figure 4: Exemplary efficiency maps of the hydraulic pump (left) and the electric drive (right)

The illustration on the right shows the typical torque behaviour of a permanent magnet synchronous machine in motor operation. The solid line shows the course of the peak torque, and the dashed line represents the course of the permanently available torque. Both torques are almost constant up to the nominal speed of the electric machine, in order to then fall off when the field weakening operation is reached at constantly increasing speeds. With increasing speeds, however, the energy efficiency of the electric drive system increases. The left illustration of **Figure 4** shows the characteristic efficiency behaviour of an axial piston pump with variable displacement volume that was used in this study. Assuming a constant pressure, one can see that high efficiency values are achieved, especially for larger swivel angles at lower speeds (which result in higher torques for the given hydraulic power demand). For the electric drive it is the opposite in that case. From this, two elementary strategies for setting a speed can be derived, which should enable a more energy-efficient overall behaviour of the electric pump drive compared to constant speed. If only the pump behaviour is considered, one would

always choose the largest possible swivel angle in order to be able to operate at least the pump in an energy-optimal range. Such an implementation can be achieved by applying the following equation for determining the pump volume flow with α close to 100%:

$$n_{min} = \frac{Q_{eff}}{\eta_V \cdot \alpha \cdot V_{gmax}} = \frac{Q_{des}}{\eta_V \cdot \alpha \cdot V_{gmax}} = \frac{Q_{des}}{\eta_V(n, \alpha, \Delta p) \cdot \alpha \cdot V_{gmax}} \tag{6}$$

Starting from the equation for describing the effective volume flow Q_{eff} of the pump, which simultaneously corresponds to the desired volume flow Q_{des}, the necessary speed is determined according to equation (6). This is the minimum necessary speed at which the volume flow can be realized and is therefore referred to as the minimum speed n_{min} in the further course. For the volumetric efficiency η_V, either a simple constant or the efficiency map of the pump can be used, which must be determined in advance and implemented on the control unit.
The minimum speed strategy impresses with its simple implementation and corresponds to the preference to adapt the speed to the actual volume flow demand. However, the losses caused by the electric drive are not considered. If you take a closer look at the efficiency behaviour of the pump and the electric drive, you can see that increasing the efficiency of both drive technologies requires opposite speeds. This suggests an optimum, which can be described by the following optimization problem P, in which the electrical input power P_{ElecIn} based on equation (5) is minimized online assuming that the hydraulic input power P_{hyd} and further system states are known:

$$n_{opt} \rightarrow P: \quad \min_n \left(\frac{p_{pmp} \cdot Q_{des}}{\eta_{Pmp}(n, \alpha, \Delta p) \cdot \eta_{eDrv}(i_{Ph}, u_{DC}, n)} \right) \quad n \in R, n_{min} \leq n \leq n_{max} \tag{7}$$

3. APPLICATION AND VALIDATION OF THE MODEL-BASED DESIGN TOOL

3.1. Dimensioning and selection of an optimal solution based on a concrete example

In order to be able to select an optimal combination from a variety of inverter, motor and pump sizes that are suitable for a specific application, it is necessary to know the typical load profile for the application, like you can see in the following **Figure 5**:

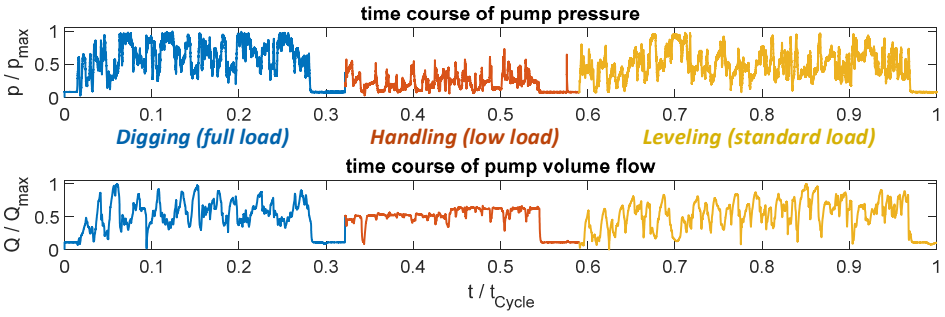


Figure 5: Composition of a load profile using the example of a mini excavator

This load profile represents the temporal pump pressure and the temporal pump volume flow of a combination of typical work tasks for an exemplarily selected mini excavator application. The values are normalized to the occurring maximum values of the cycle course. In the specific example, this would be the digging cycle as a representative of a full load task, the handling cycle, which stands for a low load, as well as the levelling cycle as standard or rather partial load task. The use of a mix of typical work tasks is particularly advisable for mobile applications, which have a large degree of

generalization regarding a range of tasks and can usually be used in their everyday work for many different activities. While the full load cycle defines reasonable performance requirements for the components, the combination of different cycles is necessary for a more realistic energy consumption evaluation. This is especially true for a design that focuses on maximizing energy efficiency.
A mathematical model of the electrified pump drive system shown schematically in **Figure 2** was developed in which the essential stationary relationships, mainly the energetic behaviour in the form of efficiency maps, but also the important dynamic effects, such as swivel angle and speed build-up dynamics, have been implemented. In addition, the speed strategies presented in Section 2.2 are implemented within this model. With the help of the provided load cycles, which act as inputs for the model, the behaviour of the electrified pump drive system can be simulated. The developed framework prepares the data of the corresponding components and generates a graphical representation of all simulated possible and permissible drive combinations. The result of such a simulation run for the exemplary selected application mini excavator and the compilation of a representative load cycle shown in **Figure 5** can be seen in the following **Figure 6**:

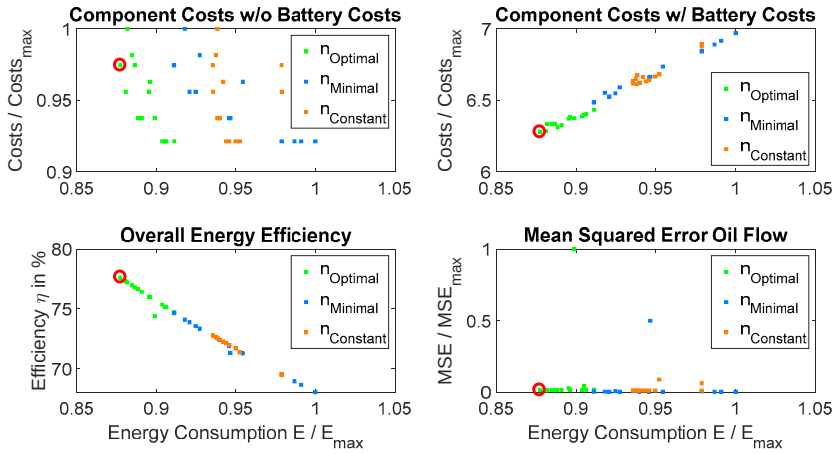


Figure 6: Presentation of the results of a simulation run of different combinations of the drive components

The results of the individual simulation runs are shown as corresponding points. The representation is again normalized here. The x-axes of all diagrams show the simulated energy consumption at the electrical input of the inverter. Further possible evaluation criteria for the design of a drive system, such as component costs with or without consideration of the battery (assumption: 8 h working time without charging), are selected as examples. The normalization of costs is carried out to the respective component cost maximum. It turns out that the selection of more expensive components of the actual powertrain may be worthwhile if a significant proportion of the total costs caused by a larger battery can possibly be saved (see drive combination marked in red in **Figure 6**).
Further statements can be made about the simulated operating strategies. The most obvious finding from the analysis of the results is that the optimal speed strategy achieves in some cases significantly higher energy savings compared with a minimum speed strategy. The contrast is even higher when looking at a constant speed strategy. Constant speed refers here to the speed at which the maximum volume flow of the cycle can be achieved. The representation of the mean squared error shows a measure of the deviation between the simulated volume flow and the predetermined volume flow Q_{des}. In this case, there are some outliers caused by effective torques that are too high. Depending on the design of the system, the pump controller can limit the torques that are too high by swivelling

back the pump, so that the drive does not stall. However, this has no effect on a possible compensation of the missing volume flow by increasing the speed. This speed adjustment can be considered, for example, when implementing the optimal speed strategy by adding a corresponding constraint during optimization.

3.2. Validation of the simulation results by measurements on a test bench

In order to validate the model-based methods, models and operating strategies under realistic operating conditions, the electrified pump drive system was set up on a test bench. The test bench setup can be taken from the following **Figure 7**:

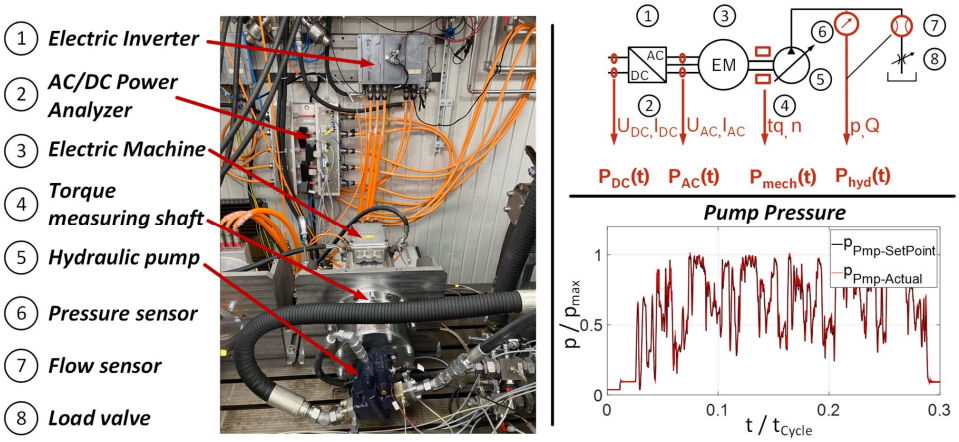


Figure 7: Test bench setup and test results

Bosch Rexroth components were used for the drive train. Specifically, this comprised an eLION EDS1 inverter, an eLION EMS1 permanent magnet synchronous machine and an A10VO axial piston pump with variable displacement and eOC Control. The software with the different operating strategies and the pump control were implemented on a BODAS RC5-6/40 control unit. The control unit received the required volume flow Q_{des}, which was set by the pump controller. The pump pressure was controlled by a pilot-operated proportional valve (8 in **Figure 7**). The trajectory of the load pressure control can be taken from the time course of the pressure (also in **Figure 7**). The following **Table 1** shows in a comparison the efficiency results obtained by simulations and measurements compared to a constant speed operation.

Table 1: Comparison of the efficiency gain relative to constant speed using the mini excavator cycles

Load cycle	**combined**		**digging**		**handling**		**levelling**	
Speed strategy	n_{Min}	n_{Opt}	n_{Min}	n_{Opt}	n_{Min}	n_{Opt}	n_{Min}	n_{Opt}
Simulation	1.6 %	3.3 %	-0.2 %	2.3 %	8.3 %	7.8 %	1.2 %	2.9 %
Testbench	1.7 %	3.1 %	-0.1 %	2.0 %	9.3 %	8.8 %	1.1 %	2.5 %

The results of the efficiency gain in percentage points show a very good agreement of the efficiency behaviour between the simulated and the measured data. As with the derivations to the operating strategies (section 5), the energy efficiency results from the division of the measured hydraulic power by the measured electrical AC power at the input of the inverter. In order to obtain a more differentiated statement, both the combined cycle and the three individual cycles from **Figure 5** were simulated and measured. Variable-speed drives show their advantages especially in cycles with low

loads, but also in cycles with high loads, significant efficiency advantages could be achieved.

3.3. Impact of speed strategies on thermal behaviour

In the following consideration, the influence of the selection of the speed strategy on the thermal behaviour of the drive system is examined in more detail. For this purpose, **Figure 8** shows characteristic maps with the operation points of the load cycles of the electrical machine and the inverter. The bottom plots show the thermal behaviour of the inverter over time. To achieve a thermally balanced state, the single digging cycle was simulated several times.

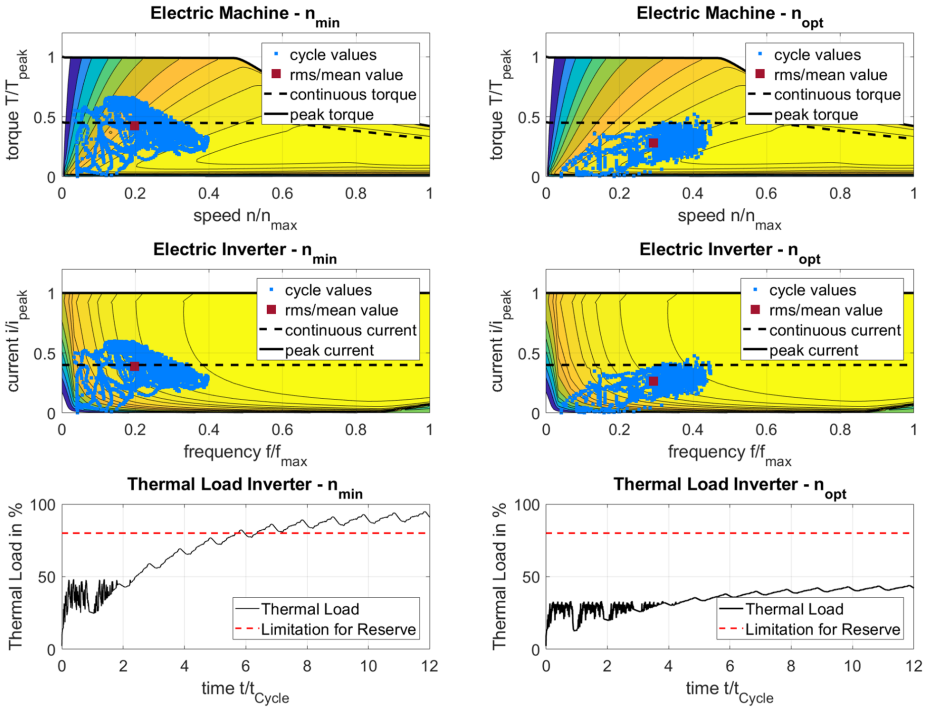


Figure 8: Effects of the speed strategies on the thermal behaviour of the drive components

The upper two figures show the stationary torque over speed behaviour of the electric machine in normalized representation. The behaviour at minimum speed is shown in the left illustration and at optimum speed in the right illustration. The point clouds reflect the resulted speed and torque operation points during the simulation, which are sampled with 10 ms. Since the thermal design should reasonably be carried out using the maximum expected loads, the consideration is based only on the digging cycle shown in **Figure 5**.

Particularly noticeable is the distribution of a significant number of working points above the continuous torque curve when using the minimal speed strategy compared to the optimal speed strategy. This circumstance results from the high swivel angles and results in an increased residence area with smaller efficiencies of the electric drive, which both leads to an increased RMS value of the torques. This is only just below the continuous torque curve, which could speak for an increased thermal utilization of the electrical machine. The same behaviour you can find in the representation of the current-frequency behaviour of the inverter, in which the torques were converted into effective values of the phase current and the rotational speed into the electrical rotational frequency of the

phase current. Again, by using n_{Min} strategy the RMS of the effective phase current values is significantly closer to the continuous curve of the phase current than when using n_{opt} strategy.
Thus, in the simulation of the thermal behaviour of the inverter, there are now very big differences. Over time the thermal load reaches almost 100% when using n_{min} strategy. At this value, the current would be limited in the real operation of the inverter, which can lead to either a stall or at least to a reduction in the volume flow conveyed. For this purpose, however, the pump controller must be designed to reduce the swivel angle in such a way that the pump torque is not higher than the reduced drive torque during a thermal overload situation of the electric drive. In both cases, however, a reduced performance would appear. Such a situation does not occur with speed strategies which utilize the degree of freedom like the optimal speed strategy where the best swivel angle and speed combination is determined based on current circumstances. The thermal behaviour when using n_{opt} strategy is not critical and the thermal utilization is well below a reserve rate of 80%, which should be aimed for design purposes at most. An advantage of an uncritical thermal behaviour when using an optimum speed strategy can be seen in the fact that a greater overload capacity of the electric drive system is maintained and, if necessary, the drive system can be optimized to use smaller dimensioned components.

4. CONCLUSIONS AND OUTLOOK

In this article, a procedure for the design and energy-efficient operation of an electrified pump drive for supplying the hydraulic work equipment of a mobile working machine with a demand-oriented volume flow is presented. Variable speed strategies play an essential role here and aim either to maximize the swivel angle alone or to minimize the loss of the entire electrified pump drivetrain. With the help of the operating strategies and a model-based approach, many different combinations of component sizes can be simulated and evaluated in their behaviour in an efficient way. This represents a very powerful tool for the users to enable a principally efficient operation of their mobile machine in combination with the use of the speed control strategies. During operation, depending on the application and the work task performed, an average of 3 to a maximum of 9 percentage points of electrical energy can be saved compared to operation at constant speed. A further advantage is obtained, when using the optimum speed strategy – also regarding the thermal behaviour of the electrical components, which is uncritical and assures the full performance of the mobile application during its operation. In our current development, we continue to work on speed variable solutions that offer an user-friendly implementation of variable-speed pump drives including pump control, inverter device and communication as well as power limitation control.

NOMENCLATURE

α	swivel angle of the variable pump	%
$cos\ \varphi$	active power factor	
EFM	Electronic Flow Matching System	
ePC	Electronic Positive Control System	
f	frequency, frequency of the three-phase rotation field of the electric machine	Hz
i_{Ph}	phase current	A
LS	Load Sensing System	
MCV	main control valve	
MSE	mean squared error (between demanded and simulated oil flow)	$(m^3/s)^2$
n	speed of pump / electric machine	rpm
n_{const}	constant speed strategy (referenced speed for maximum desired oil flow)	
n_{min}	minimal speed strategy	

n_{opt}	optimal speed strategy	
η_{eDrv}	overall efficiency electric drive	%
η_{Em}	efficiency electric machine	%
η_{Inv}	efficiency inverter	%
η_{mech}	mechanical efficiency of the pump	%
η_{Pmp}	overall efficiency of the pump	%
η_{V}	volumetric efficiency of the pump	%
p, Δp	pressure, pressure difference	Pa, bar
P_{EmIn}	input power of electric machine	W
P_{EmOut}	output power of electric machine	W
P_{ElecIn}	electrical input power at inlet of inverter	W
P_{hyd}	hydraulic power at output of electric pump drive	W
P_{InvOut}	output power of inverter	W
P_{mech}	mechanical Power at pump shaft	W
RMS	root mean square	
t_{Cycle}	cycle time	s
T_{eff}	effective torque at inlet shaft of pump	Nm
T_{ideal}	calculated ideal torque of pump based on displacement and pressure	Nm
u_{DC}	intermediate circuit voltage	V
Q_{Des}	desired / demanded oil flow	m³/s
Q_{eff}	effective oil flow at output of pump	m³/s
V_g	displacement of pump	m³
ω	circular frequency	

REFERENCES

[1] Huang H, et al (2018) *Improving the energy efficiency of a hydraulic press via variable-speed variable-displacement pump unit.* Journal of Dynamic Systems Measurement and Control 140 (11), May 2018

[2] Mayer A (2022) *Ansatz zur Bestimmung zukünftiger Einsatzgebiete von batterie-elektrischen Rad- und Teleskopladern,* VDI-Berichte 2402: Powertrain Systems in Mobile Machines 2022

[3] Mutschler S, Glas J (2023) *Diesel-elektrische Antriebstopologien für mobile Maschinen.* 9. Fachtagung Hybride und Energieeffiziente Antriebe für Mobile Arbeitsmaschinen, Karlsruhe, February 27-28, 2023, Karlsruhe, Germany

[4] Reichert W, et al (2022*) Implementation on an electro-hydraulic drive unit with two control variables in a drawing cushion application*. 28th Saxon Conference on Forming Technology SFU and the 7th International Conference on Accuracy in Forming Technology ICAFT, November 2022, Chemnitz, Germany

[5] Reidl T; et al (2021) *Loss optimal control strategy of speed and displacement variable electrohydraulic axes.* ICFP 2021, April 2021, Hangzhou, China

[6] Rui J, et al (2019) *Energy saving strategy of the variable-speed variable-displacement pump unit based on neural network*. 26th CIRP Life Cycle Engineering (LCE) Conference, May 2019, Purdue University, West Lafayette, USA

[7] Serrao L (2022) *Driving the electric transformation – The evolution of off-highway powertrain*. VDI-Berichte 2402: Powertrain Systems in Mobile Machines 2022

[8] Willkomm J (2016) *Modellprädiktive Optimierung drehzahlvariabler Verstellpumpen*. Dissertation, TU Dresden, Dresden, Germany

[9] Zheng Y, et al (2022) *Energy-efficient electro-hydraulic power source driven by variable-speed motor.* Energies 2022, 15

FUNDAMENTALS OF HYDRAULIC TRANSFORMERS

Peter Achten

INNAS BV, Nikkelstraat 15, 4823 AE Breda, the Netherlands

E-mail address: pachten@innas.com

ABSTRACT

In Common Pressure Rail systems (CPR-systems), it is no longer possible to control the loads at the primary side by means of a variable displacement pump. Instead, the loads need to be controlled directly at the load (also referred to as secondary control). Rotating loads could be controlled by variable displacement motors, but hydraulic cylinders need to be controlled by means of hydraulic transformers. The problem is, however, that these transformers are not yet available on the market.

This paper discusses and analyses the principal design options for hydraulic transformers, thereby comparing several combined hydraulic transformers (CHTs), which are combinations of two pump/motors, and the Innas Hydraulic Transformer (IHT). The emphasis is on hydraulic power transformers which can be used as a general hydraulic control component, having a continuously variable control.

Keywords: Hydraulic transformer, CPR-system

1. INTRODUCTION

Most current hydraulic systems are inefficient. There are several reasons for this inefficiency. Although pumps and motors may have a high peak efficiency, the average efficiency is much lower, especially in variable displacement machines. Hydraulic systems could also be much more efficient if energy would be recuperated. Hydro-pneumatic accumulators are excellent storage devices for capturing the recuperative energy, but in almost all hydraulic applications, energy is not recuperated. Another reason for the inefficiency of hydraulic systems are the valve losses. In most hydraulic applications, a variable displacement pump is supplying oil to several loads at different pressure levels. The resulting pressure differences are throttled away and dissipated in the valves.

It is known for many years, that hydraulic transformers could eliminate many of these losses, and would strongly increase the efficiency of hydraulic machinery. First of all, they can eliminate most valve losses. Furthermore, they offer the opportunity to recuperate energy by constantly converting the pressure in the accumulator to the pressure needed at the load. The transformers will be connected to a Common Pressure Rail or CPR. This rail separates the energy supply side from the loads. It offers the opportunity for advanced power management, thereby increasing the efficiency of the pumps (which can now have a fixed displacement) and the prime mover. Previous studies have shown that these CPR-systems can reduce the energy consumption by more than 30% [1-6].

The main reason, why hydraulic transformers are generally not used, is that there are hardly any hydraulic transformers that have reached the market [7-11]. The market is still waiting for a winning design. This paper investigates the fundamentals of hydraulic transformers, to understand what is needed to make a feasible hydraulic transformer for a market introduction. The analysis will be focused on transformers which can replace conventional valve control systems for most hydraulic applications. Such a transformer is a dynamic pressure and speed controller at a power level of 10

kW to several 100 kW, output flows in the order of 100 to 400 □/min, and pressure levels up to 420 bar.

2. BASIS EQUATIONS AND DESIGN GUIDELINES

Hydraulic transformers convert hydraulic power without any principal losses: they convert a flow Q_1 at a certain pressure difference Δp_1 into another flow Q_2 at another pressure difference Δp_2, thereby keeping the product of flow Q and pressure difference Δp (i.e. the hydraulic power) constant:

$$(p_1 - p_0)Q_1 = -(p_2 - p_0)Q_2 \tag{1}$$

In this paper, oil flowing into the transformer is defined as positive. If the pressure at the low-pressure side is defined to be zero, then Eq. (1) can be written as:

$$p_1 Q_1 = -p_2 Q_2 \tag{2}$$

The equation shows that a difference in pressure level also results in a difference in flow level. A flow of for example 100 □/min at 300 bar can be transformed to a flow of 300 □/min at 100 bar. This differs from valve operation, where the flow through the valve remains constant and energy is dissipated. In a hydraulic transformer, the flow at the output Q_2 will in general be different from the flow at the input Q_1. Only if $p_1 = p_2$, then $Q_1 = -Q_2$.

In addition to the energy continuity equation (2) there is also a flow continuity equation. Aside from Q_1 and Q_2 , the continuity equation demands that there needs to be a third flow Q_0 which is coming from the low-pressure side of the hydraulic circuitry. If we assume the density to be constant, then:

$$Q_0 + Q_1 + Q_2 = 0 \tag{3}$$

Considering this third make-up flow Q_0, the energy continuity equation can be rewritten as:

$$\sum_{i=0}^{2} p_i Q_i = p_0 Q_0 + p_1 Q_1 + p_2 Q_2 \tag{4}$$

Combining Eq(4) and (3) results in Eq.(1), whereas assuming $p_0 = 0$ results in Eq(2).

One of the earliest [12] and most simple designs of a hydraulic transformer is a multi-chamber cylinder (Fig. 1). These transformers are often called pressure amplifiers [13]. The amplifier has three chambers at three different pressure levels. If the middle ring chamber has a pressure $p_0 = 0$, then:

$$p_1 A_1 v = -p_2 A_2 v \tag{5}$$

By means of valves, the piston in the amplifier can move back and forth, for instance by means of ports in a rotating device [14, 15]. It is also possible to make the cylinder symmetrical [16], having two ring chambers. This creates the possibility for a somewhat continuous operation.

Pressure amplifiers have a few disadvantages. Due to the limited switching frequency, the power density is poor. Moreover, the transformation ratio is fixed and defined by the area ratios of the piston. It is possible to create a multi-step cylinder [17-19], which allows multiple transformation steps, but this makes the design even larger.

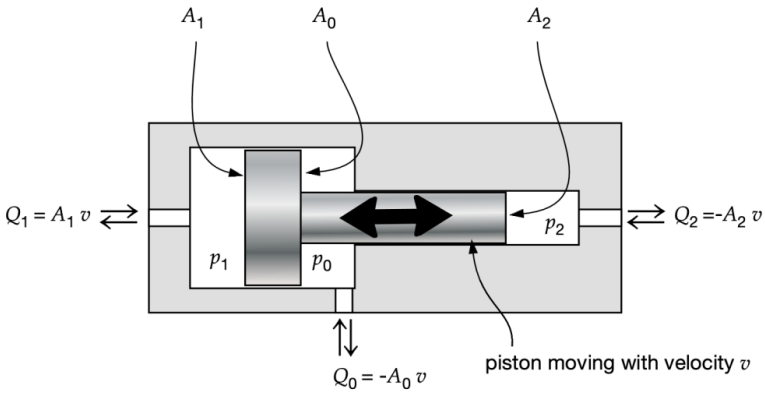


Fig. 1: Pressure amplifier

In this paper, these fixed transformers and pressure amplifiers are ignored. The focus is on hydraulic power transformers which can be used as a general hydraulic control component, i.e. as a replacement of current valve control:

1. The transformer needs at least three connections at three different pressure levels;
2. The transformer must be continuously variable;
3. The transformer should be able to regenerate hydraulic power, that is, to return hydraulic power from the load (at p_2 and Q_2) back to the high pressure rail at pressure p_1;
4. The transformer needs to be able to move oil from a lower to a higher pressure level, for instance when energy needs to be recuperated from a load pressure, which is lower that the pressure level in the high pressure accumulator. Q_0 also often needs to flow from the low pressure level p_0 to the higher pressure levels of p_1 or p_2. The flow direction of Q_0 depends on the transformation ratio p_2/p_1. If for instance $p_2 < p_1$ and Q_1 is flowing into the transformer, then Q_2 will be larger than Q_1. As a consequence, an additional flow Q_0 needs to be supplied to the transformer from the low pressure side. Also in this case, oil needs to go from a low to a higher pressure level.

In order to make oil flow from a lower to a higher pressure level, oil needs to be 'pushed' or displaced to this higher level. This requires some kind of pump or motor principle. Therefore, hydraulic transformers are hydrostatic machines, like hydrostatic pumps and motors. Taking into consideration the power, pressure, and flow demands, they need to be built around some kind of rotating displacement principle. Pressure amplifiers, such as the example in Fig. 1, do not fulfil these demands. They are not continuously variable, and their power and flow capacity is too small for most hydraulic applications.

Figure 2 shows the field of operation (the blue area) that the transformer must cover. The horizontal axis shows the load flow Q_2 relative to the maximum flow demand Q_{max}. The vertical axis shows the pressure or transformation ratio p_2/p_1. Hydraulic systems deliver their energy to either hydraulic motors or hydraulic cylinders. In most cases, these loads have stop-and-go operation, which means that the displacement principle used in the transformer needs to be operated frequently at these conditions as well. In the end, the transformer needs to be able to deliver any flow between $-Q_{max}$ and Q_{max}.

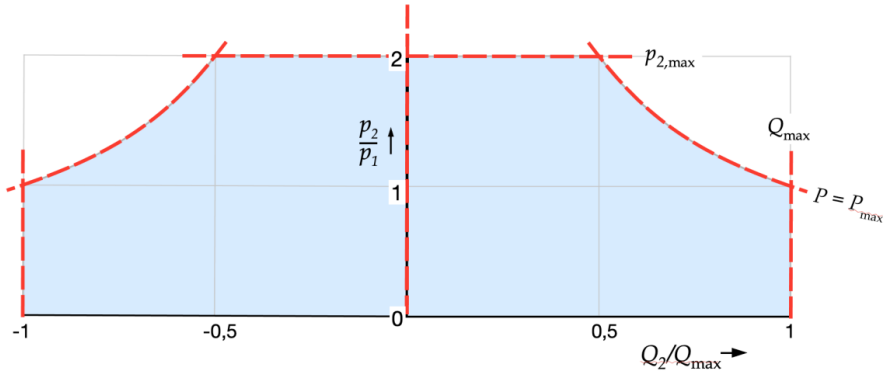


Fig. 2: Field of operation of the hydraulic transformer

The maximum pressure and power demands are also set by the load. Many, if not most hydraulically driven machines, are production machines, meaning that these are demanding applications which go to the limit of the maximum power and pressure capacity.

3. HOW TO MAKE THE TRANSFORMER VARIABLE?

There are various possibilities to make a transformer variable. One way is to connect the shafts of several fixed displacement pumps and motors, and decide by means of valves which of these pumps and motors will be active [20, 21]. Aside from becoming bulky, this design would not allow a continuously variable operation.

Another option would be to change the rotating speed. Equation (2) can be rewritten as

$$p_1 n_1 V_1 = p_2 n_2 V_2 \tag{6}$$

The minus sign on the righthand side of Eq(2) is now included in the definition of the displacement V_2. If the ratio between n_1 and n_2 could be varied, then the pressure ratio p_2/p_1 would be varied accordingly. But this would require a continuously variable mechanical transmission with the same power rating as the transformer (i.e. 10 kW to several 100 kW). Because of the size, weight, complexity and costs of such a transmission, this option will not be investigated further. In this paper the rotational speed will be considered the same for V_1 and V_2 ($n_1 = n_2$). Eq(6) is then reduced to:

$$p_1 V_1 = p_2 V_2 \tag{7}$$

The only remaining way to make the transformer variable is to change the displacement volumes:

$$p_1 \xi_1 V_1 = p_2 \xi_2 V_2 \tag{8}$$

The parameters ξ_1 and ξ_2 reflect the reduction of the maximum geometrical displacement ($|\xi| \in [0,1]$). The transformation ratio equals p_2/p_1:

$$\frac{p_2}{p_1} = \frac{\xi_1}{\xi_2}\frac{V_1}{V_2} \tag{9}$$

Two different concepts of such a transformer have been studied in the past:

1. A transformer based on a combination of two hydraulic machines (CHT), of which at least one has a variable displacement;
2. A transformer based on a single hydraulic machine with three instead of two pressure ports, as has been designed by Innas (IHT) [22-24].

Both concepts are evaluated in the next two sections of this paper.

4. COMBINATIONS OF TWO PUMP/MOTORS (CHT)

First designs of a hydraulic transformer based on two hydrostatic pump/motors showed up around 1970 [25, 26]. These designs were not variable. The first variable hydraulic transformers with two hydrostatic machines started to be designed around 1980 [27].

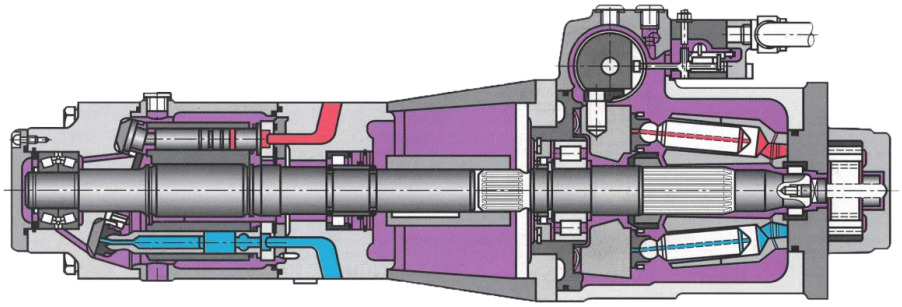


Fig. 3: Variable hydraulic transformer build as a combination of a variable displacement pump/motor and a fixed displacement pump/motor [28]

Most of these transformer designs are based on combinations of two axial piston machines, but there are also designs based on radial piston machines [29], external gear pumps [30], geroler motors [31] and vane pumps [27, 32].

These transformers can have one or two variable displacement machines. Examples of both machines being variable can be found in [30, 33]. There are various ways to connect the two machines to the p_0, p_1 and p_2-lines [29]. Figure 3 shows an example of a design in which a fixed and a variable pump/motor are combined.

A simple way, to understand a hydraulic transformer as a combination of two machines, is to first look at a transformer which is constructed as several pistons acting on a torque beam or seesaw. Figure 4 shows three configurations of a beam which is supported in the middle by a pivot, on which four pistons act. The arm lengths $\xi_1 R_1$ and $\xi_2 R_2$ of the pistons can be changed. The two pistons on the left side of the pivot point and the two pistons on the right side move in pairs. The pressure levels in Fig. 4 change between the concepts. The torque equations and the transformation ratios of each concept are shown below the drawings (assuming no friction and $p_0 = 0$).

The same idea is realized when combining two pump/motors (see figure 5). In theory, there are even more combinations possible, but those are not meaningful [34]. Figure 5 also shows the stationary torque equations, assuming no friction and other hydromechanical losses and assuming no accelerating or decelerating inertia effects.

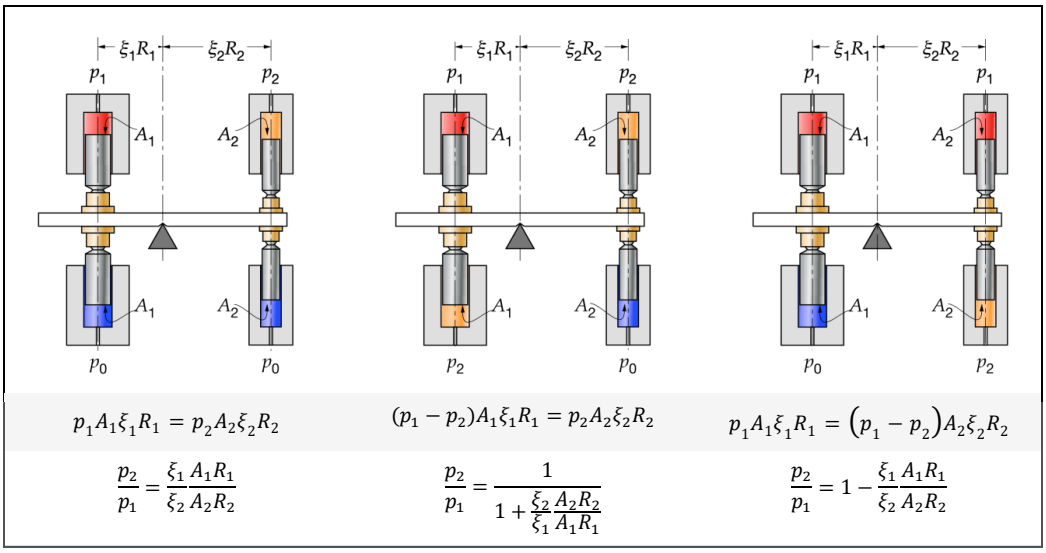


Fig. 4: Hydraulic transformers with 4 pistons achting on a torque beam. The torque equations and transformation ratios p_2/p_1 assuming no friction and $p_0 = 0$

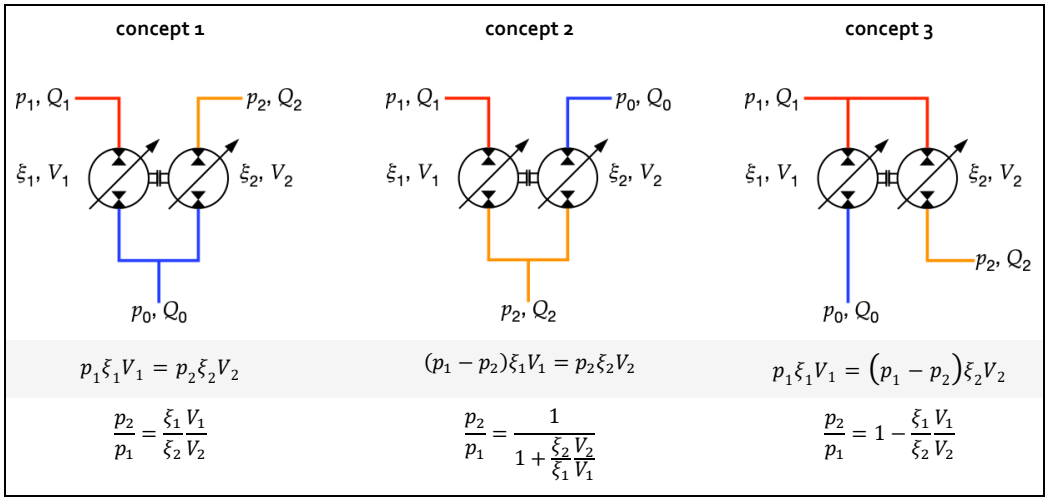


Fig. 5: Three different concepts to combine 2 hydrostatic machines. The torque equations and transformation ratios p_2/p_1 are shown below each concept, assuming no friction and no volumetric losses, and assuming $p_0 = 0$ bar

In principle, the displacement can be varied between $-1 \leq \xi \leq 1$. However, this does not make sense for concept 1, since $\xi_1/\xi_2 < 0$ would result in negative pressure ratios. All concepts in Fig. 5 can result in infinite pressure ratios (Table 1).

Table 1: Conditions when p_2/p_1 becomes infinite

concept 1:	$\xi_2 = 0$
concept 2:	$\frac{\xi_2 V_2}{\xi_1 V_1} = -1$
concept 3:	$\xi_2 = 0$

Concept 2 results in a strong non-linear relationship between the control parameters ξ_1 and ξ_2, and the transformation ratio p_2/p_1. Concept 2 also has a limited transformation range and is unable to cover the entire field of operation as shown in Fig. 2. Both concepts 1 and 3 can operate in the entire field of operation as described in Figure 2, but concept 3 would need a much more complicated over-centre design ($-1 \leq \xi_1 \leq 1$) then concept 1, which only needs ξ_1 to be varied between 0 and 1.

Because of the weaknesses and disadvantages of concepts 2 and 3, only concept 1 with V_1 being variable and V_2 being a fixed displacement machine ($\xi_2 = 1$) will be evaluated in this paper. Concepts 2 and 3 are sometimes considered as modes of concept 1 [35-37]. By means of valves, it should then become possible to choose which mode to choose, which would give the opportunity to improve the efficiency. This paper will not further pursue these options.

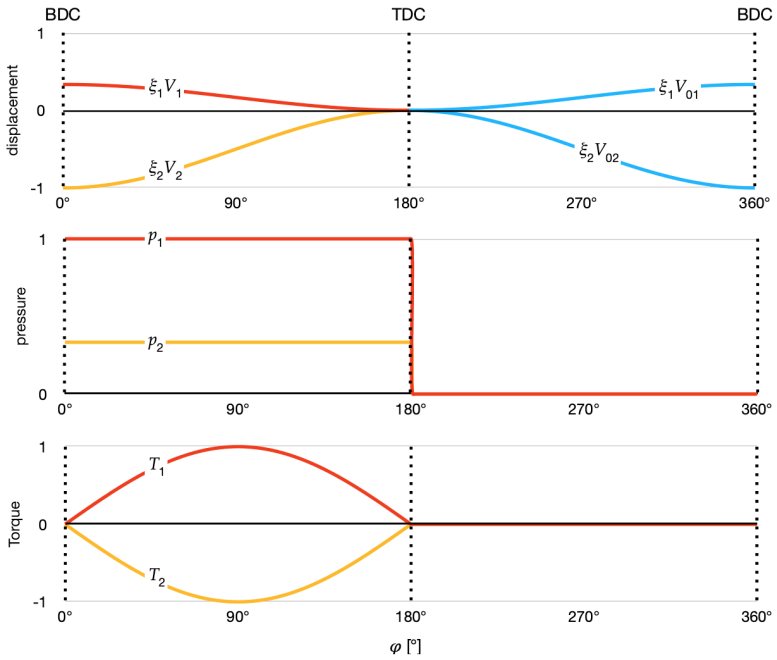


Fig. 6: Relative displacement, pressure level and torque generate of a single piston during one rotation. The diagram shows a transformer as a combination of two machines, in which machine 1 has a displacement which is 1/3 of the displacement of machine 2, and in which the pressure level during the first half of the rotation is three times as high in machine 1 as in machine 2

For transformers, which are built as a combination of two pump/motors, the principle can be explained further by plotting the volumetric displacement, the pressure level and the torque of a single displacement chamber during a single rotation (See Fig. 6). During each rotation of the shaft, each displacement chamber is connected sequentially to different ports at different pressure levels. For instance, in an axial piston pump, the cylinders are connected to the low-pressure port for about half of the rotation, and to the high-pressure port during the other half. In the example of Fig. 6, the two units have a difference of a factor 3 in displacement: $\xi_1 V_1$ of machine 1 is three times smaller than $\xi_2 V_2$ of machine 2. On the other hand, the pressure level during the first half of the rotation is three times as high for machine 1 as for machine 2. This results in a product pV i.e. a torque, which is equal in amplitude, but different in sign. The make-up flow is shown in Fig. 6 as two individual displacements $\xi_1 V_{01}$ and V_{02}. Since it is assumed that $p_0 = 0$, there is no torque generated or demanded during this second half of the rotation.

The curves which are shown in Fig. 6 are simplifications of the reality. There are no losses and the commutation in the top and bottom dead centres (TDC and BDC) are assumed to be instantaneous. In addition, in Fig. 6, the curves of machines 1 and 2 are exactly in phase. In reality, the displacement in one machine will most of the time not be synchronized with the displacement in the other. But having multiple displacement volumes (i.e. pistons and cylinders in an axial piston machine), will not change the principle of operation.

5. THE INNAS HYDRAULIC TRANSFORMER (IHT)

The conventional method of changing the displacement of piston pumps and motors is to change the amplitude of the sinusoidal piston movement, for instance by reducing the angular position of the swash plate in an inline axial piston pump (see Fig. 3). This is also the way in which, in the previous section, the transformers are controlled.

There is however another method to change the displacement, without changing the swash plate angle. In conventional pumps and motors, the filling and the delivery stroke are defined between the top and bottom dead centers (TDC and BDC). As an example, in Fig. 6, both machines have a high pressure stroke between 0 and 180°, and a low pressure stroke between 180 and 360°. But these 'windows' can be changed to different angles. Part of the stroke will then occur while passing one of the dead centres, and the effective displacement will be reduced.

Essential is, that in current pumps and motors, the port plate has a fixed position in the housing. But it is also possible to create a pump (or motor) in which the port plate can rotate around its own axis. The commutation can then occur outside the TDC and BDC areas, when the pistons are already moving with a considerable speed. As a result, the high and low pressure port of the port plate will rotate past the TDC and BDC. Both the effective suction stroke and the effective delivery stroke will become shorter, and the effective volumetric displacement will be reduced. This will effectively create a variable displacement pump or motor [38-42]. In case of a digital displacement pump or motor, the valve timing can offer the same control option [43, 44].

The Innas Hydraulic Transformer (IHT) uses the same principle. In addition, the design features a port plate with three ports instead of two (see Fig. 7). The ports are connected to the low-pressure line of the CPR-system at a pressure p_0 (port 0), the high pressure line of the CPR-system at a pressure p_1 (port 1) and the load line at a pressure p_2 (port 2).

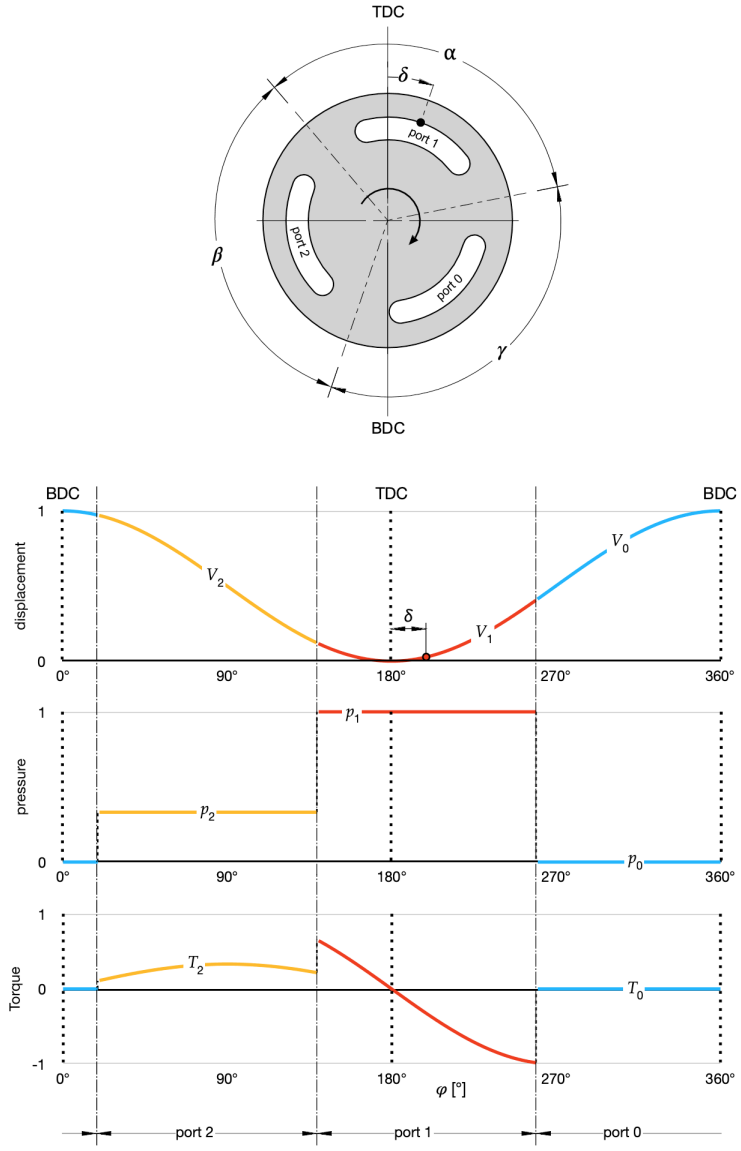


Fig. 7: The IHT which is controlled by the angular position of the port plate, and in which the port plate has three ports. The curves show the relative volumetric displacement, the relative pressure level in each of the ports and the torque generation during one revolution.

This is illustrated in Fig. 7, which shows a design of the IHT for an axial piston design. In this example, each port has the same length. Between the ports are the sealing lands which are necessary to avoid short-circuit leakage during commutation. For the port plate which is shown in Fig. 7, a cylinder block with 9 pistons is assumed, resulting in a commutation zone of about 360°/9 = 40°. Considering the movement and the size of the ports in the cylinder block, each port in the port plate has an effective length of 120°. Other configurations and designs, with different port lengths, are also

possible. For marking the rotational position of the port plate, a point in the middle of port 1 has been chosen as a reference. The angular distance between this point and the TDC-BDC-line is the control angle δ. In this drawing, the cylinder block, which rotates on top of this port plate, is assumed to rotate in clockwise direction.

The transformation ratio p_2/p_1 is not anymore dependent on the geometrical displacement V but on the effective arc lengths α, β and γ of the ports and the control angle δ:

$$\frac{p_2}{p_1} = \frac{sin\left(\frac{\alpha}{2}\right)sin(\delta) - \frac{p_0}{p_1}sin\left(\frac{\gamma}{2}\right)sin\left(\delta - \frac{\beta}{2}\right)}{sin\left(\frac{\beta}{2}\right)sin\left(\delta + \frac{\gamma}{2}\right)} \tag{10}$$

When $p_0 = 0$:

$$\frac{p_2}{p_1} = \frac{sin\left(\frac{\alpha}{2}\right)}{sin\left(\frac{\beta}{2}\right)} \frac{sin(\delta)}{sin\left(\delta + \frac{\gamma}{2}\right)} \tag{11}$$

The diagram of Fig. 8 shows the calculated relationship between p_2/p_1 and δ for two different port plate designs and $p_0 = 0$.

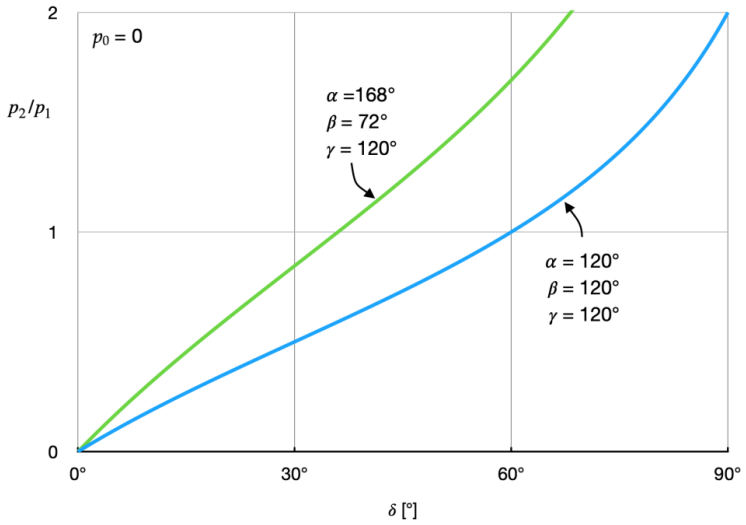


Fig. 8: Transformation ratio p_2/p_1 for the Innas Hydraulic Transformer (IHT) for two different configurations as a function of the control angle δ.

The IHT can control the complete range of transformation ratios from 0 to an amplification ratio of $p_2/p_1 = 2$, and even higher if needed. To achieve this, the port plate needs to be rotated to a maximum angle of 68° to 90°, depending on the port plate design.

Because there is no ISO-symbol for this type of hydraulic transformer, INNAS has designed and defined this symbol for the IHT:

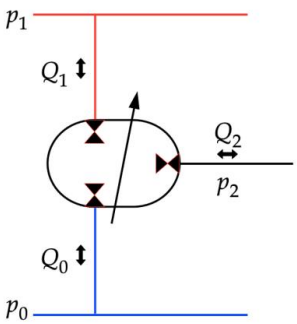


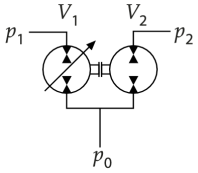
Fig. 9: Symbolic representation of the IHT

6. TRANSFORMER SIZES

The analysis of the previous section resulted in two principle designs of a hydraulic transformer that fulfil the following basic requirements:

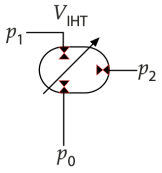
- being continuously variable;
- can be operated in the entire field of operation, as defined in Fig. 2;
- can recuperate energy

The two principles are shown below:



$$\frac{p_2}{p_1} = \xi_1 \frac{V_1}{V_2}, 0 \leq \xi_1 \leq 1$$

a. Combination of two pump/motors (CHT)



$$\frac{p_2}{p_1} = \frac{sin\left(\frac{\alpha}{2}\right)}{sin\left(\frac{\beta}{2}\right)} \frac{sin(\delta)}{sin\left(\delta + \frac{\gamma}{2}\right)}$$

b. Innas Hydraulic Transformer (IHT)

Fig. 10: Hydraulic transformer designs. The transformation ratio p_2/p_1 is valid for $p_0 = 0$

Both transformers need to achieve a transformation ratio $p_2/p_1 = 2$. For the CHT, this means that $V_1 = 2V_2$ to get to this maximum pressure ratio at a maximum displacement of V_1 (i.e. $\xi_1 = 1$). For the Innas Hydraulic Transformer (IHT), the transformation ratio is not dependent on the maximum volumetric displacement, but on the effective arc lengths of the three ports. Assuming all ports are equal ($\alpha = \beta = \gamma = 120°$) the control angle δ needs to be 90° to make $p_2/p_1 = 2$.

The flow output of the transformer based on two pump/motors is:

$$Q_2 = \frac{nV_2}{1000} \tag{12}$$

For the Innas Hydraulic Transformer the flow output equals:

$$Q_2 = \frac{nV_{IHT}}{1000 \cdot 2}\left[cos\left(\delta - \frac{\alpha}{2} - \beta\right) - cos\left(\delta - \frac{\alpha}{2}\right)\right] \tag{13}$$

In these equations the rotational speed n is defined in rpm and the flow Q in □/min. The volumetric displacement V is defined in cc/rev.

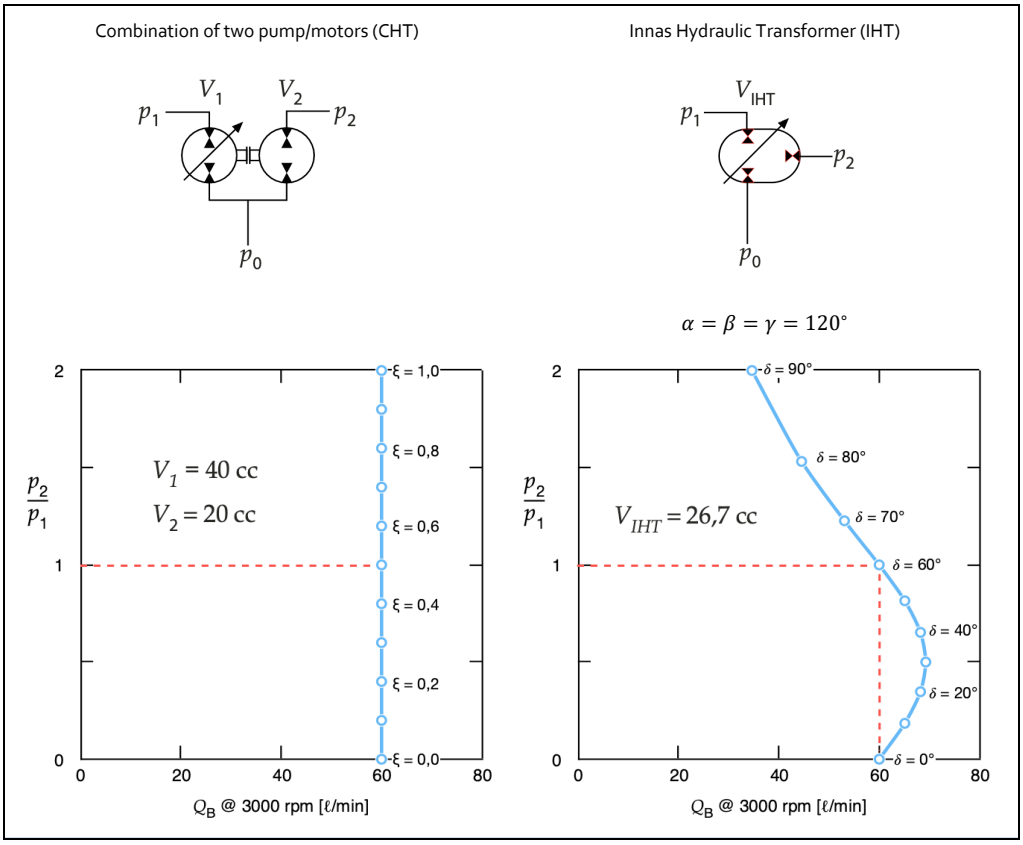


Fig. 11: Calculated flow Q_2 and transformation ratio p_2/p_1 for the two considered transformers ($p_0 = 0$). The geometrical displacements V_1, V_2 and V_{IHT} are calculated under the conditions that the transformers:

- can achieve a transformation ratio $p_2/p_1 = 2$
- can create a flow $Q_2 = 60$ □/min at $n = 3000$ rpm and $p_2/p_1 = 1$

To compare the two designs a flow capacity of 60 □/min is chosen as an example, to be reached at a rotational speed of 3000 rpm and a pressure ratio $p_2/p_1 = 1$. As before, the ports in the Innas Hydraulic Transformer are assumed to be equal ($\alpha = \beta = \gamma = 120°$). To get to the demanded flow of 60 □/min, the CHT needs a variable displacement pump/motor with a maximum displacement of 40 cc and a fixed displacement pump/motor with a displacement of 20 cc per revolution. In contrast the IHT needs a rotating group with a total geometrical displacement of only 26,7 cc per revolution. The

diagrams of Fig. 11 show the calculated relationship between the transformation ratio p_2/p_1 and the output flow Q_2.

The comparison clearly shows that, in principle, the Innas Hydraulic Transformer only needs a single rotating group which has a geometrical displacement of 26,7 cc/rev, whereas the combination of a variable displacement pump/motor and a fixed displacement pump/motor needs a total displacement of 60 cc/rev. In contrast, the IHT has a smaller flow output at the highest transformation ratio. This corresponds to the pressure and flow demands of most hydraulic applications, where the flow demand is reduced at higher load pressures. The relationship between flow and pressure demand more or less follows the hyperbolic constant power line.

If the displacement of the IHT would be increased to 46,2 cc/rev then the flow output at the maximum transformation ratio of the IHT would be equal than of the CHT. But then the maximum flow output of the IHT would go up to 120 □/min at $\delta = 30°$, twice as high as for the CHT.

Also, the IHT can result in very high transformation ratios:

$$p_2/p_1 \to \infty \text{ when } \delta \to -\frac{\gamma}{2} + k\pi \tag{14}$$

Considering that $p_2/p_1 \in [0,2]$ the control angle will always be smaller than the critical value.

7. OTHER DEMANDS

Next to the four design requirements that were discussed in section 2, there are several other more practical demands for the application of a hydraulic transformer: There are several demands for the application of hydraulic transformers:

- The hydrostatic principle needs to be variable
- It must be possible to run the transformer at close to zero speed and breakaway, even at high pressure levels and loads, without excessive wear and torque losses
- Volumetric losses need to be as small as possible, especially at low rotational speeds
- It should preferably also be possible to run the transformer at high rotational speeds
- The noise, vibration and harshness (NVH) levels need to be low
- The transformer must be capable of handling loads up to 420 bar continuously
- The overall efficiency needs to be high in a wide range of operating conditions. This includes possible losses of the displacement control

The first demand is that the displacement can be varied. This excludes all gear pumps. There is some research making gear pumps variable [45-47], but these designs are rather immature. Aside from a limited pressure, flow and power range, the most important disadvantages are the limited range in which the displacement can be varied. Other principles, which cannot be varied, are the geroler and gerotor principle, helical pumps, and multi-lobe radial piston machines.

This study focusses on hydraulic transformers as a general substitute for current valve control based hydraulic systems, i.e. to control hydraulic cylinders and possibly also hydraulic motors that have a frequent stop-and-go operation. Consequently, hydraulic transformers will also require having frequent stops and breakaways, often even at high pressure loads. Most hydrostatic principles rely on hydrodynamic lubrication and cannot run for an extended period at low rotational speeds.

Another concern is the volumetric loss and internal leakage at low rotational speeds. These losses result in a dead band for pumps: an operating range in which the pump does not deliver any net effective flow because the volumetric losses are higher than the pump displacement. In a motor, the volumetric losses result in slippage. Volumetric losses cannot be eliminated completely, but in

hydraulic transformers, the volumetric losses need to be as small as possible.

From a size perspective, it will be impossible to be as small and compact as a valve segment of a load-sensing block. Nevertheless, the transformers should preferably be small. Aside from the size constraints mentioned in the previous section, this implies that the maximum rotational speed of the transformer needs to be as high as possible. Doubling of the rotational speed not only halves the volume of the rotating parts, but also of the size of the housing and the actuators which are needed to change the displacement. Hydrostatic principles have different speed constraints. Some are due to cavitation, but also centrifugal forces and barrel tipping [48-50] in axial piston machines can limit the maximum rotational speed.

Another requirement of hydraulic transformers is noise, vibration and harshness (NVH). Most certainly in battery driven applications, but also in CPR-systems with accumulators, the noise level of the hydraulic system is of much bigger importance than in conventional drives, in which the diesel motor dominates the noise level. The noise level is generally dependent on the rotational speed, the pressure level, and the number of displacement chambers (i.e. pistons) of the rotating group.

Obviously, the hydrostatic principle for a hydraulic transformer needs to fulfil the pressure demands of current hydraulic applications. Pressure levels of 350 to 420 bar are currently the peak pressure demands. Not all hydrostatic principles can handle these pressure levels.

Finally, it is important that the hydrostatic principle has a high overall efficiency in a wide range of operating points and conditions. For instance, if, in a CHT, both machines have an efficiency of 90%, then the total efficiency of the transformer will only be 81%. If this same transformer also recuperates energy with the same efficiency, then the overall cycle efficiency will be limited to 65,6% ($0{,}9^4 \times 100\%$). Many pumps and motors don not have an overall efficiency of 90%, especially variable displacement pumps. For hydraulic transformers, the highest possible efficiency is needed.

For these reasons, INNAS has developed the new floating cup principle. It is a multi-piston principle which strongly reduces the torque ripple and improves the controllability of the transformer. Aside from having a high efficiency, it is also possible to operate floating cup machines starting from standstill and at low rotational speeds, even when the pressure is high (which it always is at the high pressure CPR-port).

8. CONCLUSION

This article derives the most important, fundamental equations for hydraulic transformers: the equations for the transformation ratio between output and input pressure, and the flow equations of the transformer. It has been shown that the principle of the Innas Hydraulic Transformer (IHT) is much more compact than a combination (CHT) of a variable displacement pump/motor and a fixed displacement pump/motor.

The demands of hydraulic transformers are different from conventional pumps and motors. The efficiency needs to be higher (also at low rotational speeds, in combination with high pressure levels), the dynamic response needs to be higher, and the noise level needs to be lower. The floating cup principle is specifically designed for these demands and is a better fit than other hydrostatic pump and motor principles. In addition. That being said, both the floating cup principle and the IHT-concept might not be the only solutions, and better designs might be found in the future. This paper gives an overview of the guidelines and demands for such new designs.

NOMENCLATURE

A	Area	m^2
n	rotational speed	rpm
p	pressure	bar
p_0	pressure level low pressure CPR-line	bar
p_1	pressure level high pressure CPR-line	bar
p_2	pressure level load line	bar
Q	flow	□/min
Q_0	flow low pressure CPR-line	bar
Q_1	flow high pressure CPR-line	bar
Q_2	flow load line	bar
R	arm length	m
v	velocity	m/s
V	geometrical displacement	cc/rev
α	effective arc-length of port 1	
β	effective arc-length of port 2	
γ	effective arc-length of port 0	
δ	control angle of the IHT port plate	
ξ	displacement ratio	-
φ	angular position	
CHT	Combined Hydraulic Transformer	
CPR	Common Pressure Rail	
IHT	Innas Hydraulic Transformer	
BDC	Bottom Dead Center	
TDC	Top Dead Center	

REFERENCES

[1] Sgro, S., M. Inderelst, and H. Murrenhoff (2010), Energy efficiency of mobile working machines, in 7. Internationales Fluidtechnisches Kolloquium, Apprimus Verlag: Aachen, Germany. p. 201-212.

[2] Achten, P.A.J., G.E.M. Vael, and K. Heybroek (2011) Efficient hydraulic pumps, motors and transformers for hydraulic hybrid systems in mobile machinery, in 1. VDI-Fachkonferenz Getriebe in Mobilen Arbeitsmaschinen, Friedrichshafen, DE, 7.-8. Jun, 2011, VDI-Wissensforum: Düsseldorf. p. 1-19.

[3] Inderelst, M., et al. (2011) Energy efficient system layout for work hydraulics of excavators, in The twelfth Scandinavian International Conference on Fluid Power SICFP'11. Tampere, Finland. p. 177-191.

[4] Heybroek, K. (2012) Hydrauliska transformatorer i hjullastare, in Hydraulikdagar, April 17-18, Linköping. 2012. p. 17-18.

[5] Heybroek, K., G.E.M. Vael, and J.-O. Palmberg (2012) Towards Resistance-free Hydraulics in Construction Machinery, in 8th International Fluid Power Conference, Dresden, March 26-28, 2012, Dresdner Verein zur Förderung der Fluidtechnik: Dresden. p. 123-138.

[6] Shen, W., H.R. Karimi, and R. Zhao (2019) Comparative analysis of component design problems for integrated hydraulic transformers. The International Journal of Advanced Manufacturing Technology.

[7] Jun, G., et al. (2020) Potential energy recovery method based on alternate recovery and utilization of multiple hydraulic cylinders. Automation in Construction, 2020. 112: p. 103105.

[8] Fassbender, D., et al. (2021) Improving the Energy Efficiency of Single Actuators with High Energy Consumption: an electro-hydraulic extension of conventional multi-actuator load-sensing systems, in The 17th Scandinavian International Conference on Fluid Power, SICFP'21: Linköping, Sweden.

[9] Li, J. and J. Zhao (2021) Energy recovery for hybrid hydraulic excavators: flywheel-based solutions. Automation in Construction, 2021. 125: p. 103648.

[10] Li, Z., et al. (2021) Study on energy efficiency characteristics of the heavy-duty manipulator driven by electro-hydraulic hybrid active-passive system. Automation in Construction, 2021. 125: p. 103646.

[11] Guo, X., J. Lengacher, and A. Vacca (2022) A Variable Pressure Multi-Pressure Rail System Design for Agricultural Applications. Energies, 2022. 15, DOI: 10.3390/en15176173.

[12] Johanson, R. (1920) Improvements in hydraulic transformers. Patent GB130975A

[13] Zardin, B., et al. (2028) Modelling and Simulation of a Cartridge Pressure Amplifier. BATH/ASME 2018 Symposium on Fluid Power and Motion Control (FPMC 2018), DOI 10.1115/FPMC2018-8913.

[14] Tucker, W.R. (1949) Hydraulic booster, U.S.P. Office, Editor, H-P-M Development Corporation, Welmington, USA, United States Patent Office 2486079.

[15] Collion, R.J. and P.A. Favrin (1959) Hydraulic transformer, U.S.P. Office, Patent 2876704.

[16] dDario, F. (2022) Pressure multiplier, in Espacenet, WIPO, Editor, Camozzi Automation s.p.a, patent WO2022254262A1.

[17] Bishop, E.D. (2007) Digital hydraulic system, WIPO, Editor, Patent appl. WO 2007/065082 A2

[18] Bishop, E.D. (2009) Digital hydraulic transformer approaching theoretical perfection in hydraulic drive efficiency, in The 11th Scandinavian International Conference on Fluid Power, SICFP'09: Linköping, Sweden.

[19] Li, B. (2018) et al. Design and Realization of a New Digital Hydraulic Transformer. in 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC).

[20] Han, Y., et al. (2021) Study on Dynamic Characteristics of Digital Hydraulic Transformers. China Mechanical Engineering, 2021. 32(03): p. 284-289.

[21] Li, W., et al. (2022) Analysis of and Experimental Research on a Hydraulic Traction System Based on a Digital Hydraulic Transformer. Sensors, 2022. 22, DOI: 10.3390/s2210.

[22] Achten, P.A.J. (1996) Pressure transformer, WIPO, Editor, Innas Free Piston, Patent nr. NL9700084W

[23] Achten, P.A.J., Z. Fu, and G.E.M. Vael (1997) Transforming future hydraulics : a new design of a hydraulic transformer, in The Fifth Scandinavian International Conference on Fluid Power (SICFP '97): Linköping, Sweden. p. 287ev

[24] Vael, G.E.M., P. Achten, and Z. Fu (2000) The Innas Hydraulic Transformer The Key to the Hydrostatic Common Pressure Rail. SAE transactions, 2000. 109: p. 109-124.

[25] Hydraulic pressure transformer (1966) Bendix Corporation, Petent nr US1023995

[26] Kouns, H.H., Hydraulic transformer (1970) U.S.P. Office, Editor, Abex Corporation, Patent nr. US3627451

[27] Központi, D. (1980) Hydraulic machines, Patent appl. nr. GB2073323A.

[28] Kordak, R. (1996) Hydrostatische Antriebe mit Sekundärregelung. Der Hydraulik Trainer. Vol. 6.: Mannesmann Rexroth GmbH, Lohr am Main. 208.

[29] Dantlgraber, J. (2001) Hydrotransformator. Patent nr. DE10037114B4.

[30] Clarke, J.M. (2002) Hydraulic transformer using a pair of variable displacement gear pumps, U.S.P. Office, Editor, Caterpillar. Patent appl. nr. US2002104313A1

[31] Grethel, M. (2012) Hydraulic transformer, WIPO, Editor, Schaeffler Technologies AG & Co.. Patent appl. nr. WO2012171519A2

[32] Zhang, F., Z. Wu, and Z. Zheng (2010) Double-acting vane type hydraulic transformer, C.N.I.P. Administration, Editor, Shandong Jiaotong University. Patent nr. CN101566177B

[33] Dantlgraber, J. and M. Robohm (1997) Hydraulic transformer with two axial piston machines with single common swash plate, E.P. Office, Editor, Mannesmann Rexroth. Patent application number EP0851121B1

[34] Gagnon, P. (2016) Configuration and performance of hydraulic transformer power distribution systems, in Mechanical Engineering, University of Minnesota.

[35] Lee, S. and P.Y. Li (2014) Trajectory tracking control using a hydraulic transformer, in 2014 International Symposium on Flexible Automation (ISFA 2014): Awaji-Island, Japan.

[36] Lee, S. and P.Y. Li (2017) A hardware-in-the-loop (HIL) testbed for hydraulic transformers reasearch, in The 15th Scandinavian International Conference on Fluid Power SICFP'17: Linköping, Sweden.

[37] Lee, S. (2018) System Configuration and Control Using Hydraulic Transformer, in Mechanical Engineering, University of Minnesota. p. 294.

[38] Stroze, M. and R. Reynolds (1990) Hydraulic pump or motor with an adjustable port plate, Sundstrand Corporation. Patent application nr. GB 2225816A

[39] Mancò, S., et al. (2004) Displacement vs Flow Control in IC Engines Lubricating Pumps, SAE International.

[40] Ericson, L., S. Kärnell, and M. Hochwallner (2017) Experimental Investigation of a Displacement-controlled Hydrostatic Pump/Motorby Means of Rotating Valve Plate, in The 15th Scandinavian International Conference on Fluid Power, SICFP'17, June 7-9, Linköping, Sweden. , Linköping University Electronic Press: Linköping. p. 19-27.

[41] Kärnell, S. (2017) An Experimental Analysis of Valve Plate Control: a New Approach to Displacement Control for Hydraulic Piston Machines. p. 89.

[42] Heeger, T. and L. Ericson (2021) A New Degree of Freedom for Variable Axial Piston Pumps with Valve Plate Rotation, in The 17:th Scandinavian International Conference on Fluid Power, SICFP'21, June 1-2, 2021, Linköping, Sweden, L.E.o.M.S. Petter Krus, Editor: Linköping. p. 117-133.

[43] Rampen, W.H.S. (1992) Digital displacement hydraulic piston pump, University of Edinburgh.

[44] Pedersen, N.H., P. Johansen, and T.O. Andersen (2017) Event-Driven Control of a Speed Varying Digital Displacement Machine. in ASME/BATH 2017 Symposium on Fluid Power and Motion Control.

[45] Vacca, A. and R.S. Devendran (2016) A Flow Control System for a Novel Concept of Variable Delivery External Gear Pump, in 10th International Fluid Power Conference, 10.IFK: Dresden, Germany. p. 263-276.

[46] Devendran, R.S. and A. Vacca (2017) Theoretical analysis for variable delivery flow external gear machines based on asymmetric gears. Mechanism and Machine Theory, 2017. 108: p. 123-141.

[47] Tankasala, S. and A. Vacca (2018) Theoretical Analysis and Design of a Variable Delivery External Gear Pump for Low and Medium Pressure Applications. Journal of Mechanical Design, 2018. 141(1).

[48] Manring, N.D. (2000) Tipping the Cylinder Block of an axial-piston swash-plate type hydrostatic machine. Transactions of the ASME, 2000. 122: p. 216-221.

[49] Achten, P. and S. Eggenkamp (2017) Barrel tipping in axial piston pumps and motors. Proc. 15th Scandinavion International Conference on Fluid Power SICFP17, June 7-9, 2017.

[50] Zhang, J., et al. (2022) Modeling and Analysis of the tilt behavior of the cylinder block in a high-speed axial piston pump. Mechanism and Machine Theory, 2022. 170: p. 104735.

DYNAMIC VALVE PLATE DESIGN FOR AN AXIAL PISTON PUMP (SERVO-LESS PUMP)

Dr. Jaromír Tvarůžek[1]*, Dr. Robert Rahmfeld[2], Carsten Fiebing[3], Aaron Krahn[4], Wilhelm Göllner[5]

[1]*Danfoss Power Solutions Open Circuit Pumps Engineering, 2800 East 13th Street Ames IO 50010, USA*
[2]*Danfoss Power Solutions Hydrostatics Engineering, Krokamp 25, 24539 Neumünster, Germany*
[3]*Danfoss Power Solutions Hydrostatics Engineering, Krokamp 25, 24539 Neumünster, Germany*
[4]*Danfoss Power Solutions Open Circuit Pumps Engineering, 14900 Technology Dr, Eden Prairie, Mn 55344; USA*
[5]*Danfoss Power Solutions Hydrostatics Engineering, Krokamp 25, 24539 Neumünster, Germany*

* Corresponding author: Tel.: +1 515 509 6785; E-mail address: jtvaruzek@danfoss.com

ABSTRACT

In times with a need to reduce CO2 emissions we are focusing on reducing fuel with improving the power-losses of hydrostatic pumps and motors. Key areas of high losses in typical designs are the cross-port at the valve plate and leakages in the servo system and at control spools. The elimination of just the cross-port is discussed in [1]. The proposed concept is a new solution using an active special valve plate porting to reduce the cross-port losses, which also controls the displacement of hydrostatic unit. The active porting strategy has in fact direct influence on the servo moments of the hydrostatic unit. With special control software, which controls the porting at the valve plate, it is possible to increase the efficiency, influence the noise and control the displacement, respectively.

Keywords: Valve plate, Efficiency, Optimized noise, Servo-less axial hydrostatic unit, Digital control, Multimode control

1. STATE OF THE ART

Today's valve plate porting design (Figure 1) uses cross-port technology to make longer time for the pressure change compared to no cross-port. With this the noise is improved while the efficiency is degraded. Typical cross-port values are between 10° to 25°. The cross-porting areas/profiles are controlling the pressure change. [2]

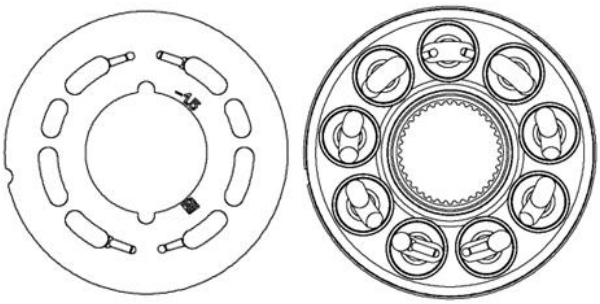


Figure 1, valve plate and cylinder block

The swash plate angle and shaft speed drive the pump flow. There are two designs of the swash plate and servo system. The servo system moves the swash plate into max stroke against the neutral return mechanism (Figure 2) or the servo system moves the swash plate from the max stroke against the maximum stroking feature (Figure 3), [6].

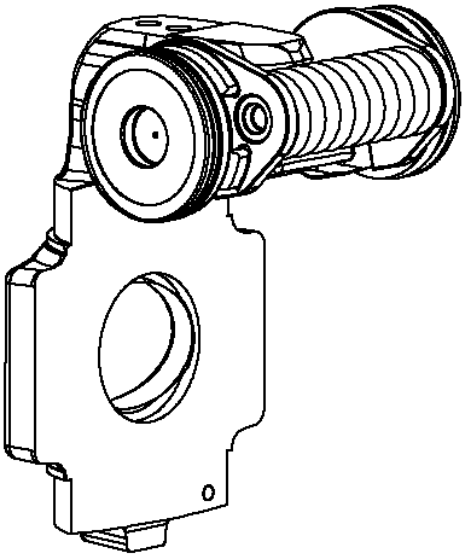


Figure 2, Neutral return mechanism

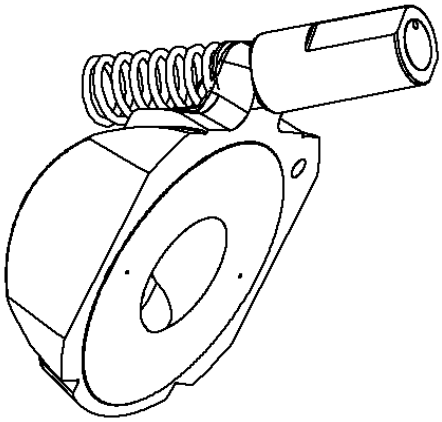


Figure 3, Max stroke feature

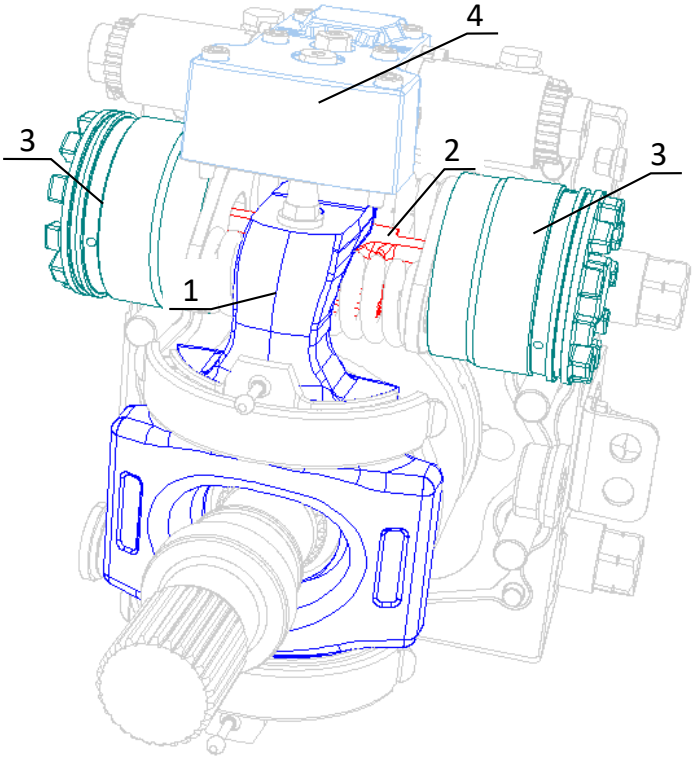
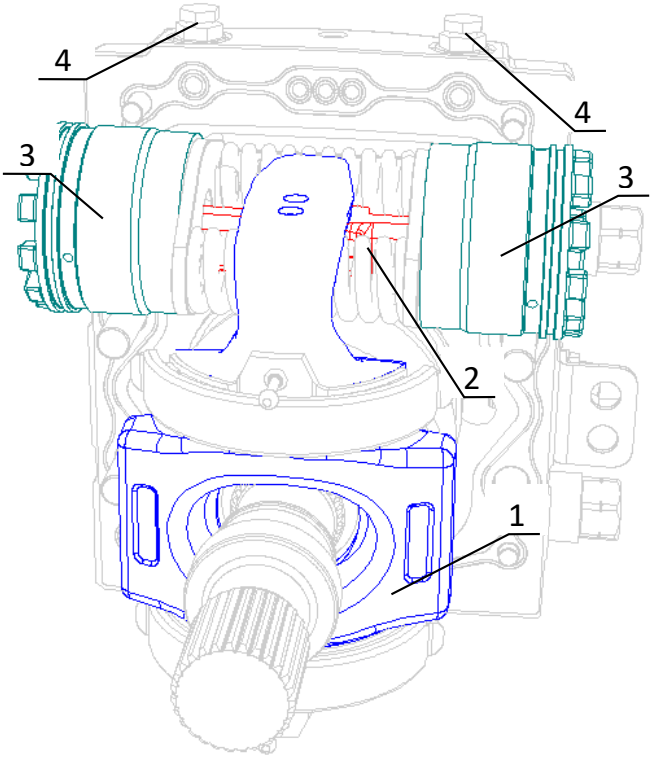


Figure 4, Close circuit servo pump

In typical closed circuit designs, the input, electric current of electronic displacement control (4), H1P service guide [5], is transferred into spool position including the mechanical feedback system. The control flow goes in and out servo cans (3). Servo piston (2) holds the swash at commanded position using servo arm (1).

2. DYNAMIC VALVE PLATE DESIGN

2.1. Concept



A standard close circuit pump is modified to verify the concept of the dynamic valve plate design. The control is removed. Servo cans (3) are drained to a housing. The servo piston assembly (2) with the springs has the function of the neutral return mechanism. Two rotary valves (4) are added into the end cap.

The rotary valves have a function of the variable orifices controlling the cross-porting flow, which has impact on the kit moments. The kit moments are capable to hold the required swash plate position at any operating conditions. More details and description in the chapter 6.

Figure 5, Dynamic (servo-less) valve plate pump

3. EFFICIENCY, NOISE AND KIT MOMENTS

The focus here is the axial hydrostatic pump efficiency. There are two group of losses: the kit losses and the control losses. The expectation is to make a significant change of the Kit losses and a step change of the control losses. The hydrostatic unit efficiency is closely connected to the noise creation, kit moments and cavitation in the rotating group.

It means that the goal of the design is to create a solution, which is the best point for all these four phenomena. [2]

3.1. Adjustable-adaptive porting design

Current designs are using a valve plate and a cylinder block kidney design, which is optimized for a specific operating conditions, such as high pressure, high rpm and max. displacement (all together). Not every unit operates in those high conditions and on vehicle base, we see high amount of time of middle to low conditions. Here the porting then is not at optimum, which means, that efficiency could be improved, and noise could be reduced.

The specific valve plate operated at max speed and mid pressure can make unit efficient and quiet, but the same valve plate operated at the low speed and high pressure is not capable to be efficient due to cross-porting leakages. Our intent is to design a valve plate, which is capable to adapt or to adjust to any operating conditions to keep the high efficiency in the whole range of the operating conditions.

3.2. Valve plate transient areas

The valve plate has two valve areas. One valve area controls the transition from low pressure to high pressure and the second valve area controls the transition from high pressure to low pressure. The traditional valve plate has a rigid timing between these two areas. The adjustable-adaptive valve plate can control independently each of those valve areas. From the flow point of view these two valve areas are separated and independent, but from the swash plate vibration and noise point of view they are coupled. The independent controllability enables the optimization of the noise of the unit.

3.3. Servo-less unit

A conventional pump with servo system needs the energy to overcome the kit moments. In case of the servo-less unit this energy burnt in the control and servo system is not needed. At conditions when the servo unit is operated at the hard stop with no servo pressure there is no space for improvement but in the case of the partial stroke or fully active control and max servo pressure the step efficiency improvement is possible.

4. PORTING DESIGN AND KIT MOMENTS

The porting design is connected to the pump performance and is impacting the design of all parts involved in the pump flow control.

The fluid dynamics of the transitions impact the cavitation and generates fluid born noise. It also defines the swash plate moment and excites the swash plate vibration. The servo system needs to overcome the swash plate moments. It specifies the needs for the servo system sizing and charge pressure value. Swash plate vibration is causing the servo piston and the servo cans vibration. This vibration generates the mechanical born noise. The servo sizing specifies the control flow needs.

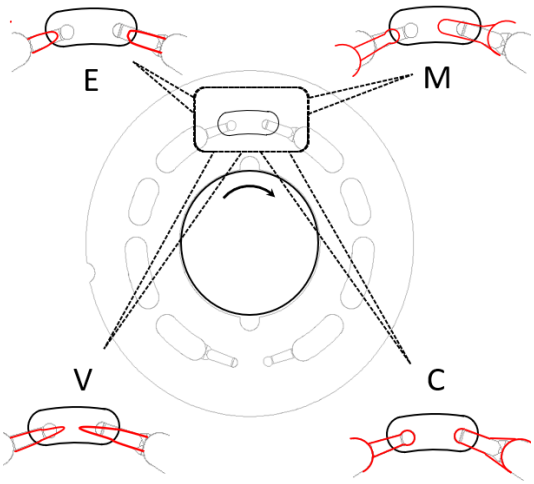


Figure 6, impact of porting design

The porting design is compromising four different areas: efficiency, noise and stability, servo sizing with charge pressure and cavitation (figure 6). In general, it is known how to make the porting design:

- (E) efficiency optimized
- (V) vibration and noise optimized
- (M) higher de-stroking (red) or lower de-stroking (black) moments
- (C) cavitation optimized

An adjustable-adaptive porting design allows to remove the servo system and at the same time enables to operate the pump at the mode optimized for the efficiency or for the noise and stability.

4.1. Kit moments – general description and terms

Kit and swash plate are shown in Figure 7. In case of the swash plate (2) design of an axial piston pump the rotating kit boundaries are the swash plate running surface (3), the valve plate running surface (4) and the cylinder block hub (6). The cylinder block hub is heavily loaded by torque – moment around the shaft axis. To meet a cylinder block (1) life expectation; any moment perpendicular to the shaft axis and transferred by the cylinder block hub (6) must be close to zero. It is achieved by having sweet point (13) aligned with the swash plate rotational axis. It means that the kit moments are balanced just by valve plate running surface (4) and swash plate running surface (3). The kit moment is the moment (7) applied on the swash plate (2) and it is equal to the moment (8) applied on the valve plate (5). Assuming there is a neutral return mechanism de-stroking swash plate (2) to the zero position by the moment (9). If the moments (9) and (7) are balanced, the swash plate is staying in the given angle and the pump provides the flow given by the kit geometrical volume and shaft speed. The moment of the neutral return mechanism (9) is balanced by the valve plate, so the final valve plate moment is the summation of the moments (8) and (10).

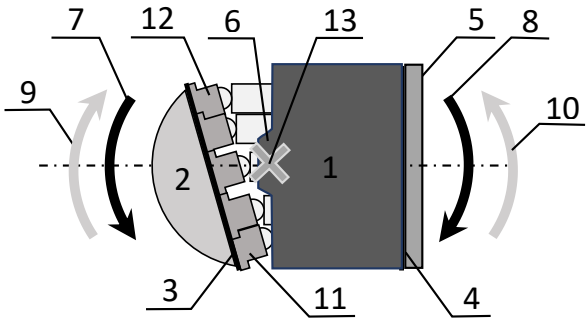


Figure 7, Kit moments

4.2. Valve plate indexing and cross-porting

There are inlet or low pressure kidneys and system or high pressure kidneys. So, there are two transient areas. In case of pumping mode, when piston is in the most inserted dead centre position IDC (11), it is transferring from the system/high pressure to the inlet/low pressure and when piston is in the most outer dead centre position ODC (12) it is transferring from the inlet/low pressure to the system/high pressure.

The indexing and cross-porting of the valve plate are two basic parameters influencing the kit moments. The figure 8 represents indexing of the valve plate. The figure 9 represents cross porting of the valve plate. The graphs are simplified to show the fundamental pressure behaviour. The overshoots and the undershoots are not shown.

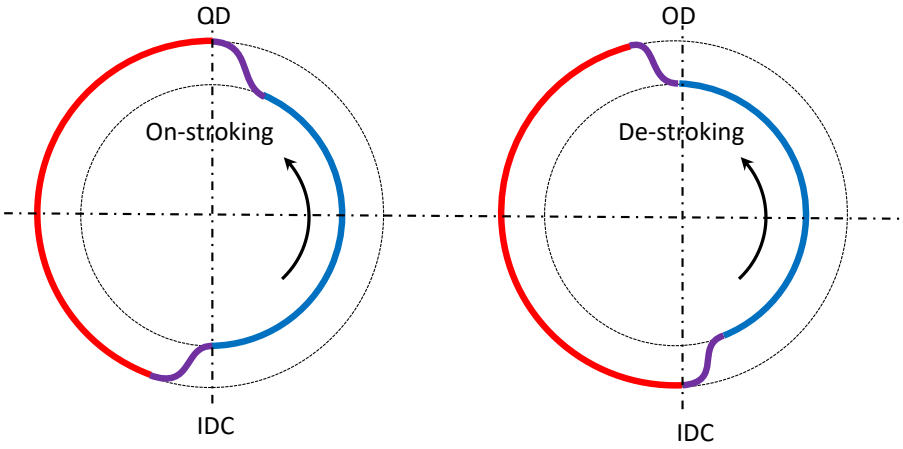


Figure 8, Indexing

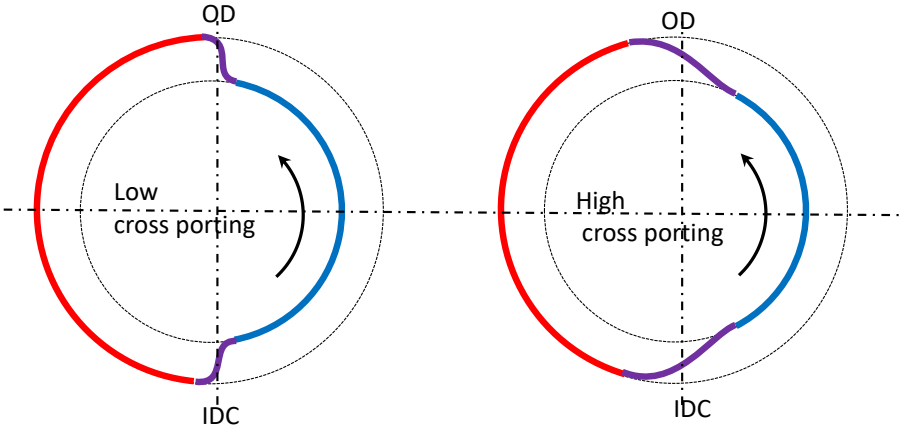


Figure 9, Cross porting

Indexing and cross porting is a result of the flow from the piston bore close to the dead centre to the coming kidney. These flows at the dead centres drive the pressure profile of each piston bore, where summation of all piston/slipper forces determines the moment applied on the swash plate.

The moments driven by pressure profile can be named the **valve plate moments**. The moment applied on the swash plate is the summation of the valve plate moments and moments due to inertia forces and piston bore frictions.

4.3. Default swash plate position

There are two options of the pump default position, the zero stroke (using the neutral return mechanism) and the full stroke (using the maximum stroke feature). Let's focus on the pump with the zero-stroke default position.

The sweet point, figure 7 (13) is well aligned with the swash plate axis. To have zero stroke as default position there is a neutral return mechanism, figure 2, which is pushing the swash plate to the zero

position plus there is a threshold moment, which needs to be overcome to push the swash plate from the zero position. The neutral return mechanism is providing default de-stroking moments. This moment can be named the **default kit moment**. It means there are in general two options of the default moment: on-stroking or de-stroking.

5. VALVE PLATE FUNCTION

A valve plate has multiple functions. From the efficiency point of view the thrust bearing function has impact on the mechanical efficiency, the seal function has impact on the volumetric efficiency. The pressure profile control function has impact on both efficiencies, the mechanical and the volumetric efficiency.

5.1. Thrust bearing and seal land

The valve plate is the thrust bearing carrying the kit axial force. The cylinder block is the rotating part, and the valve plate is a non-moving part. Each piston bore must be sealed. The size, the shape and the tribology of the seal land is designed to generate an oil film capable to carry the axial kit force and to minimize the leakages.

5.2. Pressure profile control

The position of the valve plate kidneys has the function of timing. Timing means that the cylinder block kidney is connected to the high- or low-pressure port. In addition, in the traditional valve plate, there are control grooves and or porting holes.

Usually there are two valve plate parameters, the indexing and the cross-porting. Indexing is the timing of the ODC transition from the low pressure to high pressure and of the IDC transition from the high pressure to low pressure. The valve plate is typically indexed (figure 8) in the rotation direction, which is causing bigger de-stroking kit moments, leading to a larger servo system and a safe stroke to neutral. This is a very simplified statement. In the details it is more complex. The kit moments are functions of the five independent operating variables (5D operating space) like shaft speed, inlet - and outlet- pressure, swash plate stroke and oil viscosity.

Cross-porting (figure 9) is a short cut connection between both pressures at IDC and ODC. In general, the bigger cross-porting is smoothening the pressure transition, which is reducing pressure oscillations, noise and increasing the life of the unit. The downside of bigger cross-port is increased internal leakage and lower volumetric efficiency. So, from the efficiency point of view, a smaller cross-port is better.

The complexity is, that the valve plate including control grooves is optimized at 5D operating space.

6. DYNAMIC VALVE PLATE COMPONENTS

The new valve plate porting design has the capability to modify the indexing and cross-porting according to the operating conditions to optimize the noise and the efficiency. It contains two parts: static and active.

6.1. Valve plate static parts

The static part has all functions, like a traditional valve plate but the pressure profile control is modified by added control ports (at ODC and IDC, figure 10).

- The transition part of the pressure profile is bounded by four edges.
- The ODC transient area is starting when the cylinder block piston bore kidney leaves LP ODC Edge
- The ODC transient area is finishing when cylinder block piston bore kidney hits HP ODC Edge
- The IDC transient area is starting when cylinder block piston bore kidney leaves HP IDC Edge
- The IDC transient area is finishing when cylinder block piston bore kidney hits LP IDC Edge

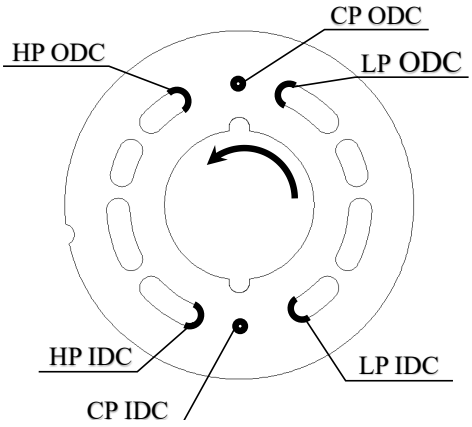


Figure 10,
Passive part of the dynamic valve plate

6.2. Valve plate active part

The active part has two control valves, which are controlling indexing and cross-porting. (figure 11)

- Control port CP ODC is connected to the high-pressure port via the flow valve CV ODC.
- Control port CP IDC is connected to the low-pressure port via the flow valve CV IDC.

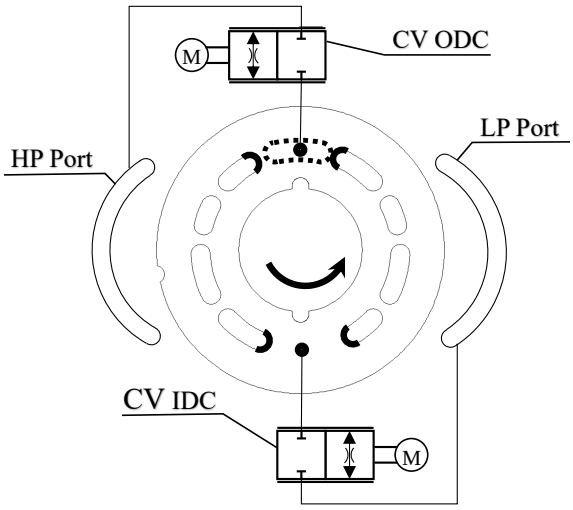


Figure 11, Active parts of the valve plate

7. THREE TRANSITIONING PHASES

At both dead centres there are three transitioning phases.

7.1. Phase A

The cylinder block kidney moves from low pressure and then connects to the CP ODC (figure 12). The piston is close to its dead centre and makes only very little axial movement / stroke.

When it comes to control port CP ODC the oil flows in the bore from the HP port through flow valve CV ODC (figure 12). The cylinder block kidney can be connected to the low-pressure kidney. In that case there is a cross-porting between the low pressure and high pressure controlled by the CV ODC valve. The second option is that the space between the control port CP ODC and low-pressure kidney is bigger than the size of the cylinder block kidney. In that case there is no cross-porting, and the oil is trapped in the cylinder block bore.

Phase A, cylinder block kidney is coming over the control port CP ODC

The angular space between Low Pressure Edge and control port is:

- bigger than cylinder block kidney, then oil is trapped in the bore for a short moment.
- smaller than cylinder block kidney, (figure 12), then there is a cross-porting controlled by control valve CV ODC.

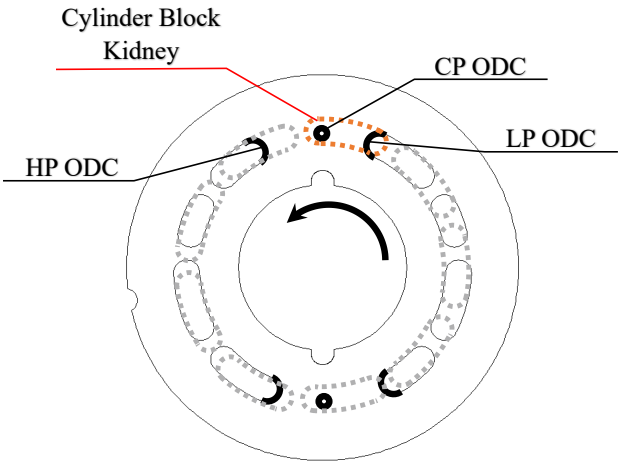


Figure 12, Phase A

7.2. Phase B

After leaving the LP ODC edge and hitting the control port CP ODC the flow in and out of the cylinder block bore is fully controlled by the CV ODC valve (figure 13).

Phase B, cylinder block kidney is going just over the control port CP ODC

The bore pressure **increase** is fully controlled by the Control valve CV ODC; there is a flow from the high-pressure port through the valve to the control port CP ODC to the piston bore.

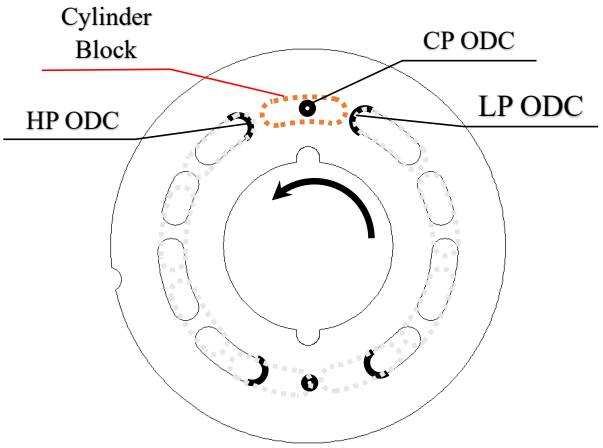


Figure 13, Phase B

7.3. Phase C

The last phase starts by hitting of the HP ODC edge. If the piston bore pressure is still lower than the system pressure, it is quicky balanced by the direct overlap with high pressure kidney.

Phase C, cylinder block kidney is coming over the edge HP ODC

The angular space between Low Pressure Edge and control port is:

- bigger than cylinder block kidney, then oil is trapped in the bore for a short moment.
- smaller than cylinder block kidney (figure 14), then there is an accelerated pressure balance if not already equal to the high pressure.

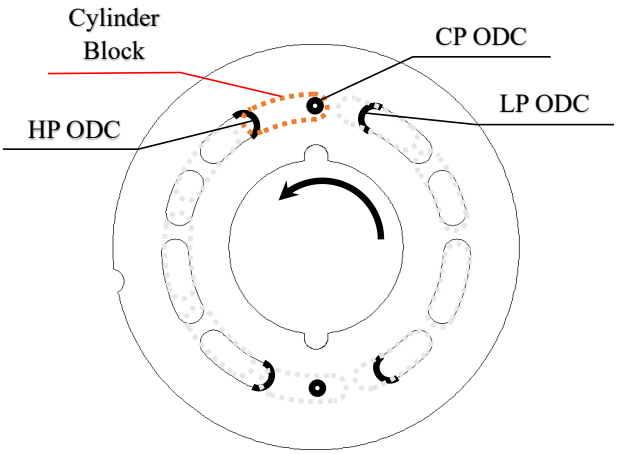


Figure 14, Phase C

At the IDC transition area there are the similar A, B and C phases like in ODC.

7.4. Relevance of the phase A, B and C

The capability to simulate and to understand the flow and pressure changes in the bore located at the dead centre are critical for the dynamic valve plate concept. There are a few milliseconds available to control the pressure profile during these three phases. Is it possible? At first this theory was tested using the digital twin of a 130cc axial piston servo pump.

The simulations were leading to the following conclusions:

- The variable orifice is capable to control the pressure profile to achieve the balance of the Kit moments and default kit moments, so to control the displacement of the pump. This means that the servo-less pump can be developed.
- The removing of the servo system leads to the unit efficiency improvement.
- The cross-porting can be optimized for all operating conditions including the variable pump displacement to optimize the efficiency.
- The non-synchronous/individual IDC and ODC valve opening is influencing the dynamics of the kit moments, which has impact on the pressure oscillations and noise.

8. DYNAMIC VALVE PLATE TESTING

The first feasibility study was done by a modified serial production closed circuit 130 cc pump.

8.1. Breadboard unit

Modified pump:

- Close circuit pump (1) with modifications described at chapter 2.1 and shown on fig 5.
- Two valves (CV IDC and CV ODC) are operated using 4 bar mechanism (2).
- The couple of 4 bar mechanisms is operated by Lego servo motors (3).
- The Lego micro-controller (4) is controlled by code created in the Lego application in IOS.
- 3D printed 4 bar mechanism is used to achieve an acceptable precision of valve position and to overcome the flow force by low torque Lego servo motors.

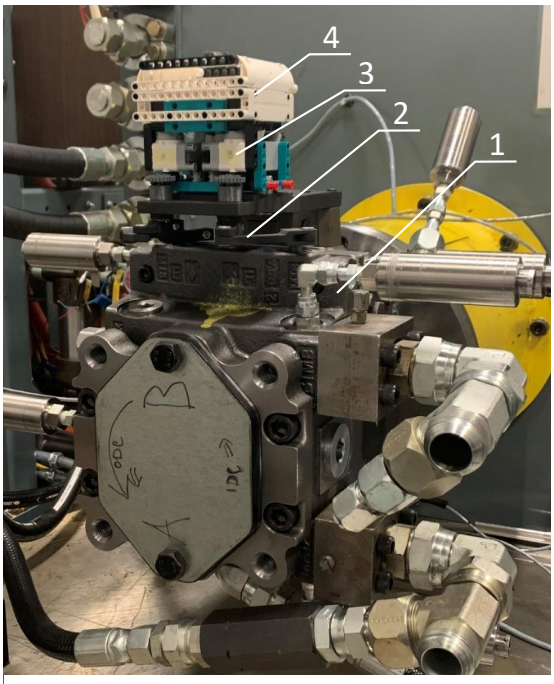


Figure 15, Modified pump

8.2. The LEGO benefit - repeatability

The active part of the dynamic valve plate consists of the two simple rotary valves and the servo motors with 4 bar linkage mechanism (figure 16). Gear (2a) is bounded with bar (2b). The rotary valve (6) angle is driven by the rotation of the gear (5). Each valve has its own 4 bar linkage mechanism and servo motor.

Using the LEGO battery and servo motors show the new concept capability to replace the servo system powered by a 26cc charge pump.

Using of LEGO app GUI allowed us to run multiple tests with high repeatability of the commands. It was fast, easy, and straight forward.

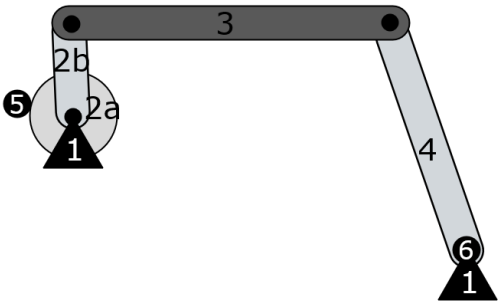


Figure 16, 4 bar mechanism

8.3. Example of the test results

There were done multiple types of the tests to study the controllability of the pump.

For example, the Lego servo motors were commanded to make six steps to close the rotary valves and six steps to open them to the original position.

The LEGO application set on the iPad was simple.

The servo motors are linear, but the 3D printed 4 bar mechanism is not.

Figure 17 shows the table of the commanded steps. Each step the servo motors turned 50 degrees. The last row shows the calculated orifice diameter corresponding to the ODC and IDC valve opening at each step. At the shaft speed 2000 rpm and delta pressure 100 bar, the swash plate starts to de-stroke at about 2.37mm orifice size. At 20% displacement the corresponding orifice size is only 1.06mm. For the delta pressure 200 bar the swash plate starts to de-stroke at about 1.84mm. The standard valve plate of this pump has over 3mm fixed corresponding orifice size. The result of the smaller cross-porting is volumetric efficiency improvement.

	step 1	step 2	step 3	step 4	step 5	step 6	step 7	step 8	step 9	step 10	step 11	step 12	step 13
position [degrees]	0	50	100	150	200	250	300	250	200	150	100	50	0
Orifice diameter [mm]	2.6	2.37	2.12	1.84	1.5	1.06	0	1.06	1.5	1.84	2.12	2.37	2.6

S

Figure 17, commanded steps

Figure 19 shows the flow of the pump. There are shown three load pressure setting (figure 18).

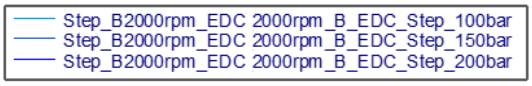


Figure 18, pressure levels

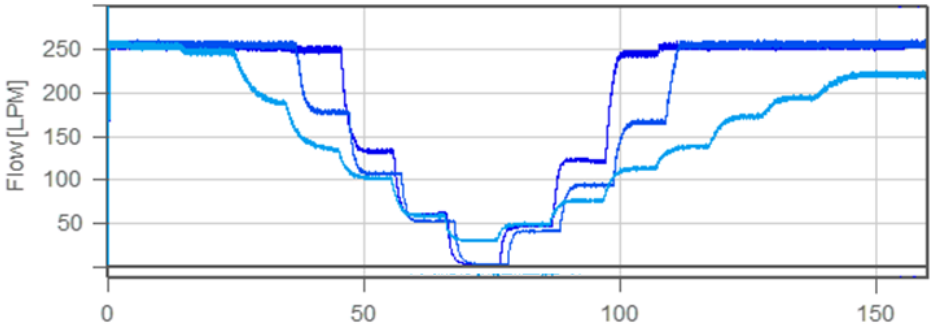


Figure 19, Pump flow at 2000 rpm

To present the smooth behavior of the pump, the low-pressure and the high-pressure in time domain is shown with the 5 kHz sampling frequency and not filtered (figure 20). The low system pressure oscillation is good indicator of the low swash plate vibration. With these results the controllability of the servo-less pump is basically proven.

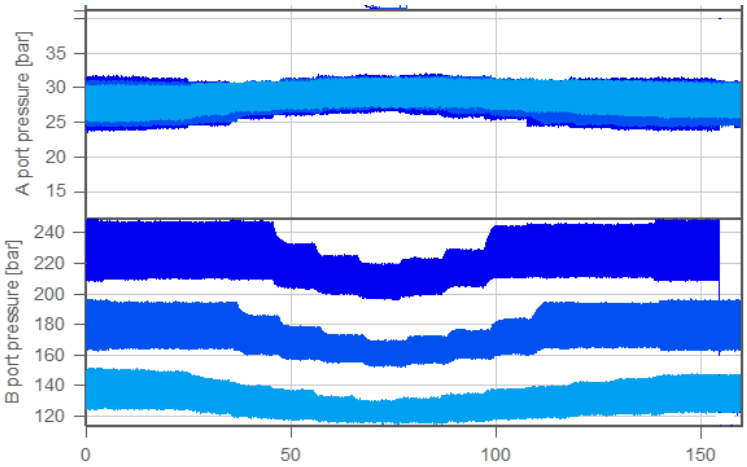


Figure 20, System pressure

9. SUMMARY AND CONCLUSION

- There is high potential and customer value propositions for developing of servo-less pumps.
  - Faster flow requirements response.
  - More compact unit, less components, lower weight.
- There is a good potential to reduce the cross-port losses of today's pumps especially at partial strokes and at higher pressures.
  - At lower stroke the ODC and IDC smaller orifice is reducing cross-porting losses.
  - At higher pressure and lower speed, the CV ODC and the CV IDC smaller orifice is needed, higher efficiency.
- The individual control of the valves has a significant impact on the system pressure oscillations, and it can be used to optimize the noise.
  - By holding of the CV IDC (or CV ODC) and varying the CV ODC (or CV IDC) have significant impact on the noise and sound quality.
- The results of the testing are especially convincing to use the dynamic valve plate concept by applying PLUS+1® software.
  - The vision is to use inputs from multiple sensors. The code will process inputs from sensors and command from the machine operator to run the CV IDC and CV ODC. The dynamic valve plate enables the pump to operate at optimized noise, efficiency and stability conditions.
  - Using of the pressure and the swash plate angle sensor allows to run the pump at the displacement or pressure or power control mode based on the needs of the machine.
- It is already a significant innovative result that it was possible to replace the servo system powered by charge pressure and flow at about 2-4 kW just by a LEGO battery. This concept enables the advantage of the reduced cross-porting flow at lower stroke, to have significant impact on the improved kit efficiency. In addition, together with the smart hydraulic system

utilizing multiple sensors it has in fact the potential to be a market game changer. The last but not least important point is that this technology is using existing series production pump hardware.

10. ABBREVIATIONS

- ODC – the kit piston position at the dead center with the biggest oil volume in the cylinder block bore
- IDC - the kit piston position at the dead center with the smallest oil volume in the cylinder block bore
- CP ODC – the control port near ODC position
- CP IDC - the control port near IDC position
- CV ODC – the control valve for the kit piston at ODC position
- CV IDC– the control valve for the kit piston at IDC position
- EDC – electronic displacement
- B_EDC – electronic displacement pump running with high pressure at B port
- IOS -apple operating system
- GUI – graphic user interface
- HP IDC – high pressure near the IDC position
- HP ODC – high pressure near the ODC position
- LP IDC – low pressure near the IDC position
- LP ODC – low pressure near the ODC position

11. LITERATURE

1) Ivantysyn J. and Ivantysynova M. 2001. Hydrostatic Pumps and Motors. Academic Books International. New Delhi.
2) Ivantysynova M., Impact of valve plate design on noise, volumetric efficiency, and control effort in axial piston pump, ASME 2006
3) Computational Valve Plate Design in Axial Piston Pumps/Motors, Received 04 February 2019; Accepted 28 August 2019; Publication 19 September 2019
4) Achten P., Shuttle Technology for noise reduction & efficiency improvement of hydraulic machine, FPMC 2021
5) Danfoss, Service guide, H1 045/053 Tandem Axial Piston Pumps Parts Manual, AX152886482422en-000802.pdf, December 2021
6) Danfoss, User guide, Series 45 Axial Piston Open Circuit Pump Technical Information, BC152886483703en-001301, March 2023
7) Katharina Schmitz, Hubertus Murrenhoff, Fundamentals of Fluid Power Part 1: Hydraulics, 2nd edition 2018, ISBN: 978-3-8440-2826-3
8) Daniil Dumnov, Jérémie Lagarde, Efficiency of a Digital Displacement Pump Operating With Partial Strokes, ASME/BATH 2023 Symposium on Fluid Power and Motion Control, FPMC2023-111388, V001T01A017, October 16–18, 2023
9) Rahmfeld, R. 2020. Displacement Machines - Key Elements of Future Technology. General lecture. 12. IFK (Int. Fluid Power Conference), Dresden, Germany.

Chapter 19

System Layouts in Mobile Machines

ELECTRO-HYDRAULIC DAMPING STRATEGIES FOR HYDRO-PNEUMATIC SUSPENSIONS

Steffen Antoni[1]*, Dr.-Ing. Lars Brinkschulte[1]

[1]*ARGO-HYTOS GmbH, Industriestraße 9, 76703 Kraichtal*

* Tel.: +49 7250 76 542; E-mail address: s.antoni@argo-hytos.com

ABSTRACT

Operators of mobile machines are exposed to high vibration loads while on the machine. To reduce this, seat, cab, axle and wheel suspension systems are used today. In addition to a uniform ground pressure required to steer and brake the machines safely, the main aim is to reduce whole-body vibrations. The driving comfort can be improved and component stress can be reduced. The wide range of applications for such machines means that the suspension systems have to meet special requirements in various driving and working conditions. Due to characteristics suitable for mobile machinery, hydro-pneumatic suspension systems are increasingly being used in such machines. Adjusting the damping - either adaptively, semi-actively or actively - is one way to react quickly to changing conditions. This paper will focus on presenting and evaluating different damping strategies for hydro-pneumatics suspension systems. After a general overview of hydro-pneumatic suspension systems, the various damping strategies are discussed in detail. The effects of three damping strategies are explained in detail using the example of a tractor cab suspension.

Keywords: hydro-pneumatic suspension, damping strategies, mobile machine, tractor cab

1. INTRODUCTION

As long as mobile machines are not fully automated, operator comfort will remain one important requirement for those machines. In modern agricultural tractors, for example, the cab has become more than just the place for the driver to sit – nowadays, it is a workplace for many hours a day, especially during peak working seasons. In addition to comfort and efficiency, safety at work, thus the protection of operators' health, has gained importance. Diverse standards and regulations determine the dose of vibration to which a driver may be exposed within a certain period of time and how this dose can be measured [1].

In addition to driving comfort, controllability and tire-to-ground contact are also important. To ensure this, machine manufacturers are continuously improving their machines with active seats, suspended axles or wheels and specially designed cab suspensions.

2. STATE OF THE ART

The basic function of a suspension system is to isolate certain parts of a machine (usually the chassis), or the operator him- or herself from vibrations from the ground. The frequency range of suspension systems is below 25 Hz. In the case of vehicle dynamics, "operator comfort" is related to only this part of the frequency range [2]. With suspension systems, stress for driver and components can be reduced. ISO2631-1 [3] lists several diseases caused by vibrations on the human body – for example, excitations in the frequency range of 0.5 Hz to 0.75 Hz can be responsible for motion sickness.

As a result, suspension systems are used in a wide range of mobile applications. While mechanical spring and damper systems (see **Figure 1**, left) are typically used in passenger cars because of their simple design and low load changes, pneumatic systems are used in commercial vehicles. In these applications, additional features, e.g. lowering the chassis of buses for a better entry of passengers, are requested. Hydro-pneumatic systems (see **Figure 1**, right) are widely used in mobile machines. Here, especially the rough conditions and the higher load changes are main drivers for this technology. In contrast to pneumatic suspensions, hydro-pneumatic systems have separated media (oil and gas), which allows hydro-pneumatic suspensions to be blocked easily.

The main difference between mechanical springs and (hydro-)pneumatic springs is the progression of the spring rate. Mechanical springs usually have a linear characteristic (progressive or degressive is also possible). Pneumatic and hydro-pneumatic springs typically show a progressive behavior, which leads to a higher spring rate when the load is increasing and vice versa. In this way, the suspension characteristics are adapted automatically to the axle load. Depending on the layout, an approximation to a linear curve is possible with these systems, as well.

Apart from the main advantage of adapting the spring rate to the load, hydro-pneumatic suspensions are also beneficial when [4]:

- the level control has to deal with high load changes
- the suspension needs to be blocked
- the level control needs to react frequently and quickly
- the spring rate should be adjustable
- a compact and robust design is required

Tractors often have hydro-pneumatic suspension systems for front axles and cabs. Therefore, this type of machine is used as the reference application in this publication but should generally be seen as one of many mobile machines.

2.1 Hydro-pneumatic suspension systems

The suspension effect of hydro-pneumatic springs is mainly generated by the compressibility of the gas inside the accumulator. The compressibility of the hydraulic fluid (~1,2 % @ 200 bar [5]) and the elasticity of the hydraulic hoses only contribute a negligible amount. The spring rate is determined by the size of the accumulator, its precharge pressure and the load on the suspension system. Because of the wide load change of mobile machines, it makes sense to adapt the spring rate according to the actual load condition. In combination with higher spring rates, higher damping settings can increase comfort. For the front axle of tractors, the load can change depending on the work being done. A plow on the rear hitch reduces the front axle load, while a loaded front loader increases it. By using a differential cylinder, the spring rate can be influenced by a preload pressure in the rod side of this cylinder. Increasing the preload in the rod side leads to an increased spring rate of the suspension and vice versa [4].

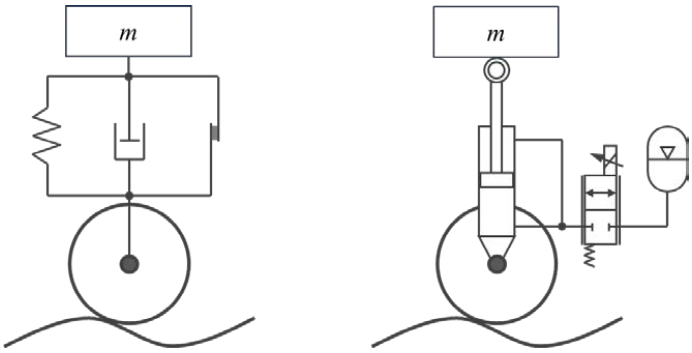


Figure 1: Mechanical (left) and hydro-pneumatic (right) suspension system

In addition to the suspension effect, unwanted damping in a hydro-pneumatic suspension can be caused by the following:

- solid friction in seals and bearings and
- viscous friction in tubes and pipes.

Friction causes a pressure difference that acts upon the active area of the cylinder, creating a retarding force. The viscous friction in tubes and lines should be kept rather low. However, under certain conditions it is helpful to generate additional damping. This is done by adding a flow restrictor between the cylinder chamber and the accumulator. An adjustable flow restrictor, like a variable orifice or an electronically controlled proportional valve, allows damping settings to be adjusted to the driving and working conditions.

Another core function of hydro-pneumatic suspensions is level control. Sensors detect a deviation of the suspension position, e.g. after a load change. An ECU is controlling a valve to reposition the suspension cylinder – see **Figure 2**.

This article focuses primarily on the various damping strategies. Detailed information on level control and adjusting the spring stiffness of hydro-pneumatic suspension systems can be found in [6] and [7].

2.2 Damping strategies for hydro-pneumatic suspension systems

The damping strategies for suspension systems can be classified into four different categories:

- passive systems
- adaptive systems
- semi-active systems
- fully-active systems

Passive systems are characterized by constant spring stiffness and fixed dampers. A soft spring rate and low damper settings offer a good body isolation, while a stiff spring rate and firm damper settings create good controllability on the machine body motion. Due to that fact, the behavior of passive systems is always a compromise between riding comfort and machine stability. Especially when there is a higher load change, these systems are designed to be rather rigid, creating performance losses in comfort.

Adaptive systems were developed to overcome these challenges. Depending on the working or driving conditions – e.g. driving on an even road vs. driving in the field – the damping settings can be adapted.

The easiest way is if the operator adjusts the settings himself. Usually, the operator can select between a soft, medium and stiff setup. In certain situations, the operator can even block the suspension. An advantage of adaptive systems is that the operator can adjust the damping settings according to his wishes; however, it is not possible to adjust properly to temporary changes. More advanced adaptive systems use electronic controls to switch the suspension characteristics automatically based on sensory information. Since only steady state changes (e.g. rough terrain vs. smooth terrain) are considered, low bandwidth controllers are sufficient.

Beyond that, ***semi-active systems*** react even to temporary changes, adjusting the damper setting more dynamically. By using adjustable dampers, the rate of energy dissipation can be influenced depending on the suspension motion. Aside from the more dynamic damper adjustments, the closed loop control of these settings is the main difference to adaptive approaches. The actively controlled forces in a semi-active system are only created by the damper, hence they only act against the direction of the suspension motion. The most popular semi-active suspension approaches are skyhook control and relative control. While dampers in conventional spring-damper systems reduce the relative speed between the isolated mass and the excitation, these more sophisticated solutions reduce the absolute speed of the suspended mass, thus its acceleration. With these approaches, a better combination of resonance damping and high-frequency isolation and consequently a more direct and quick reaction can be achieved. However, this is offset by higher costs for faster processors and components, as well as additional sensors that may be required.

In ***fully active systems,*** spring and damper are displaced by a force actuator. This actuator is placed directly between sprung mass and unsprung mass of the machine. The force actuator also allows to feed energy into the system independent of the direction of the suspension movement. To do so, special control algorithms are used, which energize the force actuator based on sensory feedback. The full bandwidth of the system provides good driving comfort and, at the same time, good controllability. In addition to the full bandwidth, they have a high-power consumption. This is a big disadvantage of these systems, especially against the backdrop of machine electrification. On top of this issue, the complex design makes this solution expensive, as well. Because of these reasons, fully active systems are only used in special applications.

A combination of different approaches can be used for the final application. One example is shown in **chapter 3** for a tractor cab suspension.

3. EXEMPLARY APPLICATION: TRACTOR CAB SUSPENSION

Nearly every manually driven mobile machine has an operator’s cab. This is why cab suspensions offer good opportunities to improve comfort. Since tractors are very versatile machines, they are used in a wide range of operating conditions: from working scenarios on the farm, driving on flat roads to working in rough terrain on the field. On the road, tractors are often used to transport implements to/from the field or to transport agricultural goods with a trailer. The driving speed is rather high, the controllability and the driving stability are more important than comfort. Thus, the cab suspension should be rather stiff. When driving in the field for hours with lower speeds, comfort and reduction of transmitted whole-body vibration are most important. For this type of work, the cab suspension should be set to a soft setting. This could be achieved with so-called adaptive damping: the operator selects between driving on the road or in the field. The damping setting is then switching between low and high, making the suspension feel soft or stiff. But to relieve the driver, solutions are shown that react automatically (semi-actively) to different driving conditions, without the need of the operator to interfere.

Figure 2 shows an exemplary cab suspension of a tractor and the hydraulic design. The hydraulic

manifold contains the proportional damping valve (no. 3), as well as two valves for leveling the suspension (no. 1 and no. 2). The cab can be lifted by activating valve no. 1. If the cab is supposed to be lowered, valve no. 2 needs to be opened. The damping can be proportionally adjusted by valve no. 3. There is also a small bypass orifice in parallel to protect the suspension from pressure peaks if the damping valve is completely closed. In order to guarantee the best possible body isolation, the cab suspension is typically laid out rather soft. In this case, this means a basic damping setting of 55 % (55 % is the percentage ratio between min. and max. valve opening). The soft suspension typically has a long suspension travel, which can lead to tricky situations in some driving and working conditions: excitations close to the natural frequency, heavy bumps, etc. Different concepts from chapter 2.2 are therefore combined to set up a sufficient suspension system.

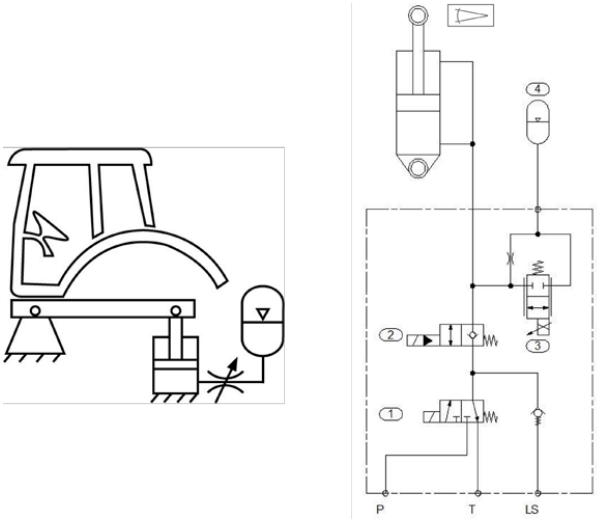


Figure 2: Cab suspension of a tractor

3.1 Relative control

To better adapt the performance of the suspension to the existing vibration conditions, the relative control approach can be used. The relative control is a modification of the skyhook control approach. To generate the connection of the suspended mass to the sky – “the hook” – the skyhook control uses an acceleration sensor. Additionally, the relative speed between unsuspended and suspended mass is measured by a second sensor. This makes the skyhook control quite complex and expensive. The relative control only needs the existing position sensor already installed for level control of the suspension. This makes this solution much more cost-effective.

Similar to the skyhook control, the approach behind the relative control [8] is to use the damping forces whenever they are beneficial to reduce the acceleration of the suspended mass. The relative control therefore optimizes comfort by reducing these accelerations. It can be used as an extension to the standard damping and applies only at high speeds.

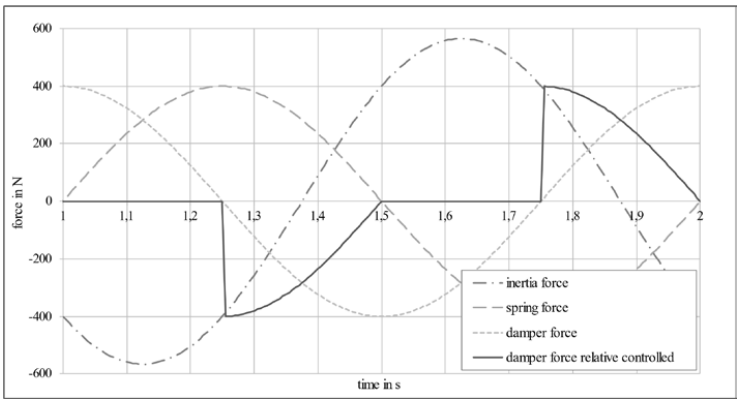


Figure 3: Relative controlled damping force in a vibration isolator based on [8]

Figure 3 shows the oscillation force curve of a passive oscillation isolator (see **Figure 1** left). In addition to the force of inertia, the spring and damper forces are shown. In certain time periods of the oscillation, the dampers´ force and the spring force have equal directions. In other periods the direction of action is different.

Because the inertia force F_I ($F_\mathrm{I} = m \cdot \ddot{x}$) of the mass corresponds to the sum of spring ($F_\mathrm{S} \sim x$) and damper force ($F_\mathrm{D} \sim \dot{x}$), the following equations explain the context:

$$|\ddot{x}| = \frac{|F_S| + |F_D|}{m}, t \in \left[t_0;\ t_0 + \frac{T}{4}\right] \cap \left[t_0 + \frac{T}{2};\ t_0 + \frac{3T}{4}\right] \tag{1}$$

$$|\ddot{x}| = \frac{|F_S| - |F_D|}{m}, t \in \left[t_0 + \frac{T}{4};\ t_0 + \frac{T}{2}\right] \cap \left[t_0 + \frac{3T}{4};\ t_0 + T\right] \tag{2}$$

Because the acceleration of the mass, thus the force of inertia, corresponds to the sum of spring and damper force, the damping force reduces the acceleration of the mass in certain sections and amplifies it in others. When focusing on equation (2), it can be seen that the damping force reduces the acceleration of the mass in the time frame for which the equation is defined. Because of this, the damping forces must only be applied to the mass in the appropriate time frames in order to reduce its acceleration. This leads to an on-off control of the damper. Because the spring force is not constant, the damping force should be continuously adjustable in accordance with the current spring force to optimize this approach [8]. **Figure 4** shows the continuous damping adjustments for the position curve of a cab suspension when driving on a dirt track with 20 kph with the relative control algorithm. The static stroke position of the suspension is 50 %. The static damping is unchanged at 55 %. The relative damping is added at the desired conditions, see equation (2).

As the correlation between spring force and compression of a hydro-pneumatic spring depends mostly on the system design (accumulator size, precharge pressure, etc.) and is often nonlinear – the system could further be optimized by taking this into consideration. This would help to better adapt the damping forces to the actual spring forces.

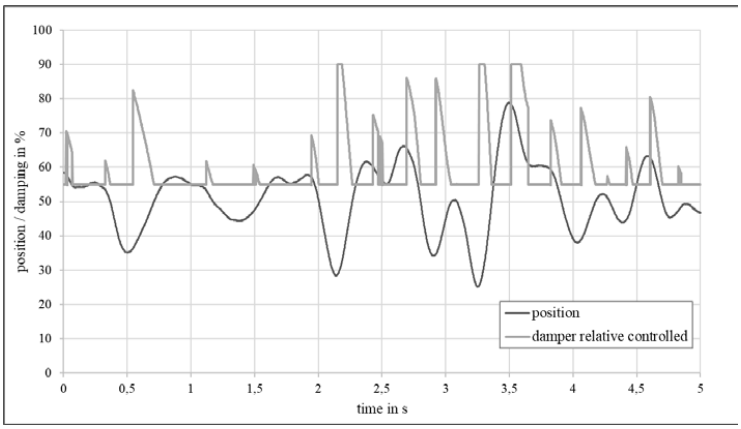


Figure 4: Damping setting relative controlled

3.2 End-Stroke-Damping

The wide range of usage of a tractor can cause situations in which a soft laid out suspension can get to its limits. High force peaks coming from the chassis, e.g. when driving over a bump or a pothole, can deflect the suspension cylinder in a soft suspension to its end stop, resulting in increased component stress or even damage to the system. Furthermore, in such situations, the forces are transferred more or less directly into the cab, which may cause injuries of the driver or loss of machine control. There are cylinders with end stop damping available, but incorporating these means extra costs for each machine. Moreover, if a damping valve is already available in the suspension circuit, it can be used for a software-based end stop damping.

Whether end-stroke-damping – which causes a deceleration force of the piston – is activated depends on two system conditions:

- the position of the suspension cylinder: the closer the piston is to one of the two end stops, the higher the risk of reaching one of them. For the suspension, this means that the damping needs to be increased, the closer the piston gets to one of the end stops.
- the speed in the direction of an end stop: the higher the speed, the more kinetic energy is stored in the system. Assuming a certain position outside the middle stroke position, higher energy means higher risk of reaching the end stop. Therefore, more damping is necessary to reduce this risk by reducing the speed.

These two conditions need to be carefully harmonized so that the damping only intervenes when necessary. Otherwise, the suspension could become too stiff under normal driving conditions, or even worse, a hydraulic end stop could occur: If the damping valve is closed too much, the oil cannot flow into the accumulator, causing a high pressure peak in front of the spool lands. Instead of hitting the mechanical end stop, the piston would hit a “hydraulic” end stop.

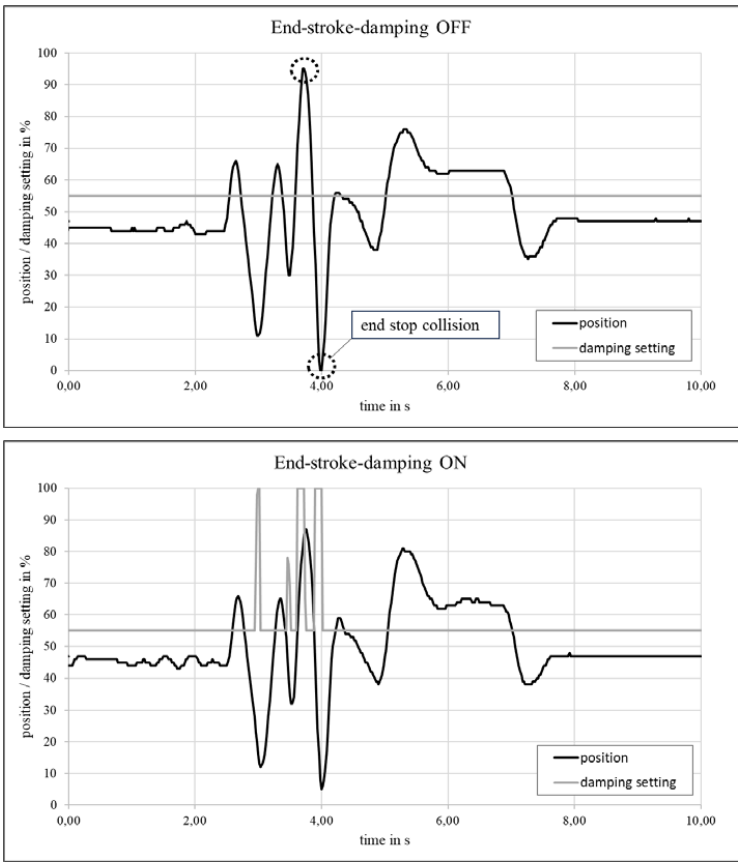


Figure 5: stroke position w/o end-stroke-damping

To explain how the damping strategy works, a test run was carried out - driving over an obstacle (height: 80 mm, length: 1200 mm) with 12 kph, see **Figure 5**. In the diagrams, the first excitations (see time at 3 s) come from the front axle, which reached the obstacle first. When the rear axle moved over the obstacle (see time at 4 s), the excitations for the cab suspension were even higher. The activated position control compensated for long-lasting deviations from the target position. This had no effect on the large oscillations, but an effect can be seen at the end of the test run.

The upper diagram shows the stroke position (black line) with the end-stroke-damping shut off. Again, a constant damping setting of 55 % was applied to the suspension. The spring system showed a relatively soft behavior, as desired. When the rear axle passed over the obstacle, the lower end stop was reached and there was a hard impact inside the cab.

The lower diagram shows the influence of the activated end-stroke-damping. The constant damping is set to 55 %, as well. With end-stroke-damping, the damping was increased, temporarily and only when necessary, to prevent end stop collision.

3.3 Harmonic Oscillation Damping

Road and driving conditions occasion the suspension systems of a mobile machine to oscillate to a greater or lesser extent. If the frequencies of the excitations are close to the natural frequency of the suspension the oscillations are superimposed. Sometimes these oscillations last for a longer period of time (e.g. if the road conditions do not change), which leads to a very unpleasant condition for the operator and could cause so-called seasickness.

To get rid of these disruptive oscillations, energy needs to be extracted with the damping valve. The harmonic oscillation damping approach detects oscillations within a certain frequency spectrum by using the position sensor signal. If these oscillations do not subside after a certain period of time, the damping setting is increased step by step, until the oscillation energy is reduced to such an extent that the continuous oscillations disappear or are reduced to a tolerable level. Damping setting will then be decreased. If the oscillations start intensifying again, the damping is increased again, as well. The advantage of this approach is that the suspension is only stiff when needed, thus achieving good decoupling under regular operating conditions.

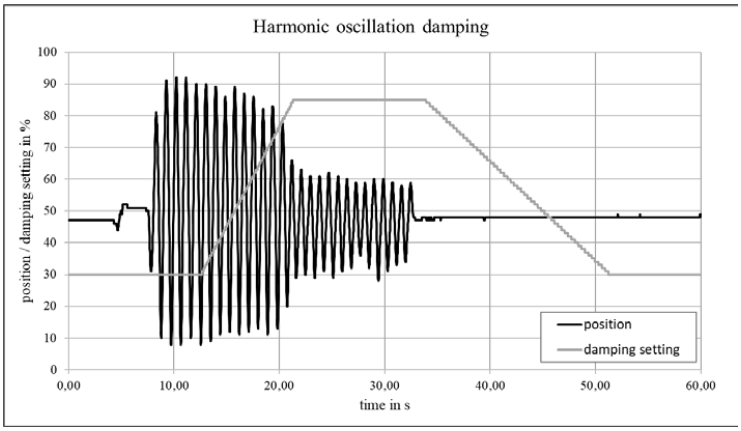


Figure 6: Oscillation reduction with harmonic oscillation damping

The diagram in **Figure 6** shows the impact of the harmonic oscillation damping. To better see the impact of the harmonic oscillation damping, the suspension was artificially deflected with a high amplitude. This was done by rhythmic acceleration of the tractor with the gas pedal. Even with excitations up to 30 s, the increasing damping setting reduced the oscillations, resulting in less movement. After the excitations stopped and the suspension movement came to a rest, the harmonic oscillation damping control reduced the damping settings to the lower level, in this case 30 % again. Depending on the setting of the gradient, the damping makes the suspension more or less rigid.

4. CONCLUSIONS

Mobile machines have a wide range of applications. This is why their suspension systems must meet a variety of requirements. The adjustability of the damping settings of a hydro-pneumatic suspension via software offers a great opportunity for individual solutions that are tailored to different operating conditions. There are numerous approaches to optimize the suspension performance.

The measurements for a tractor cab suspension showed that with end-stroke-damping and harmonic oscillation damping, there are two control strategies, which can increase the comfort in special operating conditions. At the same time, component stress can actively be prevented, extending the service life of the machine and particularly the suspension parts. The relative control can be used

without having to implement complex and expensive additional equipment. Further optimization steps could be done regarding the time response of the damping valve in order to generate better timed forces. Furthermore, the nonlinear spring curve of the hydro-pneumatic spring could be taken into account to further optimize the damping forces according to the theory.

NOMENCLATURE

F_D	Damper force	N
F_I	Inertia force	N
F_S	Spring force	N
m	Mass	kg
t	Time	s
t_0	Start time	s
T	Periodic duration	s
x	Movement	m
$\dot{x}$	Speed	m/s
$\ddot{x}$	Acceleration	m/s²

REFERENCES

[1] ISO 10326:2002-07 Traktoren und Landmaschinen auf Rädern – Messung der Ganzkörperschwingungen, 1st edition. ISO, Switzerland

[2] Sarami S (2009) Development and Evaluation of a Semi-active Suspension System for Full Suspension Tractors, page 8. TU Berlin, Berlin, Germany

[3] ISO2631-1:1997-05 Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration, 2nd edition. ISO, Switzerland

[4] Bauer W (2008) Hydropneumatische Federungssysteme, 1st edition. Springer, Berlin

[5] Findeisen D (2006) Ölhydraulik – Handbuch für die hydrostatische Leistungsübertragung in der Fluidtechnik, 5th edition. Springer, Berlin

[6] Bauer W (2014) 9th International Fluid Power Conference. Aachen, March 24-26, 2014, Aachen, Germany

[7] Antoni S (2022) 13th International Fluid Power Conference. Aachen, June 13-15, 2022, Aachen, Germany

[8] Hauck M (2001) Geregelte Dämpfung für Traktor-Fahrersitze, page 57-59. TU Berlin, Berlin, Germany

ENERGY EFFICIENT EXCAVATOR FUNCTIONS BASED ON ELECTRO-HYDRAULIC VARIABLE-SPEED DRIVE NETWORK

peer reviewed

Lasse Schmidt[1*], Mikkel van Binsbergen-Galán[1], Reiner Knöll[2], Moritz Riedmann[2], Bruno Schneider[2], Edwin Heemskerk[2]

[1] *AAU Energy, Aalborg University, Pontoppidanstraede 111, 9220 Aalborg, Denmark*
[2] *Bosch Rexroth AG, Zum Eisengießer 1, 97816 Lohr am Main, Germany*

* Corresponding author: Tel.: +45 2232-2622; E-mail address: lsc@energy.aau.dk

ABSTRACT

Electrification of mobile working machines is subject to increasing focus in both industry and academia. At this stage, focus has been the replacement of conventional internal combustion engines with cable or battery fed electric motors driving the main pump(s), and the replacement of rotary functions with electro-mechanical drive solutions. However, the linear functions remain controlled by hydraulic control valves resulting in substantial throttle losses, which in turn necessitates large battery sizes and/or low machine uptimes. Alternatively, the valve-controlled hydraulic cylinders may be replaced with electro-mechanical solutions in applications with limited forces, whereas heavy duty working machines such as medium/large excavators may benefit from standalone electro-hydraulic primary controlled drives, i.e., variable-speed standalone drives. The use of such solutions will substantially increase efficiency due to the absent/limited throttle control and the ability to share power through the electric supply/DC-bus. A main drawback is that each axis needs to be designed to meet both the maximum force and maximum speed, hence in the case of using single motor standalone drives, each motor and associated inverter needs to be designed to meet both the maximum force and maximum speed, potentially rendering these somewhat large. Alternatively, dual motor standalone drives can be applied, enabling power distribution via more motors. However, the use of numerous motors requires more extensive system integration and potentially large motor power installations considering industrially available non-specialized components. This paper presents a novel so-called electro-hydraulic variable-speed drive network, applied for actuation of three linear functions of an excavator implement. Cylinder chamber short-circuiting's and electro-hydraulic variable-speed units constitute a drive network allowing both electric and hydraulic power sharing. The drive network is realized with Bosch Rexroth A2 displacement units and eLION electric motors as its core components. Results demonstrate that the proposed drive network is realizable with similar energy efficiency as a standalone dual motor electro-hydraulic drive solution, but with less motor power and with fewer motors, displacement units and integration effort, rendering this a more sustainable and cost-efficient solution. Finally, it is shown that the proposed drive network is superior in terms of installed displacement, electric motor power and energy efficiency, compared to a separate metering valve drive supplied by a battery fed electro-hydraulic pump.

Keywords: Electro-hydraulic Drive Networks, Energy Efficiency, Excavator Implement Drives

1. INTRODUCTION

In the endeavours to improve the energy efficiency of hydraulic working machines, the current state of technology is the replacement of the conventional internal combustion engines with batteries, electric motors and inverter drives. However, inefficient valve controlled hydraulic systems remains

the current standard resulting in low machine uptimes or large battery capacity requirements[1]. To negate these unfortunate features, obvious drive technologies to consider are electro-mechanical or electro-hydraulic drive solutions. The former has already been proposed for small and medium sized machines whereas the latter may be a feasible alternative in larger machines. Electro-hydraulic drive research and developments have historically been focused on standalone drives for actuation of (mainly) differential cylinders and, a fairly large number of such drives have been introduced in literature [1-5]. Even though many hydraulic systems contain more than a single cylinder, limited attention has so far been given to dedicated multi-axis electro-hydraulic drives. This field has, however, begun to emerge in the last few years ranging from solutions combining displacement units in variable-speed and/or displacement-controlled cylinders with valve-controlled cylinders in a mix [6,7] to more disruptive drive design approaches such as the so-called HHEA [8,9] and so-called electro-hydraulic variable-speed drive networks [10-12]. The latter approach generally contains a tremendous amount of possible drive architectures, with this number increasing exponentially with the number of cylinders (or motors) in a system. Electro-hydraulic variable-speed drive networks do basically not contain any control valves to realize the drive functionality and may be realized with short-circuited chambers and with the only active components being variable-speed displacement units in network configurations. The absence of control valves, consequently the absence of conceptual losses, renders such drive networks highly efficient, whereas the networked drive system and chamber short-circuits may allow for substantially reduced realistic power installations compared to other electro-hydraulic (and electro-mechanical) drive solutions.

This paper presents a novel electro-hydraulic variable-speed drive network (EDN) for actuation of the main linear functions of an excavator implement. The EDN is put in the context of a separate metering valve drive solution (SMV) supplied by a variable-speed pump feasible for electrified machines and a drive system based on dual displacement unit standalone electro-hydraulic drives (DEH's), also with no conceptual losses. Based on measured digging cycles from a representative 17-19t wheeled excavator, the basic drive requirements are established, and the main components subsequently determined with a focus on Bosch Rexroth A2 displacement units and Bosch Rexroth eLION motors. Hereafter, component loss models based on experimental measurements are included and the resulting power consumption and loss distribution is elaborated. Results imply that the EDN energy efficiency is superior to the SMV and comparable to the DEH, but with lower installed displacement and electric motor power in comparison.

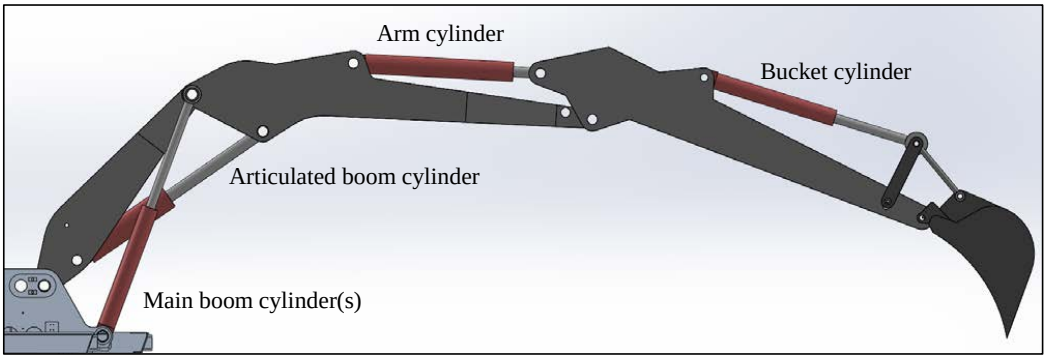


Figure 1: Illustration of excavator implement, emphasizing the cylinders in consideration.

[1] As experienced by the authors at the fair BAUMA in Munich, Germany in October 2022.

The study presented in the following is restricted to the main cylinders of the implement of a 17-19t wheeled excavator, with these being the two main boom cylinders (excluding the articulated boom cylinder), the arm cylinder and the bucket cylinder as illustrated in Figure 1.

An excavator, especially a wheeled excavator, is used for a variety of tasks which, for an industrially applicable design, should all be considered. In some of these tasks, like levelling, the controllability is the main focus and in other the energy efficiency is most important. The latter is true for digging, a predominant excavator task. For the sake of simplification only eight measured digging cycles are considered here, containing piston velocities and forces as depicted on Figure 2.

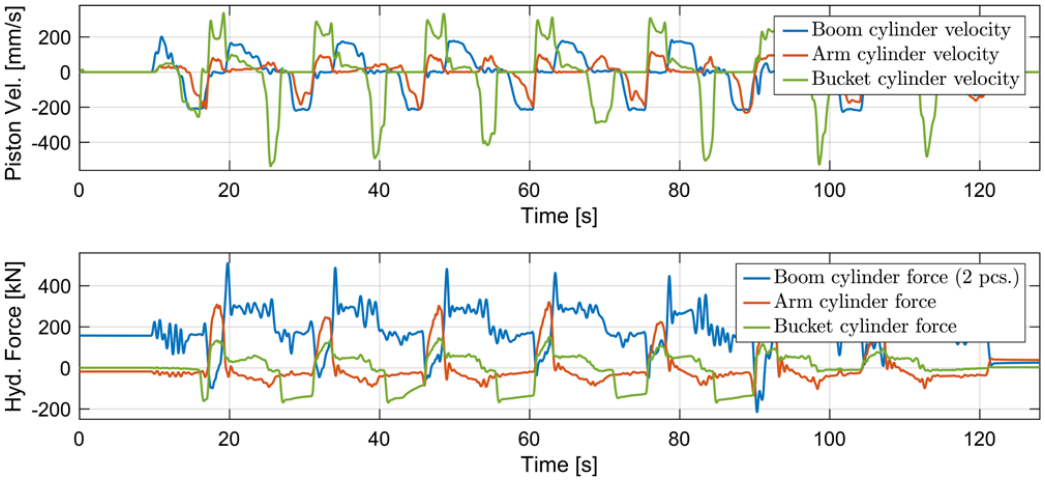


Figure 2: Load cycles for implement cylinder piston velocities and hydraulic pressure forces.

2. ELECTRO-HYDRAULIC DRIVE NETWORK FOR EXCAVATOR IMPLEMENT

Based on the chosen measurement data (Figure 2) the proposed electro-hydraulic drive network (EDN) considered for actuation of the excavator implement cylinders is depicted in Figure 3. This drive network is a so-called minimal realizable EDN, meaning that it is realized with the minimal number of displacement units allowing the control of cylinder motions and the system pressure level. Hence, with three linear functions (noting that there are two parallel boom cylinders), there are four motor inputs and four outputs to control.

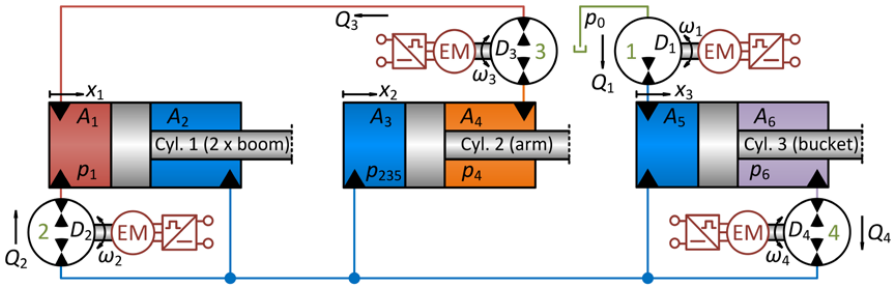


Figure 3: Proposed electro-hydraulic drive network (EDN) for excavator implements.

The proposed EDN encompasses two chamber short-circuiting's being the short-circuit of the boom rod side with arm as well as the bucket piston side chambers. The consequences of these short-circuiting's are identical pressures in the connected chambers but also that hydraulic fluid can flow unrestricted between the connected chambers. Hence, during simultaneous cylinder motions, power in terms of flow under pressure can be transmitted nearly loss free between the cylinders. The network of variable-speed displacement units (VsD's) allows to control the fluid exchange between the control volumes of the drive system. However, due to the way these are interconnected, the ability to control the drive system relies on a combined effort of all VsD's.

2.1. Component Requirements

Considering the EDN schematic of Figure 3, the flow continuity equations expressed as the pressure dynamics are given by Eqs. (1), (2), assuming identical fluid bulk modulii β.

$$\dot{p}_1 = \frac{\beta}{V_1}(Q_1 + Q_3 - A_1\dot{x}_1) \quad , \quad \dot{p}_4 = \frac{\beta}{V_4}(A_4\dot{x}_2 - Q_3) \quad , \quad \dot{p}_6 = \frac{\beta}{V_6}(A_6\dot{x}_3 - Q_4) \tag{1}$$

$$\dot{p}_{235} = \frac{\beta}{V_{235}}(Q_1 - Q_2 + Q_4 + A_2\dot{x}_1 - A_3\dot{x}_2 - A_5\dot{x}_3) \tag{2}$$

Solving Eqs. (1), (2) for Q_1, Q_2, Q_3, Q_4 under steady state conditions, Eqs. (3), (4) are obtained.

$$Q_1 = (A_1 - A_2)\dot{x}_1 + (A_3 - A_4)\dot{x}_2 + (A_5 - A_6)\dot{x}_3 \tag{3}$$

$$Q_2 = A_1\dot{x}_1 - A_4\dot{x}_2 \quad , \quad Q_3 = A_4\dot{x}_2 \quad , \quad Q_4 = A_6\dot{x}_3 \tag{4}$$

The displacement unit pressure differences for the four units are given by Eqs. (5), (6).

$$\Delta p_1 = p_{235} - p_0 \quad , \quad \Delta p_2 = p_1 - p_{235} \tag{5}$$

$$\Delta p_3 = p_1 - p_4 \quad , \quad \Delta p_4 = p_{235} - p_6 \tag{6}$$

Combining Eqs. (3)-(6) with the load cycles of Figure 2, the maximum and minimum required displacement unit flows and pressure differences appear as tabularized in Table 1.

Table 1: Max. and min. requirements for VsD flows and pressure differences of the EDN.

	Unit 1	Unit 2	Unit 3	Unit 4
Max. flow:	111 [l/min]	198 [l/min]	50 [l/min]	98 [l/min]
Min. flow:	-196 [l/min]	-227 [l/min]	-101 [l/min]	-155 [l/min]
Max. Δp:	273 [bar]	266 [bar]	217 [bar]	107 [bar]
Min. Δp:	20 [bar]	-240 [bar]	-346 [bar]	-367 [bar]

The requirements of Table 1 will be used for component sizing in Section 4.

3. BENCHMARK DRIVE SYSTEMS FOR EXCAVATOR IMPLEMENT

The two benchmark drive systems are chosen from their suitability in electrified machinery applications, their energy efficiency perspectives and ability to control piston motions as well as the lower system pressure. These benchmark drive systems are described in the following.

3.1. Benchmark Drive System # 1 & Main Component Requirements

The first benchmark drive system relies on separate metering control as depicted in Figure 4 and is

here denoted SMV. Besides being a separate metering valve control drive system, this also includes flow regenerative valves on the boom and arm functions, and it is supplied by an electro-hydraulic variable-speed pump unit realizable with electric load-sensing function.

The SMV valve flows are given by Eqs. (9)-(11) where $p_{lim} = 1$ [bar] and $\dot{x}_{lim} = 10$ [mm/s], with the pump flow given by Eq. (12).

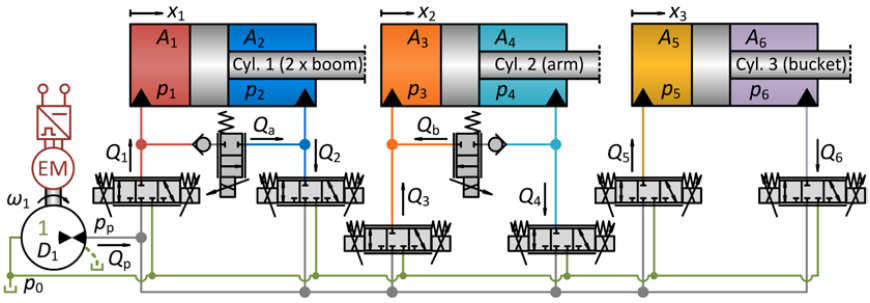


Figure 4: Separate metering valve drive system with electro-hydraulic supply (SMV).

$$Q_1 = A_1\dot{x}_1 + Q_a, Q_2 = A_2\dot{x}_1 + Q_a, Q_3 = A_3\dot{x}_2 - Q_b, Q_4 = A_4\dot{x}_2 - Q_b, Q_5 = A_5\dot{x}_3, Q_6 = A_6\dot{x}_3 \quad (9)$$

$$Q_a = \begin{cases} -A_2(\dot{x}_1 + \dot{x}_{lim}) & \text{for} \quad \dot{x}_1 < -\dot{x}_{lim}, \ p_1 - p_2 > p_{lim} \\ 0 & \text{else} \end{cases} \quad (10)$$

$$Q_b = \begin{cases} A_4(\dot{x}_2 - \dot{x}_{lim}) & \text{for} \quad \dot{x}_2 > \dot{x}_{lim}, \ p_4 - p_3 > p_{lim} \\ 0 & \text{else} \end{cases} \quad (11)$$

$$Q_p = \bar{Q}_1 + \bar{Q}_2 + \bar{Q}_3 + \bar{Q}_4 + \bar{Q}_5 + \bar{Q}_6 \quad (12)$$

$$\bar{Q}_1 = \begin{cases} Q_1 & \text{for} \quad Q_1 > 0 \\ 0 & \text{for} \quad Q_1 < 0 \end{cases}, \ \bar{Q}_5 = \begin{cases} Q_5 & \text{for} \quad Q_5 > 0 \\ 0 & \text{for} \quad Q_5 < 0 \end{cases}, \ \bar{Q}_3 = \begin{cases} Q_3 & \text{for} \quad Q_3 > 0 \\ 0 & \text{for,} \quad Q_3 < 0 \end{cases} \quad (13)$$

$$\bar{Q}_2 = \begin{cases} 0 & \text{for} \quad Q_2 > 0 \\ -Q_2 & \text{for} \quad Q_2 < 0 \end{cases}, \ \bar{Q}_4 = \begin{cases} 0 & \text{for} \quad Q_4 > 0 \\ -Q_4 & \text{for} \quad Q_4 < 0 \end{cases}, \ \bar{Q}_6 = \begin{cases} 0 & \text{for} \quad Q_6 > 0 \\ -Q_6 & \text{for} \quad Q_6 < 0 \end{cases} \quad (14)$$

The pump outlet pressure is controlled via electric load sensing, adjusted according to Eq. (15), where p_{po} = 7 [bar] is the pressure overhead allowing for valve control under relevant loads.

$$p_p = \max(p_1, p_2, p_3, p_4, p_5, p_6) + p_{po} \quad (15)$$

From Eqs. (9)-(15) and the load cycles in Figure 2, the ideal maximum and minimum pump flow and pump pressure difference appear as outlined in Table 2.

Table 2: Max. and min. requirements for the VsD flow and pressure difference of the SMV.

	Unit 1
Max. flow:	250 [l/min]
Min. flow:	0 [l/min]
Max. Δp:	394 [bar]
Min. Δp:	46 [bar]

3.2. Benchmark Drive System # 2 & Main Component Requirements

The second benchmark drive system is based on dual pump electro-hydraulic standalone drives (DEH's) and is depicted in Figure 5. The DEH do not contain any conceptual losses, rendering this one of the most efficient variable-speed electro-hydraulic drives introduced in literature. This drive system can share the electric supply across all VsD's and has the ability to control the lower chamber pressure level on all cylinders.

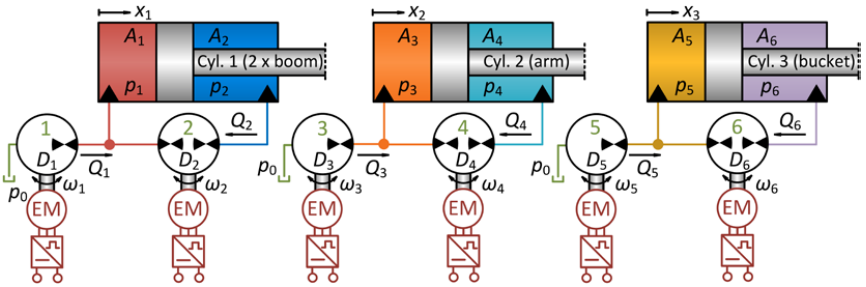


Figure 5: Drive system based on dual pump electro-hydraulic standalone drives (DEH's).

The displacement unit flows and pressure differences for the DEH's are given by Eqs. (16)-(19).

$$Q_1 = (A_1 - A_2)\dot{x}_1\,,\quad Q_2 = A_2\dot{x}_1\,,\quad Q_3 = (A_3 - A_4)\dot{x}_2 \tag{16}$$

$$Q_4 = A_4\dot{x}_2\,,\quad Q_5 = (A_5 - A_6)\dot{x}_3\,,\quad Q_6 = A_6\dot{x}_3 \tag{17}$$

$$\Delta p_1 = p_1 - p_0\,,\quad \Delta p_2 = p_1 - p_2\,,\quad \Delta p_3 = p_3 - p_0 \tag{18}$$

$$\Delta p_4 = p_3 - p_4\,,\quad \Delta p_5 = p_5 - p_0\,,\quad \Delta p_6 = p_5 - p_6 \tag{19}$$

Combining Eqs. (16)-(19) with the load cycles of Figure 2, the ideal required maximum and minimum displacement unit flows and pressure differences appear as shown in Table 3.

Table 3: Max. and min. requirements for the VsD flows and pressure differences of the DEH.

	Q1	Q2	Q3	Q4	Q5	Q6
Max. flow:	94 [l/min]	117 [l/min]	35 [l/min]	50 [l/min]	78 [l/min]	98 [l/min]
Min. flow:	-105 [l/min]	-131 [l/min]	-70 [l/min]	-101 [l/min]	-124 [l/min]	-155 [l/min]
Max. Δp:	306 [bar]	286 [bar]	273 [bar]	253 [bar]	184 [bar]	164 [bar]
Min. Δp:	20 [bar]	-240 [bar]	20 [bar]	-198 [bar]	20 [bar]	-367 [bar]

4. IMPLEMENT DRIVE COMPONENT SIZINGS

The sizing of components is conducted according to the maximum and minimum displacement unit flows and pressure differences obtained for the EDN, SMV and DEV.

4.1. Sizing of Displacement Units

In all cases, the displacement units considered are Bosch Rexroth A2 bent axis pumps or motors[2], due to their proven application history. It is assumed that the lower pressure level of any control volume is controlled to 20 [bar]. For any displacement unit not connected to a vented fluid reservoir, the suction pressure conditions are not violated for any of the considered units and may therefore operate in all four quadrants. Hence, displacement units not connected to fluid reservoirs are chosen as A2FM hydraulic motors. Displacement units that are connected to a fluid reservoir are subject to suction restrictions and may be operated in two quadrants. Hence, A2FO hydraulic pumps are more feasible than A2FM's. Also, cavitation may occur when fluid is pumped from a vented reservoir to a pressurized control volume. Consequently, A2FO sizing's are based on the maximum positive flow requirements. Furthermore, flow losses are not included in the sizing, and hence the sizing is based on nominal shaft speeds but with an upper limit of 6000 [rpm].

The corresponding choices of displacement units chosen for the three drive systems are tabularized in Table 4 along with actual geometric displacements to be installed. Here, the EDN and DEH are subject to the lowest installed displacement, in both cases are more than 9 [%] lower than the SMV.

Table 4: Summary of chosen displacement units and total displacements for the three drive systems.

Drive	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6	Total Displ.
SMV	A2FO 180	-	-	-	-	-	180,0 [ccm]
DEH	A2FO 45	A2FM 23	A2FO 12	A2FM 23	A2FO 32	A2FM 28	163,5 [ccm]
EDN	A2FO 56	A2FM 56	A2FM 23	A2FM 28	-	-	163,7 [ccm]

4.2. Sizing of Electric Motors

Based on the displacement unit sizes of Table 4, the ideal shaft speed and torque requirements for the electric motors can be established for the three drives as specified in Table 5.

Table 5: Required motor speeds and torques for the three drive systems.

SMV	Unit 1					
$\max(\lvert n \rvert)$:	1384 [rpm]	-	-	-	-	-
$\max(\lvert \tau \rvert)$:	1129 [Nm]	-	-	-	-	-
DEH	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6
$\max(\lvert n \rvert)$:	2294 [rpm]	5709 [rpm]	5822 [rpm]	4398 [rpm]	3867 [rpm]	5505 [rpm]
$\max(\lvert \tau \rvert)$:	222 [Nm]	104 [Nm]	53 [Nm]	93 [Nm]	94 [Nm]	165 [Nm]
EDN	Unit 1	Unit 2	Unit 3	Unit 4		
$\max(\lvert n \rvert)$:	3486 [rpm]	4010 [rpm]	4398 [rpm]	5505 [rpm]	-	-
$\max(\lvert \tau \rvert)$:	244 [Nm]	240 [Nm]	127 [Nm]	165 [Nm]	-	-

The flow and torque losses are not included in the following electric motor sizing examples and hence these are sized considering the motor S2 torques as the maximum design torques and the shaft speeds well below the maximum motor speeds. Considering the eLION EMS1 motor portfolio[3] and the specifications in Table 5, the motor choices are presented in Table 6 along with the resulting total rated motor power.

[2] Based on data sheets "RE 91071/2021-05-17, Bosch Rexroth AG" and "RE 91401/06.2012, Bosch Rexroth AG".
[3] RE98771/2022 04 26, Bosch Rexroth AG.

The deviations in the total rated (motor) power results especially from either relatively high required torques or the use of a relatively large number of VsD's. In case of a high maximum motor torque, the associated rated power tends to be correspondingly large as well, attributed the associated speed range. Similarly, the use of a relatively large number of VsD's tends to result in a relatively large total installed torque overhead, with this increasing with number of units used. A reasonable level of installed torque and power is therefore achieved with few units and with a reasonable ratio between required speed and torque for each unit.

Table 6: Choice of eLION motors EMS1 along with the resulting total rated (motor) power to be installed.

Drive	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6	Total rated power
SMV	EMS1-16	-	-	-	-	-	74 [kW]
DEH	EMS1-10	EMS1-10	EMS1-10	EMS1-10	EMS1-10	EMS1-10	54 [kW]
EDN	EMS1-10	EMS1-10	EMS1-10	EMS1-10	-	-	44 [kW]

In summary, the total rated motor power for the EDN is 18,5 [%] lower than the DEH and 40,5 [%] lower than the SMV, owed to the fewer units applied compared to the DEH and, indirectly, to the substantially lower torque requirement compared to the SMV.

4.3. Tank Sizing Considerations

Besides the component sizes, hence the drive sustainability in terms of material usage, another important aspect is the sizing of the fluid reservoir/tank. The EDN, as well as the DEH, do not rely of throttle control, and hence the fluid degasification and fluid cooling requirements are substantially reduced, as the "throttling" taking place in these drive systems are associated only with cross-port leakage and drain flows, whereas it for the SMV is the full pump flow that is throttled. A rule of thumb suggests that the tank size should be chosen as three times the average (throttled) flow to allow for heat dissipation and degasification[4]. Using this rule, the theoretical SMV tank size may be obtained as Eq. (20). Applying the same rule for the EDN (and the DEH), and accounting for the cylinder piston volumes, the theoretical tank size is obtained as Eq. (21).

$$V_{tank,SMV} = 3\,\mathrm{mean}(Q_p) \tag{20}$$

$$V_{tank,EDN} = 3\,\mathrm{mean}(\Sigma Q_D + \Sigma|Q_L|) + (A_1 - A_2)x_{1\max} + (A_3 - A_4)x_{2\max} + (A_5 - A_6)x_{3\max} \tag{21}$$

Using the loss models described in Section 5, the resulting theoretical tank volumes are outlined in Table 7, suggesting significant reductions in the EDN and DEH tank volumes compared to the SMV, hence also in the required amount of fluid in the system. Conventional tank designs used in mobile machinery are often optimized in various ways, and hence the tank volume could be as low as half of the volume proposed for the SMV in Table 7. However, even in such a case, the proposed theoretical volumes for the EDN and DEH are still reduced by more than 80 [%].

Table 7: Ideal fluid reservoir/tank sizes and relative reduction compared to SMV.

	EDN	DEH	SMV
Ideal fluid reservoir/tank size	30,7 [l]	29,4 [l]	330,4 [l]
Reduction relative to SMV	↓ 90,7 [%]	↓ 91,1 [%]	-

[4] https://www.powermotiontech.com/hydraulics/reservoirs-accessories/article/21882642/fundamentals-of-hydraulic-reservoirs (assessed 26. October 2023).

5. ENERGY EFFICIENCIES & LOSS DISTRIBUTION

Having decided on the main component sizes of the proposed EDN and the benchmark drive systems SMV and DEH, their energy efficiency and loss distributions are considered in the following as well as the loss models applied for the analyses.

5.1. Loss Models Used in Case Studies

The loss model used as reference for the displacement units is based on measured loss maps for the A2FMM 32 [ccm/rev], assuming this also representative for the A2FO units. The loss map used as reference for the eLION EMS1's is based on measured losses of an EMS1-20 type motor. In all cases, the measured data has been smoothened to mitigate the impact of outlier measurement points, and the hydraulic losses extrapolated for pressure differences above 350 [bar], hence representing approximate loss measures. The approximate loss diagrams are shown in Figure 6.

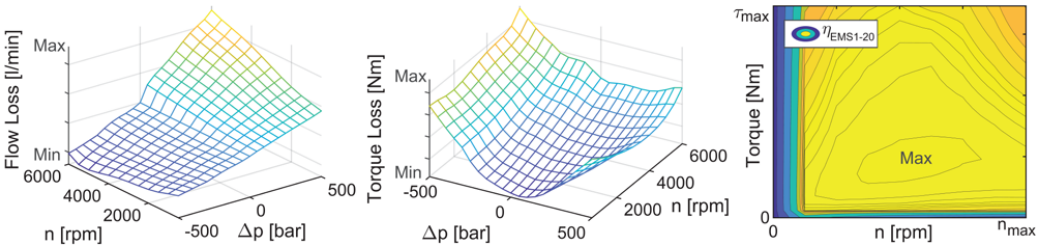


Figure 6: Approximate loss diagrams. **Left plot**: 2Q flow loss map for A2FMM 32. **Center plot**: 2Q torque loss map for A2FMM 32. **Right plot**: Power loss map for eLION EMS1-20 component series motor.

Regarding the A2FM, the measured flow losses account for both cross-port leakage flow and drain flows. It is assumed that the flow losses are evenly distributed in the reference measurements and that these can be scaled to relevant displacement unit sizes using scaling laws [2,10]. The scaling of the EMS1 assumes that the efficiency map is invariant with respect to the motor size i.e., that any motor has the same efficiency map as the reference EMS1-20, with the axes scaled to the max. torque and speed of the motor in consideration. Furthermore, for the EMS1 motors it is assumed that the efficiency map is valid for all four quadrants. The total inverter losses $P_{inv,loss}$ and DC-bus losses $P_{dc,loss}$ are estimated according to Eqs. (22), where R_i is the electrical resistance of the i^{th} inverter, R_{dc} is the electrical DC-bus resistance, $U_{dc} = 700$ [Vdc] the DC-bus voltage, $I_{inv,nom,i}$ the nominal current of the i^{th} inverter, and $I_{dc,nom} = \sum_{i=1}^{n} I_{inv,nom,i}$ the nominal DC-bus current. Finally, the inverter and DC-bus efficiencies are assumed to be $\eta_{inv} = 0{,}98$ [-], $\eta_{dc} = 0{,}98$ [-] and the battery efficiency assumed to be $\eta_{bat} = 0{,}90$ [-].

$$P_{inv,loss} = \sum_{i=1}^{n} R_i \frac{P_{m,i}^2}{U_{dc}^2}, P_{dc,loss} = R_{dc} \frac{P_{inv,loss}^2}{U_{dc}^2}, R_i = \left(\frac{1}{\eta_{inv}} - 1\right) \frac{U_{dc}}{I_{inv,nom}}, R_{dc} = \left(\frac{1}{\eta_{dc}} - 1\right) \frac{U_{dc}}{I_{dc,nom}} \quad (22)$$

5.2. Loss Distribution & Power Consumption

The overall average loss distribution of the three drive systems, using the load cycles of Figure 2, are tabularized in Table 8 along with the average and peak power consumption. The deviation in losses and power consumption of the EDN and DEH relative to the SMV drive are shown in Table 9. From this it is evident that significant reductions in all losses are achieved with DEH and EDN drives compared to the SMV drive, except for the displacement unit friction measures which are higher,

attributed the use of more displacement units and the associated specific loads.

Table 8: Average losses and average and peak power consumption.

Drive Type	Avg. Hydraulic Losses	Avg. Friction Losses	Avg. Electric Losses	Avg. Battery Losses	Peak Power Consumption	Avg. Power Consumption
SMV	15,6 [kW]	1,3 [kW]	2,8 [kW]	4,0 [kW]	116,7 [kW]	39,8 [kW]
DEH	1,2 [kW]	2,8 [kW]	2,1 [kW]	2,4 [kW]	83,7 [kW]	18,3 [kW]
EDN	1,2 [kW]	2,5 [kW]	2,0 [kW]	2,5 [kW]	84,6 [kW]	18,1 [kW]

Furthermore, it is notable that the average energy consumption by the DEH and EDN drives are reduced by more than 54 [%] compared to the SMV. If the load cycles considered here are assumed generally representative for the implement function, then for an 8 hour shift the SMV drive would require a 318 [kWh] battery, whereas the DEH and EDN drives would only require 146 [kWh] and 145 [kWh] batteries, respectively.

Table 9: Relative differences in losses and power consumption of the DEH and EDN compared to SMV.

Drive Type	Avg. Hydraulic Losses	Avg. Friction Losses	Avg. Electric Losses	Avg. Battery Losses	Peak Power Consumption	Avg. Power Consumption
DEH	↓ 92,3 %	↑ 115,4 %	↓ 25,0 %	↓ 40,0 %	↓ 28,3 %	↓ 54,0 %
EDN	↓ 92,3 %	↑ 92,3 %	↓ 28,6 %	↓ 37,5 %	↓ 27,5 %	↓ 54,5 %

Finally, it is notable that the drive energy efficiencies, in terms of the ratio between the average battery power and the average hydraulic piston power (≈ 9,9 [kW]), are 24,8 [%], 54,0 [%] and 54,5 [%], for the SMV, DEH and EDN drive systems, respectively. Hence, the DEH and EDN drive systems are subject to energy efficiency increases of ≈118 [%] and ≈120 [%], respectively.

6. CONCLUSION

A novel electro-hydraulic variable-speed drive network is proposed, specifically intended for actuation of the cylinder functions in excavator implements, namely the main boom cylinders, the arm and bucket cylinders. The proposed drive network encompasses the short circuiting of the boom cylinder rod sides and the arm and bucket cylinder piston sides. Consequently, the three-cylinder system contains four effective control volumes, which are actuated by four variable-speed displacement units in a network configuration. The proposed drive network is placed in context of a separate metering drive system supplied by a variable-speed fixed displacement pump as well as a dual pump electro-hydraulic standalone drive system. Main component requirements are established, and components subsequently chosen from the Bosch Rexroth A2 hydraulic pump and motor series, and the eLION electric motor series. Steady state models relying on measured component losses are established, and the power consumption and loss distribution evaluated. The resulting key measures in terms of total energy efficiency, power consumption, total installed displacement and electric motor power are evaluated, and the relative differences for the proposed drive network compared to the benchmark drive systems are outlined in Table 10.

Table 10: Relative differences in key figures of EDN compared to the DEH and SMV drive systems.

Drive Type	Avg. Total Energy Efficiency	Avg. Power Consumption	Total Installed Displacement	Total Installed Motor Power
EDN vs. DEH	↑ 0,9 %	↓ 1,1 %	↑ 0,1 %	↓ 9,1 %
EDN vs. SMV	↑ 119,2 %	↓ 54,5 %	↓ 11,6 %	↓ 40,5 %

It is found that the proposed drive network may be realized with substantially less installed electric motor power and similar energy efficiency compared to the dual pump electro-hydraulic standalone benchmark drive system. Furthermore, compared to the separate metering valve drive system the proposed drive network may be realized with less installed displacement, and significantly less installed electric motor power, and with more than twice the energy efficiency. Also, the proposed drive network may be realized with an ideal tank volume reduction of more than 80 [%] compared to the separate metering valve drive system due to the substantially reduced fluid degasification requirements. The results emphasize the potential significance of electro-hydraulic drive networks in the ongoing electrification trends and efficiency improvements of hydraulic working machines.

REFERENCES

[1] Jensen KJ, Ebbesen MK, Hansen MR (2021) Novel Concept for Electro-Hydrostatic Actuators for Motion Control of Hydraulic Manipulators. Energies 14:6566. https://doi.org/10.3390/en14206566

[2] Padovani D, Ketelsen S, Hagen D, Schmidt L (2019) A Self-Contained Electro-Hydraulic Cylinder with Passive Load-Holding Capability. Energies 12:292. https://doi.org/10.3390/en12020292

[3] Casoli P, Scolari F, Minav T, Rundo M (2020) Comparative Energy Analysis of a Load Sensing System and a Zonal Hydraulics for a 9-Tonne Excavator. Actuators 9:39. https://doi.org/10.3390/act9020039

[4] Schmidt, L, Ketelsen, S, Padovani, D, & Mortensen, KA (2019) Improving the Efficiency and Dynamic Properties of a Flow Control Unit in a Self-Locking Compact Electro-Hydraulic Cylinder Drive. Proceedings of the ASME/BATH Symposium on Fluid Power and Motion Control. https://doi.org/10.1115/FPMC2019-1671

[5] Ketelsen S, Michel S, Andersen TO, et al (2021) Thermo-Hydraulic Modelling and Experimental Validation of an Electro-Hydraulic Compact Drive. Energies 14:2375. https://doi.org/10.3390/en14092375

[6] Fassbender D, Zakharov V, Minav T (2021) Utilization of electric prime movers in hydraulic heavy-duty-mobile-machine implement systems. Automation in Construction 132. https://doi.org/10.1016/j.autcon.2021.103964

[7] Fassbender D, Brach C, Minav T (2023) Experimental Investigations of Partially Valve-, Partially Displacement-Controlled Electrified Telehandler Implements. Actuators 12:50. https://doi.org/10.3390/act12020050

[8] Siefert J, Li PY (2022) Optimal Control of the Energy-Saving Hybrid Hydraulic-Electric Architecture (HHEA) for Off-Highway Mobile Machines. IEEE Transactions on Control Systems Technology 30:5. https://doi.org/10.1109/tcst.2021.3131435

[9] Khandekar AA, Siefert J, Li PY (2021) Co-Design of a Fully Electric Hybrid Hydraulic-Electric Architecture (FE-HHEA) for Off-Road Mobile Machines. Proceedings of the American Control Conference. https://doi.org/10.23919/acc50511.2021.9482615

[10] Schmidt L, Hansen KV (2022) Electro-Hydraulic Variable-Speed Drive Networks—Idea, Perspectives, and Energy Saving Potentials. Energies 15:1228. https://doi.org/10.3390/en15031228

[11] Schmidt, L, Ketelsen, S, Hansen, KV (2022) Perspectives on Component Downsizing in Electro-Hydraulic Variable-Speed Drive Networks. Proceedings of the BATH/ASME Symposium on Fluid Power and Motion Control. https://doi.org/10.1115/FPMC2022-89547

[12] Schmidt, L, Ketelsen, S, & Hansen, KV (2022) State Decoupling & Stability Considerations in

Electro-Hydraulic Variable-Speed Drive Networks. Proceedings of the BATH/ASME 2022 Symposium on Fluid Power and Motion Control. https://doi.org/10.1115/FPMC2022-89548

A COMPREHENSIVE REVIEW OF ELECTRONICALLY CONTROLLED IMPLEMENT ARCHITECTURES FOR MOBILE MACHINERY USING SECONDARY CONTROL

Edwin Heemskerk[1]*, Michael Brand[2]

[1] *Bosch Rexroth AG, Partensteinerstraße 23, 97276 Lohr am Main, Deutschland*
[2] *Bosch Rexroth AG, Glockeraustraße 4, 89275 Elchingen, Deutschland*

* Corresponding author: Tel.: +49 9352 183127; E-mail address: edwin.heemskerk@boschrexroth.de

ABSTRACT

Transmission architectures for drive and implement functions in mobile machinery are currently diverting. On the one hand new market requirements, as alternative energy sources, impact the system design, and so do direct or indirect market drivers as efficiency, controllability, predictability, and driver comfort. On the other hand, the ability to electronically control components in alternative ways enable new architectures as well.

The persons in charge of the machine design must find the balance between considering the new demands and relying on known building blocks to reduce risks and safeguard valuable resources. To help the decision making, an overview of new and promising architectures is presented, utilizing secondary control approaches for drive and work functions. These architectures target to recover kinetic energy, reduce throttling losses, and operate components in their sweet spot.

Decisive criteria for the shown architectures are their maturity, degree of fulfilment of the market demands, minimized risk criteria via proven sub-components and feasibility of handling the transmission variants for varying market demands.

Keywords: Electro-hydraulics, secondary control, mobile machinery, energy recovery

1. INTRODUCTION

Our increasing demands to protect our environment in a sustainable way, to optimize material resources like energy, to reduce CO_2 and to ensure human well-being are well known. This results in more diverse machine architectures for segmented market demands e.g., with various power sources like battery driven machines, e-fuel base driven machines and conventional *ICE*s and well adapted transmission architectures depending on load cycles and machine usage. At the same time machine manufacturers must consider decreasing engineering resources out of cost and availability reasons. This rising gap in increasing demands versus available resources needs to be solved via standardization and usage of overarching building block technologies in general and also for transmissions architectures and their required components.

Beneficially, technological improvements allow new approaches in mechatronic optimizations of components. Sensors get more cost effective, smaller, and more reliable. In the field of hydraulic transmissions, this allows to transit from hydromechanics actuation and controls to more precise and more adjustable electronic controls.

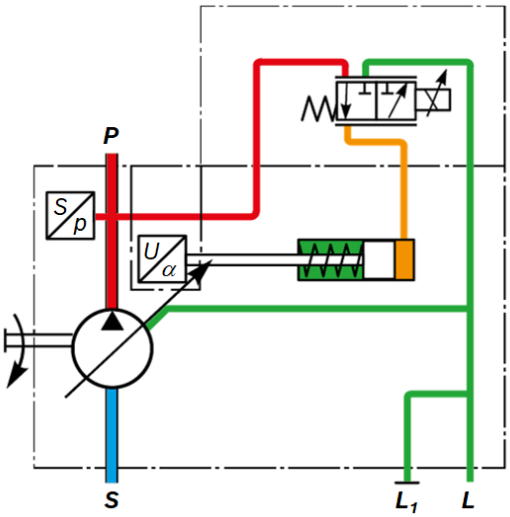


Figure 1: Electrohydraulic control schematic of an electronified open circuit pump

A good example for this transition is the electronified open circuit (*eOC*) pump control for open circuit hydraulic pumps. Its control concept is shown in **Figure 1**. An electro-proportional valve controls the swivel speed of the pump. The pump is equipped with a pressure and a displacement sensor. The valve is controlled via software. Available commands are pressure control, displacement control, torque control and the gradient of these command values as seen in **Figure 2**. Gradients of actual values can be limited.

The concept allows to operate the pump much closer to the physical limits since tolerances are compensated and static or dynamic critical working points can be avoided. Load cycles can be respected. With this mechatronic optimization, the least hydro-mechanic effort reaches the maximum performance.

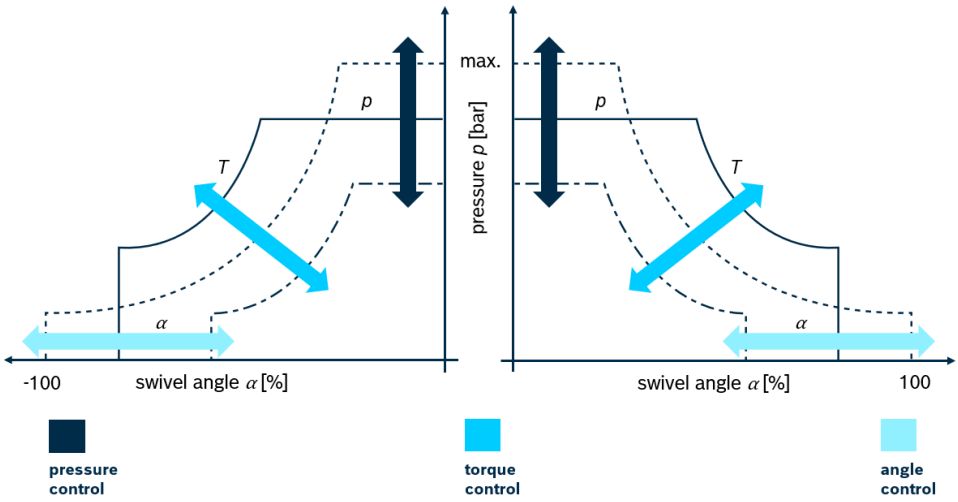


Figure 2: Control options for electronified open circuit pumps

This control approach allows open circuit pumps to be operated in two quadrants as shown in **Figure 2**. This means that without relevant additional hardware effort the pump can supply oil in pumping mode and can receive oil as a hydro motor and transfer this hydraulic energy into mechanical energy at the input shaft of the pump [1].

A second technology trend fosters this mechatronic approach. It becomes easier, more known and accepted and technology wise supported to allocate the component-based software functionality on central control units in an efficient way. This gives the machine manufacturer a maximal degree of freedom to master and adjust the component function to any needs of the transmission architecture with maximum efficiency and in a cost attractive way.

The mechatronic optimization principle is valid also for closed circuit pumps and valves. As important is that the best technology is used for a respective component feature and to follow a pragmatic and not a dogmatic approach. Engineering disciplines for developing components get more heterogenous and competitiveness will be strongly influenced by the ability to adapt to these opportunities.

On a component function level for open or closed circuit hydraulic pumps the potential can be summarized in the following way:

- Pressure, displacement and torque control and limitations
- Dynamic and gradient control for command and disturbance values
- Component protection
- Adjustability on various functional demands resulting on low variance and stock demands
- Self-calibration to reduce or avoid EOL or service efforts
- Speaking failures
- Reduced function ability in failure situations

Focus of the paper are in the following new transmission architectures that are now possible based on this new component control approach.

2. NEW POSSIBILITIES WHEN USING A MECHATRONICALLY OPTIMISED PUMP

As explained in the introduction, a mechatronically optimised pump has benefits in dynamics, flexibility, efficiency, and performance. Depending on the application different aspects can be interesting. In this paper the excavator is used as an example to showcase different possibilities.

2.1. Power Control

In many mobile machines the hydraulic power needs to be limited, as the hydraulic system can take more power than the primary drive train (i.e., a diesel engine or an electric motor) is able to supply. Typically, this is done be a two-spring regulator. These springs as chosen in such a way, that the deviation from the hyperbolic power curve is minimized in the main operation points of the machine. Unfortunately, the maximum power of an engine is afflicted with tolerances. These are caused i.e., by the manufacturing, the fuel used or the operating altitude of the machine and the actual operating speed. To prevent the engine from stalling under all conditions, the hydraulic power limitation is set lower than the lowest expected engine power, as shown in **Figure 3**, while also considering the spring tolerances.
As the electronified open circuit pump has a pressure and swivel angle sensor (see **Figure 1**), the system always knows the load it is putting on the engine. The maximum allowed load can be dynamically communicated to the pump via *CAN* bus message and the electronic pump control ensures that this limit is respected accordingly. In addition, if other power consumers (like the air

conditioning or an auxiliary pump) are switched on, or if the engine speed is operated at other (lower) engine speeds, this can be considered in the pump control.

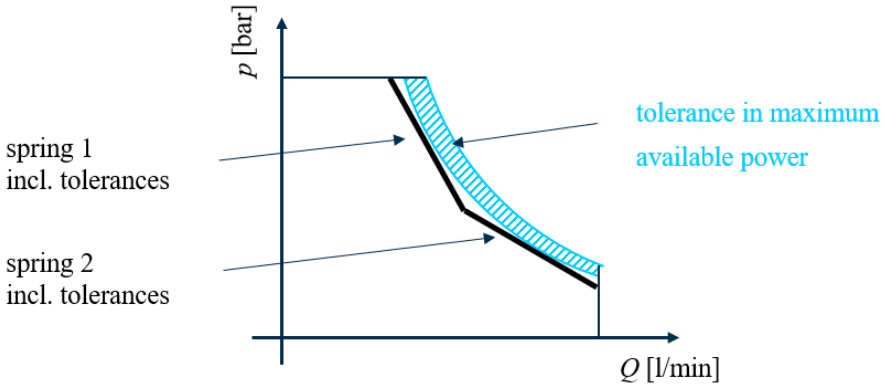


Figure 3: Power control of a hydraulic pump

In case the engine is capable to communicate its rotational speed and remaining torque capacities to the pump controller, the hydraulic output power of the pump can be directly aligned with the load limit of the engine. This integrated load limiting control allows to eliminate any safety margin in the torque setting of the pump and leads to up to 10% increased hydraulic output power. In comparison to conventional load limiting control solutions based on electronic displacement control, there is no negative impact of hysteresis.

2.2. Crawler machines with 1-loop systems

Many mini and midi excavators in the world are equipped with a 1-loop flow sharing system (see **Figure 4** for a crawler excavator setup). This is a well-suited system architecture to achieve a good compromise between performance, controllability, and cost, especially for machines that work with different attachments.

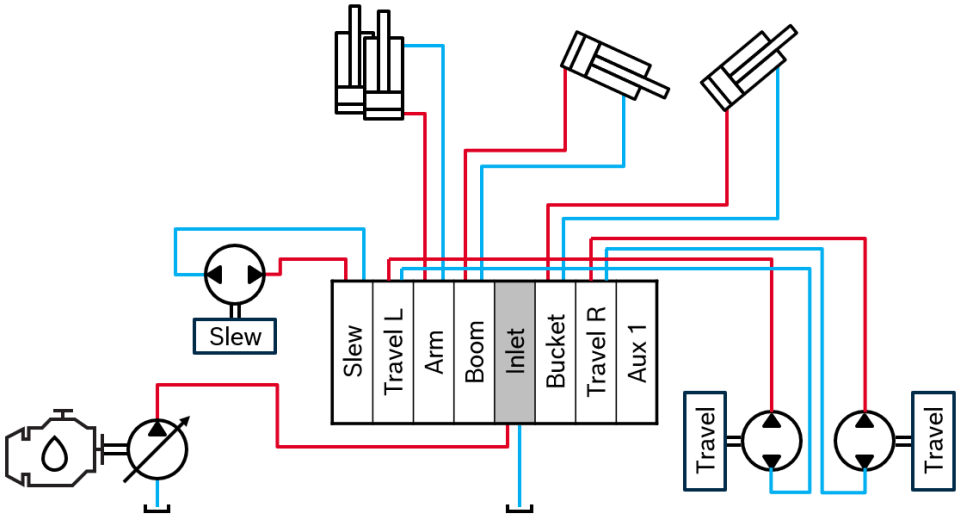


Figure 4: Typical configuration of the hydraulic system in a mini excavator

The biggest challenge for a 1-loop system in a crawler excavator is a light turn while driving at full speed. **Figure 5** shows the transition from driving a straight line to a slow turn.

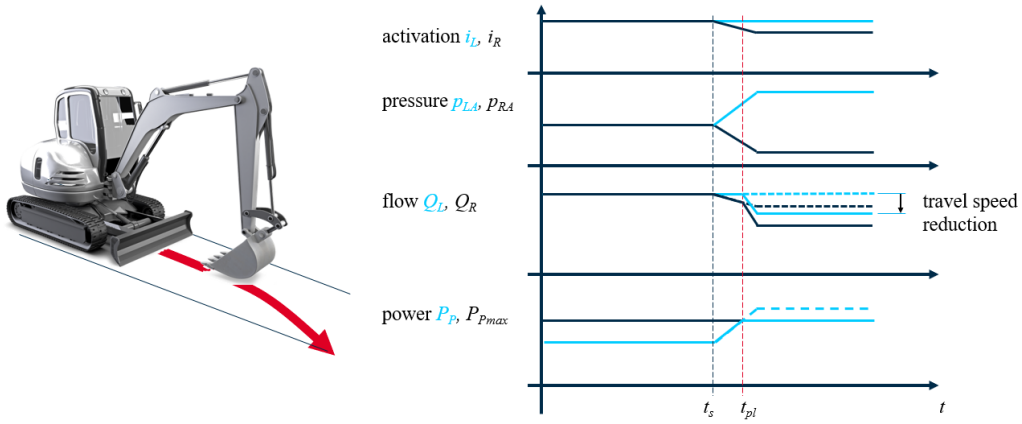


Figure 5: Slow turn with a crawler excavator in a 1-loop hydraulic system

As soon as the right track is slowed down enough for the machine to turn slightly to the right, the left track meter-in pressure rises significantly, while the meter-in pressure of the right track goes down. This is a hydraulic 1-loop system, so also the pump pressure rises accordingly. As it is only a slow turn, the total flow request (the sum of Q_L and Q_R) for the pump is only slightly lower compared to a straight travel. In consequence the hydraulic power is significantly increased. If this increased power need is bigger than the total available power (at t_{pl} in **Figure 5**), the hydraulic power limitation will reduce the total pump flow, slowing down the machine driving speed (the dashed line shows the flows Q_L and Q_R in case of no power limitation).

When replacing the normal fixed displacement travel motors with electronified open circuit pump/motors and operating these secondary controlled, as shown in **Figure 6**, this behaviour can now be significantly improved.

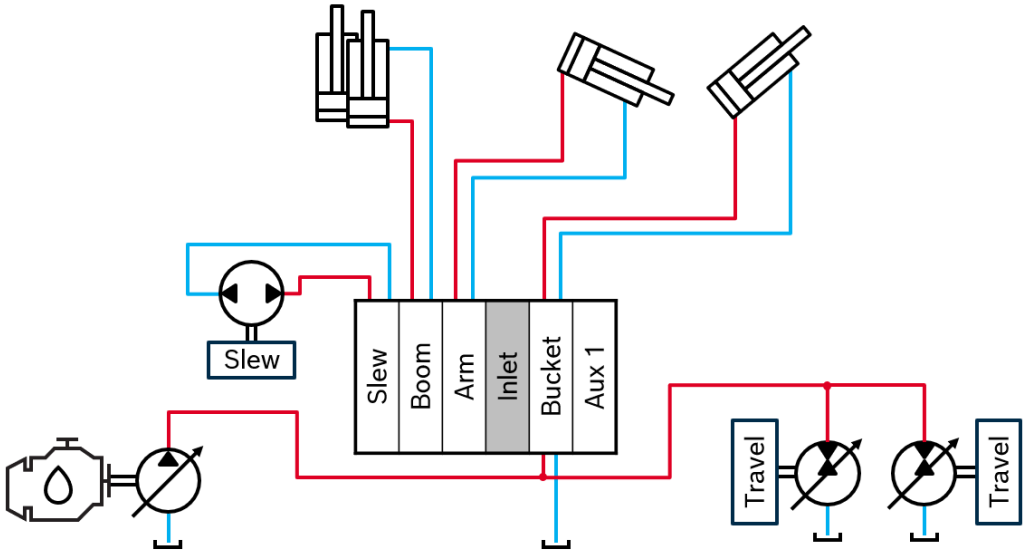


Figure 6: Modified hydraulic system in a mini excavator using secondary controlled travel motors

An operator demand for turning results in different torque demands to both motors. These are realized with an adequate pressure level of the pump together with the calculated displacements on the left and right hydro motor. No throttle losses arise in the main control valve (*MCV*) for both functions and no throttle losses arise for the hydro motor side with the lower torque demand. The flow demand for the same machine speed is lower than compared with the conventional architecture. This way of actuation is only possible based on the pump and hydro motor function described in chapter 1. On machine level it results in faster and more efficient drive manoeuvres.

2.3. Slew control

Like the travel motors, also the slew motor, typically a fixed displacement motor, can be replaced by a secondary controlled pump/motor. Here the target is not an improved machine behaviour, but energy saving.

Also, in this case the electronified open circuit pump/motor is directly connected to the main pump line, with no valve in between, eliminating throttling losses in parallel movement. The swash angle of the slew pump/motor is dependent of the current pump pressure, both when accelerating and de- accelerating. When slowing down the slew, the slew pump/motor operates as a pump, supplying oil to all other functions used in parallel, reducing the energy consumption of the hydraulic system [2]. The degee of freedom of the slew pump/motor displacement allows to adapt the pressure level to other functions and therefore reduces throttle losses.

2.4. Boom energy recovery

Apart from reducing the throttling losses in a system, many machine also have the possibility to recover energy while lowering a load (i.e., reach stackers or industry excavators) or just by lowering its own machine structure. The force needed to lift the empty boom of an excavator is typically about 20 to 50 % of the force needed to lift a full bucket. So just lowering the empty boom already offers potential to recover energy. **Figure 7** shows a hydraulic 2-loop system for a battery driven wheeled excavator that allows for boom energy recovery. The main difference to a normal 2-loop system, that is used in bigger excavators, is the additional power regenerative boom module (*PRB*) and the fact that the left pump is an electronified open circuit pump, as described in chapter one.

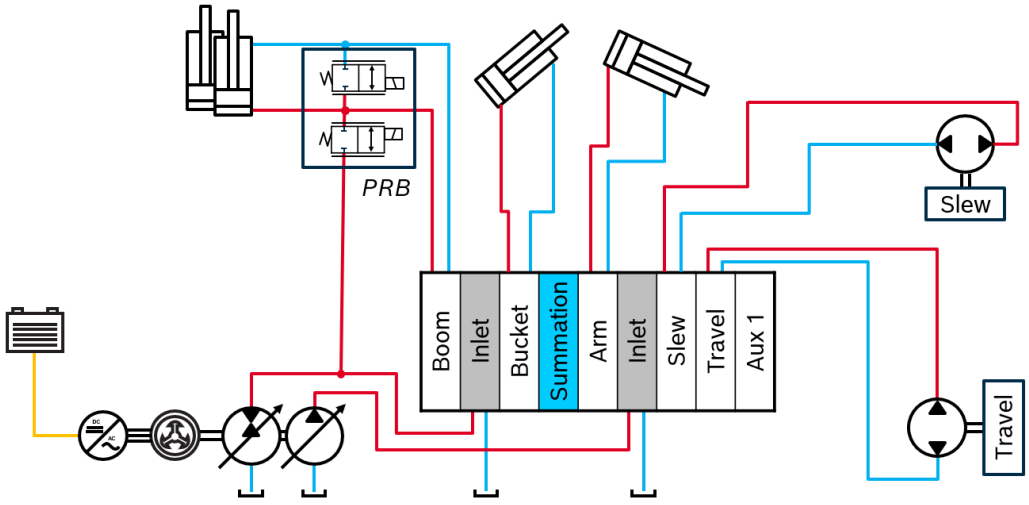


Figure 7: 2-loop hydraulic system with boom energy recovery

With the top valve of the power regenerative boom module, the head and rod side of the boom cylinder is connected. The lower valve connects the boom head side with the electronified open circuit pump/motor. With this hydraulic setup we can increase the pressure and reduce the flow from the boom cylinder to the pump/motor, working as a motor in this condition. So even when lowering the boom at full speed it can recover nearly all the boom down energy, without the need to install a bigger pump, which would reduce the energy efficiency in all operating situations. Also, this function requires the component features introduced in Chapter 1.

2.5. Active ride control

Most mobile machines, like wheeled loaders and wheeled excavators are not equipped with suspensions at the wheel axles. Instead, passive (*PRC*) and active ride control (*ARC*) modules are used [3]. The *PRC* typically use an accumulator and a hydraulic restriction to dampen the oscillation. This lost free solution is optimized for one load conditions and that has reduced functionality in other load conditions. *ARC* systems can adapt to different load conditions, but typically at the cost of energy.

The hydraulic setup shown in **Figure 7** also allows for an *ARC* without additional components. While driving the *PRB* valve connects the boom to the electronified open circuit pump/motor. This pump/motor is operated in pressure control, either taking out oil of the boom or adding oil, to keep the load pressure constant [4]. When working as a pump, it uses energy from the drive train, when working as a motor, it recovers energy, resulting in a near loss free active ride control (*ARC*).

3. CONCLUSION AND OUTLOOK

The mobile machinery market demands for variable architectures to support the different needs of the customers, as alternative energy sources, efficiency, controllability, predictability, and driver comfort. The machine *OEM*s must find the balance between considering the new demands and relying on known building blocks to reduce risks and safeguard valuable resources. New electronified components offer possibilities to handle the transmission variants for the different market demands.

This paper presents different examples of hydraulic architectures, that utilize the dynamics and the control flexibility of an electronified open circuit pump/motor, that improve the performance and/or the energy efficiency of an excavator. These architectures are examples to stimulate machine *OEM*s to find new individual solutions to improve their specific machines based on enhanced component functionalities.

NOMENCLATURE

α	Swivel angle percentage of the max. pump swivel angle	%
A	Area	m^2
ARC	Active ride control	
CAN	Controller area network	
CO_2	Carbon dioxide	
eOC	Electronified open circuit	
EOL	End of line	
ICE	Internal combustion engine	
i_L, i_R	Joystick activation of the left and right track	%
L, L1	Leakage port on the pump	
MCV	Main control valve	
OEM	Original equipment manufacturer	
p	Pressure	bar

P	Main port of the pump	
p_{LA}, p_{RA}	Motor inlet pressure of left and right travel motor	bar
P_P	Hydraulic power at the pump	kW
P_{Pmax}	Maximum available power at the pump	kW
PRB	Power regenerative boom	
PRC	Passive ride control	
Q	Flow	l/min
Q_L, Q_R	Flow to the left and right travel motor	l/min
S	Suction port on the pump	
S	Sent interface (electronic protocol)	
t	Time	s
t_s	Starting time of the steering movement	s
t_{pl}	Time the power limitation limits the flow	s
T	Torque	Nm
U	Voltage	V

REFERENCES

[1] Tetik S (2022) Secondary Control Based Hydraulic Architectures For Mobile Applications. The 13th International Fluid Power Conference, 13.IFK, March 21-23, 2022, Aachen, Germany

[2] Pettersson K (2009) Secondary control in construction machinery – design and evaulation of an excavator swing drive. 11th Scandinavian International Conference on Fluid Power, SICFP'09, June 2-4, 2009, Linköping, Sweden

[3] Madau R (2019) Active Ride Control for Construction Machines Based on Pressure Feedback. ASME/BATH 2019 Symposium on Fluid Power and Motion Control, October 7–9, 2019, Longboat Key, USA

[4] Cardoso R (2022) Active Ride Control Based on an Over-Center Pump. 2022 Maha Fluid Power Conference, May 17-19, 2022, Purdue, USA

WEIGHT SAVING POTENTIALS OF PRESSURE INCREASE IN CYLINDERS OF MOBILE MACHINES KINEMATICS

peer reviewed

Tobias Radermacher[1]*

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*

* Corresponding author: Tel.: +49 351 463-33618; E-mail address: tobias.radermacher@tu-dresden.de

ABSTRACT

Many mobile machines operate primarily through the use of boom structures, which account for a large share of energy consumption. The purpose of this work is to investigate if and to what extent the selection of the system pressure level in hydraulically driven boom structures can contribute to the reduction of the moving weight and thus to the reduction of the overall energy demand. The paper focusses on the choice of optimal system-pressure with regard to the cylinder-weight. A calculation method is derived that allows for the analytical calculation of cylinder weight with regard to system pressure to profit from the lightermost cylinders taking into account all sizes of differential cylinders, material properties , safety-factors and cylinder-ratio. It is shown that by specifying the load force, cylinder material, safety factor and area ratio as input parameters the system pressure level that allows the lightest cylinders can be determined. Application of the method shows that there are ideal pressure levels for hydraulic cylinders that are almost independent of the cylinder force and therefore machine size. The results show that a proper pairing of cylinder size and system pressure may have big weight-saving potentials. Taking into account high-strength steel (tensile strength $\sigma = 500\,N/mm^2$) enormeous weight savings may be achieved with rising system pressures.

Keywords: mobile hydraulics, mobile machinery, energy efficiency, payload, boom design, hydraulic optimization, lightweight cylinder

1. INTRODUCTION

Energy optimization in the sector of mobile machines is an important issue. The population of 8..40 ton excacators worldwide is about 1.000.000 and needs more energy than the annually production of offshore windmills worldwide [1]. Together with wheel-loaders excavators have a share of around 60% in mobile construction machineries fuel need [2]. Especially the hydraulic-driven boom structures have –due to a bad payload-to-weight ratio- a big share of the energy consumption in these machines. The improvement of energy efficiency in the operation of boom structures is the subject of numerous research projects which may be classified to optimisations of the diesel motor [3, 4], the task-related design of the kinematics, structural optimisations regarding material and construction [5, 6] the use of compensators [7–10] and the hydraulic drivetrain [11–15]. Within the hydraulic drivetrain special attention is being paid to the efficient operation of the pump(s), the cirquit design and the proper adjustment of the cylinder size to the kinematics. For the optimisation of an energy-efficient system all aspects have to be considered because they are strongly linked.

In mobile machines there is a trend for lifting system pressures that can exemplarily be seen in hydraulic traction drives which today work with pressures up to 450 bar [16, 17]. Along with the idea

to increase power density the question arises which are the most important restrictions and if lifting system pressure would be a good solution for making boom structures more powerful. In addition to the approaches stated above an improvement of the payload-weight ratio may be achieved by a proper choice of the system pressure in system design, which may allow for lighter cylinders.

After presenting the basic assumptions and design goals in chapter 2 the mathematical solution is presented in chapter 3. The possibilities and limitations of the methodology are discussed in chapter 4 before drawing a conclusion and giving an outlook.

2. BASIC ASSUMPTIONS AND DESIGN GOALS

One key benefit oy hydraulic actuation systems is the high power density which may further be improved with higher system pressures. Concerning the increase of system pressure the question arises if there are additional effects that may be lead to an increased energy efficiency of the overall system. Reducing the cylinder weight can be advantageous for:

- Reduction of energy consumption (less mass to be lifted + machine weight)
- Reduction of counterweights (cost + machine weight)
- Reduction of machine weight → more power in same weight-class possible

In the very price-sensitive market of mobile machines, these advantages can be exploited if production cost does not rise significantly, so changes in design and construction of hydraulic cylinders are not in the scope of this work.

The guiding question for this work is whether there is a lightest standard cylinder for accomplishing a work task and how this can be calculated depending on the sytem pressure level or, viceversa how system pressure may be appropriately chosen for the lightest (standard) cylinders possible.

2.1. Cylinder design

The design of boom structures is strongly driven by a high payload-to-weight ratio and the restrictions given by the working tasks (working area, movements, sensibility). The basic idea is to reduce the cylinder weight by optimizing the cylinder in stroke, size and rated pressure taking into account the applicable directives[18]. Even though many restrictions like the tasks to fulfil, the available design space, the strength of the mechanical linkage and many more limit the design space for cylinder size, the knowledge of a minimal weight-design may add a design-factor that could be considered in kinematics development.

Cylinders of different sizes may be compared by their work. Cylinders of equal work fulfill the equation:

$$F_c \cdot l_c = const. \tag{1}$$

with the cylinder force F_c and the length l_c. In other words, a cylinder that is mounted closer to the joint has a shorter stroke but needs to supply more force, what in general leads to a bigger diameter. We will refer to different cylinder sizes fulfilling equation (1) as “equal-work cylinders” according to **Figure 1** (left).

Assuming that cylinder parts that are stroke-independent like hinges have a defined length, the cylinder can be calculated according to its position. The weight of the cylinder, which is the variable to be optimized here, is determined by the design of its components. In the following, we calculate performance-equivalent cylinders in order to find the lightest cylinder depending on system pressure and material for the given application. We will only consider parallel displacements of the cylinders according to Figure 1 (left) here. A change of the cylinder ratio α, which is defined as the ratio

between rod and piston area, or any further degrees of freedom in the cylinder design is not considered in this work to get comparable results.

3. CALCULATION OF CYLINDER WEIGHT

Single-acting cylinders consist of a piston, a piston rod, a piston tube, as well as a piston head at both ends and mounting elements on the piston rod and piston bottom. These need be dimensioned during the design process. The cylinder weight then is calculated of its elements according to Figure 1 with the material constants $\rho = 7850\ kg/m\text{^}3$ (steel) and $\rho_{oil} = 880\ kg/m^3$(mineral oil).

$$m_c \cong \rho \cdot \pi \cdot (4 \cdot h \cdot r^2 + l \cdot (2 \cdot r \cdot s - s^2) + l \cdot r_s^2\ ) + oil\ fillling \tag{2}$$

with the cylinder mass m_c, and nomenclature according to Figure 1 (right). Hereby it is considered that the cylinder rod and the end caps have the same height h and the stroke of the cylinder can be approximized as $x = l - h$.

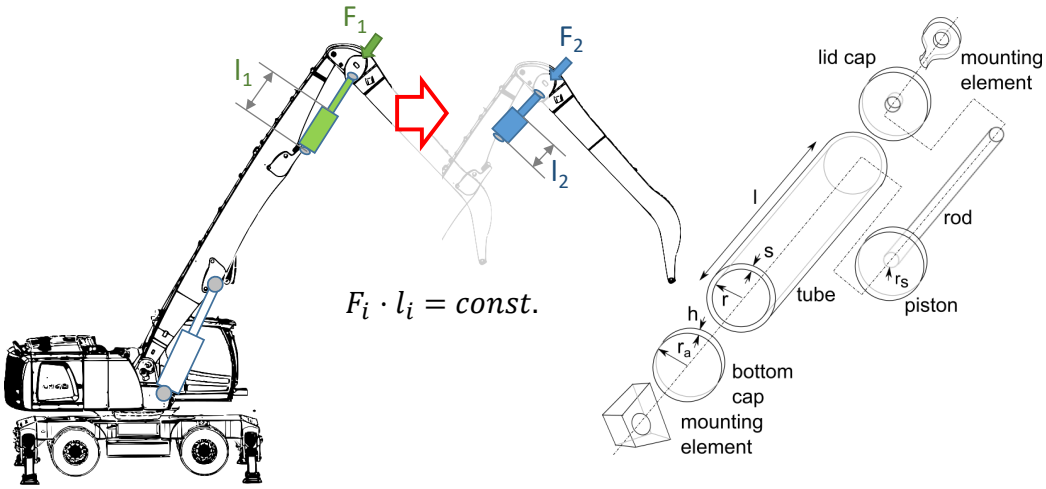


Figure 1: cylinders with $F \cdot l = const$ (left) and cylinder elements (right)

The cylinder rod with radius r_s can be calculated on the basis of the maximum stroke and the acting force F_c on buckling with buckling factor $\frac{l_k}{l} = 0{,}7$ whereas:

$$r_s = \sqrt[4]{\frac{F_c \cdot 4 \cdot \frac{l_k}{l} \cdot sf \cdot l_{0,9}^2}{E \cdot \pi^3}} \tag{3}$$

whereas $l_{0,9}$ is 90% of the rod length and the elastic modulus of the material: E=2.1E10 N/m^2, the security factor which is assumed to be $sf = 1{,}5$ for hydraulic cylinders in mobile applications and $l_k/l = 0{,}7$ for the Euler-buckling factor.

The cylinder inner radius r then is calculated for a given maximum system pressure p and force F_c via:

$$r = \sqrt{\frac{F_c}{p \cdot \pi}} \tag{4}$$

Complying the restriction of a given aspect ratio α there may be the need to choose a bigger radius

for the cylinder according to:

$$r = \sqrt{\frac{\alpha \cdot r_s^2}{\alpha - 1}} \tag{5}$$

Note: Formulae (4) and (5) show that choosing the right cylinder ratio α may reduce the cylinder diameter without any further optimisation.

From the sizing of rod and inner cylinder diameter the whole cylinder can be calculated for all pressure levels. Dimensioning the piston tube according to DIN 2413:202-04 is used if the ratio of outer and inner diameter is $d_a/d_i < 2$. This may not be applied for higher pressures, therefore an unequal distribution of tension along the wall thickness shall be assumed. We assume linear-elasticity and ideal plasticity along the wall. In comparison of the calculation methods: normal stress hypothesis, shear stress hypothesis or shape change energy hypothesis the latter gives nearly exact values for the maximum pressure p_i at the inner radius r_i. In consequence the outer radius r_a of the cylinder will be calculated with formula (6) according to [19]. In comparison with the standard calculation methods for cylinders ($d_a/d_i < 2$) this formula is accurate in pressure ranges above 500 bar.

$$p_i = \frac{R_e}{sf} \cdot \frac{r_a^2 - r^2}{\sqrt{3} \cdot r_a^2} \tag{6}$$

With $\frac{R_e}{sf} \geq \sigma$ The cylinder cap height h follows the maximum tangential tension

$$\sigma_t \leq \sigma_{t,max} = f \cdot p \cdot \frac{r^2}{h^2} \tag{7}$$

with the factor $f = 0.8$ considering the cap as a panel with clamped edge (welded construction). The same height is applied to the piston and the lid with minor error. The attachments at both sides are considered with the weight of one lid.

With these formulae the lowest cylinder mass for equal-work cylinders under the given restrictions of material and security factor can be calculated depending on the system pressure according to (1) with $s = r_a - r$.

4. RESULTS AND DISCUSSION

Relevant changes of cylinder mass result from variation of force, stroke, maximum pressure level and material. Calculation results are visualized in **Figure 2** for the example of cylinder ratio of $\alpha = 2$.

On the left hand side results for cylinders made of standard steel ($\sigma = 235\ N/mm^2$) are shown, on the right hand side cylinders made of steel with high tensile strength ($\sigma = 500\ N/mm^2$). All diagrams show different cylinder configurations: each line represents possible configurations for a kinematic with given work scenario. Along each line the lightest equal-work cylinders for different pressure levels are shown. The lower diagram shows the corresponding lengths.

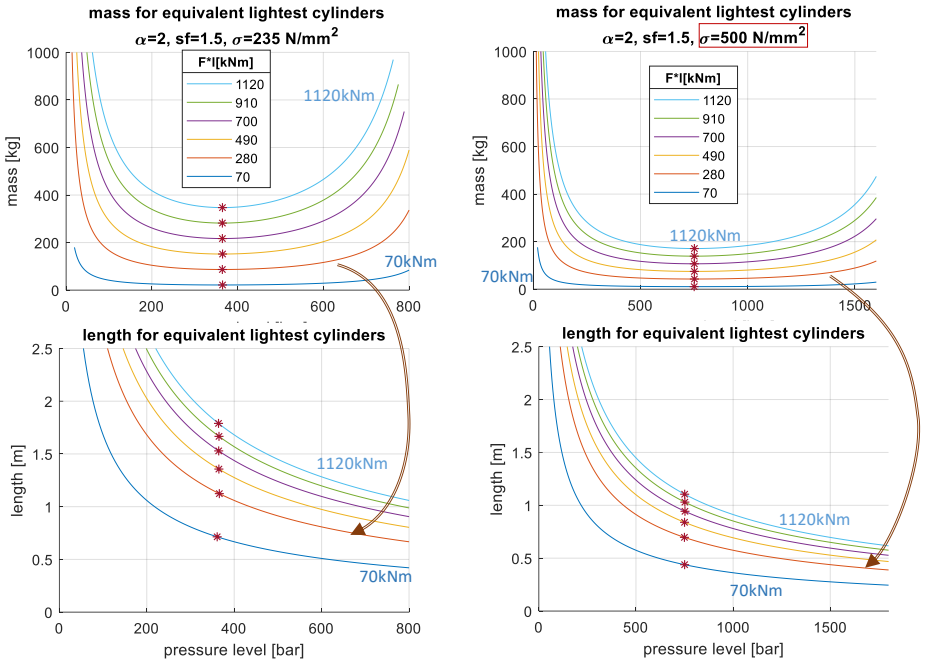


Figure 2: mass and length for lightest equal-work cylinders

The lightest standard industrial steel cylinders with an area ratio $\alpha = 2$ have an ideal pressure level of between 300 and 450 bar with a minimum around 370 bar. The ideal pressure for saving cylinder weight is thereby nearly independent of the machine size. The ideal cylinder length grows with the cylinder power. The calculations show that manufacturing cylinders from high-strength steel halves the weight of a comparable cylinder from standard steel. The main reason for that effect is the shorter stroke. The distribution of weight for equal-work cylinders with 700 kNm (compare Figure 2, purple lines) is shown in **Figure 3**. Weight of piston, caps, attachments and tube grow with rising pressure (even length decreases) whereas weight of rod and oil volume decreases resulting in a defined minimum of weight. As buckling depends on elasticity-module (which is nearly constant for all steel materials) the weight of the piston rod is independent from the strength of the chosen material.

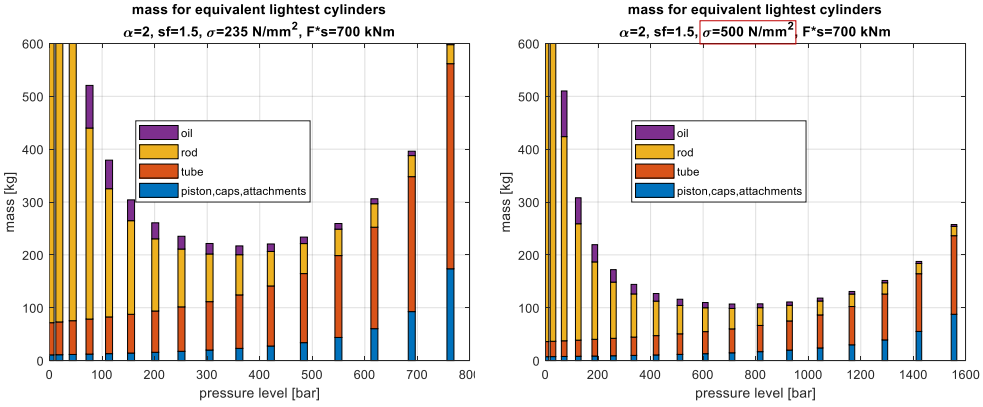


Figure 3: mass distribution for equal-work cylinders with 700 kNm and different material strength

Amongst material and area ratio the minima also depend on the chosen security factor. As a rule of

thumb weight scales proportional with the security factor and optimal pressure halves.

Figure 4 shows the points of minimal mass of equal-work cylinders (red asterisks in Figure 2) for cylinders with varying area ratios α. For standard-steel on the left hand and for high performance steel on the right hand side.

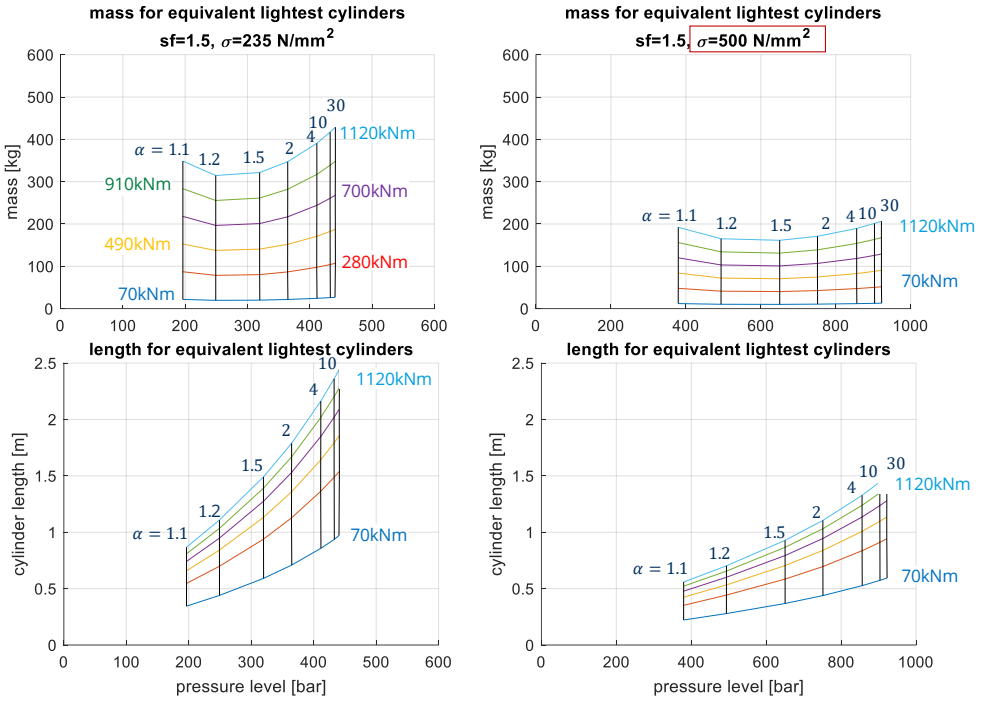


Figure 4: mass and length for lightest equal-work cylinders

The figure shows that weight heavily varies with the strength of the steel but the points of optimal weight stay constant: Cylinders with an area ratio of between $\alpha = 1.2$ and $\alpha = 1.5$ have the best power-to-weight ratio, independent of the material and the cylinder power needed. The smaller the area ratio α is, the lower is the ideal working pressure with respect to cylinder weight. With regard to this outcome the optimal pressure for the lightest cylinders can be plotted over the area ratio.

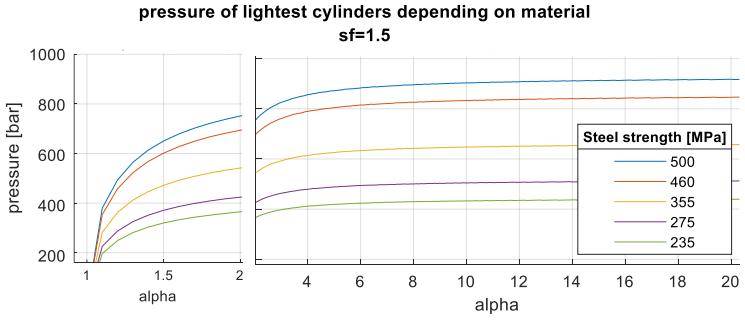


Figure 5: ideal system pressure for minimal-weight cylinders with regard to area ratio α

Figure 5: shows the ideal pressure for the lightest cylinders produced from different materials. As a rule of thumb it can be stated here that the optimal system pressure scales proportional to the strength of the used material.

5. SUMMARY AND CONCLUSION

With respect to hydraulic-driven kinematics the results give insights in how moving mass may be reduced if system pressure and cylinder size are matched to each other. Rising system pressure without using different materials may have the potential to save cylinder weight independent of the machine size in many applications. If high-strength steel is used, enormous weight savings may be realised with rising system pressures. The comparison of different system pressure levels in **Figure 6** exemplarily shows possible weight reductions of about 27% when rising system pressure from 160 bar to 410 bar and twice as large if a high-strength steel is used. The weight-saving potential increases with higher security-factors sf and slightly with bigger area ratios α.

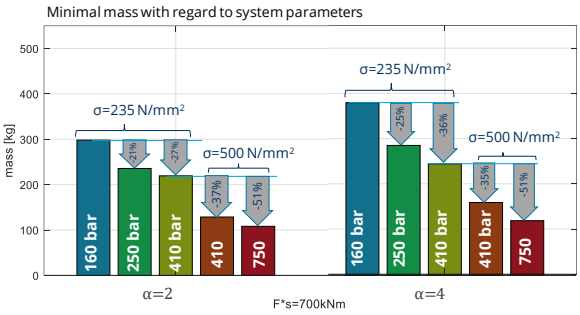


Figure 6: weight savings for cylinders with different materials and pressure levels

The availability of components for increasing the system pressure above 500 bar is currently not (yet) given. However, the investigation shows the potential for significant weight savings. Savings in the weight of the boom are directly reflected in additional handling capacity and or a reduced mass of the counterweight (approximately factor 3). As a result, the power density of the machines may be increased, alternatively machines of the same power get lighter. There might be drawbacks in application: Even if calculations for some practical cases show that simple strengthening of the cylinder mounting points is sufficient, there may be cases where forces become too high. Precision may also be lost if cylinders become very short. Changes in stiffness and natural frequency of high pressure systems should also be further investigated.

ACKNOWLEDGEMENT

The author thanks the Research Association for Construction Machinery of the German Engineering Federation VDMA for its financial support. Special gratitude is expressed to the participating companies and their representatives in the accompanying industrial committee for their advisory and technical support in the IGF project No. 21840 BR/1 “Systemdruckniveau in mobilen Arbeitsmaschinen”.

NOMENCLATURE

σ	Tensile strength	N/m^2
α	cylinder ratio	$-$
F	Force	N
l	length	m
$_{\cdot c}$	cylinder	$-$
ρ	Density	kg/m^3
m	mass	kg

$_{oil}$	Oil	$-$
h	height	m
s	wall thickness	m
r_i	Inner radius	m
r	(outer) radius	m
x	Cylinder stroke	m
r_s	Rod radius	N
sf	Security factor	$-$
l_k/l	critical (buckling) length	$-$
E	Elasticity modulus	N/m^2
p	Pressure	N/m^2
R_e	Yield strength	N/m^2
F	Force	N
f	Form factor	$-$

REFERENCES

1. Caldwell N (2019) Digital displacement Technology. 2019 CCEFP Summit Presentation:

2. Helmus M, Fecke M (2015) Standardisierung definierter Lastzyklen und Messmethoden für die Energieverbrauchsermittlung von Baumaschinen. Bergische Universität Wuppertal, Interdisziplinäres Zentrum III, Wuppertal

3. Schneider M, Koch O, Weber J (2014) Green Wheel Loader - Operating Strategy of an Energy Efficient Hybrid Drive Train. In: SAE 2014 Commercial Vehicle Engineering Congress. Las Vegas, NV, USA, pp 2014-01–2400

4. Jähne H (2013) Struktursystematik und Effizienzpotentiale hydraulischer Fahrantriebe unter Einbeziehung der Applikation. Dissertation, Technische Universität Dresden

5. n.n. N (2010) Leichtbau Baggerstiel, Gebrauchsmusterschrift, DE 20 2010 003 617 U1

6. Hufenbach W, Helms O, Ulbricht A (2010) Development of Lightweight Hydraulic Actuators in Innovative Fibre Composite Design. In: 7th Internationsl Fluid Power Conference. Aachen

7. Mantsinen, nn. In: Mantsinen Hybrilift explained. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjbhay5yfiDAxU1cvEDHbTGCncQwqsBegQIChAF&url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3Doiq_9xjol-k&usg=AOvVaw3T4nHGGNaS-6xVdfLPkIe7&opi=89978449. Accessed 10 Jan 2024

8. Sennebogen.com Sennebogen ERC. https://www.sennebogen.com/technik/energierueckgewinnungssystem-green-hybrid. Accessed 10 Jan 2024

9. Stanger S, Wagner B, Asam D, Lavergne H-P Arbeitsgerät, EU Patent EP 2 148 014 A2

10. Xia L, Quan L, Ge L, et al (2020) Research on efficient driving method of heavy hydraulic excavator boom. In: 12th International Fluid Power Conference. Technische Universität Dresden, pp 481–488

11. Huova M, Tammisto J, Linjama M, Tervonen J (2018) Fuel Efficiency Analysis of Selected Hydraulic Hybrids in a Wheel Loader Application. In: BATH/ASME 2018 Symposium on Fluid

Power and Motion Control. American Society of Mechanical Engineers, Bath, UK, p V001T01A035

12. Ivantysyn R, Weber J (2014) Novel Open Circuit Displacement Control Architecture in Heavy Machinery. In: 8th FPNI Ph.D Symposium on Fluid Power. American Society of Mechanical Engineers, Lappeenranta, Finland, p V001T02A002

13. Heybroek K (2017) On Energy Efficient Mobile Hydraulic Systems: with Focus on Linear Actuation. PhD Thesis, Linköping University Electronic Press

14. Vukovic M, Prof. Murrenhoff, Prof. Geimer (2017) Hydraulic Hybrid Systems for Excavators. Phdthesis, Shaker Verlag

15. Lübbert J, Weber J, Strauch C, et al (2020) Modular independent metering system for mobile applications providing smooth mode transition. In: 12th International Fluid Power Conference. Technische Universität Dresden, pp 239–248

16. Mothes M, Finzel R (2018) Elektronifizierung von Baumaschinen Neue elektrohydraulische Fahrantriebe. In: 7. Fachtagung Baumaschinentechnik 2018. Dresden

17. Heidrich L (038) (2017) Unimog-Geräteträger Leistungssteigerung Hydrostatischer Fahrantrieb UNIMOG durch Anhebung Systemdruck auf 500 bar. In: Fachtagung MOBIMA Karlsruhe 2017. Karlsruhe

18. Druckgeräte-RL 97/23/EG

19. Läpple V (2016) Einführung in die Festigkeitslehre. Springer Fachmedien Wiesbaden, Wiesbaden

PROGRAMMING EXAMPLE

```
%MATLAB® example code for calculation of cylinder mass
% Author Tobias Radermacher, TU Dresden, Chair of Fluid-Mechatronic Systems
% License: GNU GPL
% download: github.com/boing0815
clc,clear, close all
%calculation of mass for F*l=const
sig=235E6; %N/m^2,max tension,reversibel
Re=sig;
sf=1.5;%security factor
alpha=2; %ratio
leg="";
l_light=[];
force=700; %niveau of force for f*l

f=figure(); hold on; grid on;
% calculation of lines with F*l=const
l=[0.1:0.3:1.6];
skal=0.02:0.02:4; %scaling ofo F*l
for h=1:length(l)
    ls=l(h)./skal;
    fs=force.*skal;
    for k=1:length(skal) %for every variant with f*l constant
        sp(k)=pfl(fs(k)*1000,ls(k),alpha,sf); %neccesary pressure level in bar
        m(k)=zylindermasse_von_r(sig,Re,sp(k)*1E5,fs(k)*1000,ls(k),alpha,sf); %calc cyl. mass; change force to N
    end
    [m_min_f_mal_l(h),ind(h)]=min(m);
    p_min_f_mal_l(h)=sp(ind(h));
    l_light(h)=ls(ind(h));
    f_light(h)=fs(ind(h));

    pl=plot(sp,m);
    uistack(pl,"bottom");
    test=convertCharsToStrings(sprintf('%.0f', force*l(h)));
    leg=[leg,test];
end
title(["mass for equivalent lightest cylinders","\alpha=2, sf=1.5, \sigma=235 N/mm^2"]);
```

```
xlabel('pressure level [bar]');
ylabel('mass [kg]');
xlim([0,1800])
ylim([0,600])
%marking min.
plot(p_min_f_mal_l,m_min_f_mal_l,'*')
%scale position for paper size
f.Position(3:4) = [400 300];
%make legend
le=legend(leg(1,2:end));
le.String=flip(le.String); %flip legend
le.Position = [0.4 0.5 0.1 0.2];
saveas(f,"lightest cylinders_mass.fig");
saveas(f,"lightest cylinders_mass.emf");

function p=pfl(F,l,alpha, sf)
    %calc of p for defined area ratio
    E=2.1E10; %E-module in N/m^2;
    %go for it
    % min. radius der of rod against buckling
    % buckling length 90%
    r_s=(F*sf*2*(0.9*l)^2/(E*pi^3))^(1/4);
    r_i=sqrt(alpha/(alpha-1))*r_s; %area ratio
    p=F/(r_i^2*pi);
    p=p/1E5; %in bar
end

function [m] = zylindermasse_von_r(sig,Re,p,F,l,alpha,sf)
%change mass to 1E-12, if useless geometry
% sig, Re in N/m^2, p in N/m^2, F in N, l in m

%input:
    %mat. properties
    rho=7850;%kg/m^3, dens. steel
    rho_oel=880; %kg/m^3
    E=2.1E10; %E-module in N/m^2
%calculation
    %rod radius
    r_s=(F*sf*2*(0.9*l)^2/(E*pi^3))^(1/4);
    %piston radius
    r_i=sqrt(F/(p*pi));
    %wall thickness (fat, plasticity=yes)
    r_a=r_i/sqrt(1-p*sf*sqrt(3)/sig);
    s=r_a-r_i;
    %wall thickness (thin, elastic):
    %s=r_i*p/(sig-p)
%cap heigth
    h=(r_i+s)*sqrt(0.8*p/(sig/sf));
%finally: the result:
    %mass:
    m_platten=rho*pi*(4*h*r_a^2);
    m_rohr=rho*pi*(2*r_a*s-s^2)*l;
    m_stange=rho*pi*l*r_s^2;
    mz=m_platten+m_rohr+m_stange;
    m_oel=(l-h)*r_i^2*pi*rho_oel;
    m=mz+m_oel;
    if s>0.1 || s<=0 %delete useless wall values
        m = NaN;
    end
end
```

Chapter 20

Digitalization

PRECISE HYDROSTATIC CYLINDER DRIVE WITH INCREASED PRESSURE LEVEL FOR INDUSTRIAL APPLICATIONS

Ralf Bonefeld[1]*, Thomas Gellner[1], Klaus Bauer[2]

[1] *Bosch Rexroth AG, Zum Eisengiesser 1, 97816 Lohr am Main*
[2] *Werkzeugmaschinenfabrik WALDRICH COBURG GmbH, Hahnweg 116, 96450 Coburg*

* Corresponding author: Tel.: +49 9352 18-2309; E-mail address: ralf.bonefeld@boschrexroth.de

ABSTRACT

Primary controlled hydrostatic drives, mainly based on speed-variable pumps, are increasingly used in various industrial applications, efficiency and controllability being the main advantages. Still, it is an open question, if the simple substitution of a proportional valve by a speed-variable pump leads in general to a qualitatively comparable system behaviour, or if additional measures are necessary.

Especially, for the 4-quadrant control of single rod cylinders, the use of a servo-driven double-pump is advantageous in terms of efficiency. Nevertheless, the sizing of the pump-unit with respect to the cylinder geometry is a non-trivial task. In the following it will be shown that it is essential for a servo-quality performance, to firstly ensure a flow balance matching the cylinder geometry and secondly to achieve an increased pressure level in order to take over load forces without exceeding pressure limits. To this aim, a control of the sum pressure is known to be appropriate [1]. Introducing another controllable motor-pump-unit allows for coping with both aspects at a time with a purely primary controlled architecture. Within the following a generic systematics of corresponding setups is shown and the example of a cylinder drive for a large machine table proves the servo quality of the solution, making it suitable even for precision machining.

Keywords: Primary control, efficiency, single rod cylinder, 4-quadrant-operation, hydraulic circuit, sum pressure, load stiffness, machine tool

1. INTRODUCTION

Hydraulic cylinders are in general considered as a competitive, efficient and robust component for linear motion in a wide range of actuator applications. Exemplary numbers of the global market size in 2023 show, that cylinders contribute with approximately 18 billion US-$ [2] to nearly 1/3rd of the total world actuator market (~ 55 billion US-$, [3]). Even with these rough, unverified numbers it can be assumed that hydraulic actuation is of considerable importance for present and future applications. However, there are some valid reasons for improvements in hydraulic actuation, most of them resulting from the general ways to generate, distribute and control hydraulic power. If the role of hydrostatics is restricted to power transmission only, while control and power distribution is based on switched power electronics and AC-(servo-)motors, efficiency can be increased significantly.
This leads to the well-known concept of a hydrostatic gear (left in **Figure 1**), where cylinder and speed-variable pump are directly connected in a closed circuit, thus in principle enabling for a full 4-quadrant operability. With no valves involved, a single pump can only operate a double-rod cylinder, while a single-rod cylinder requires a second pump to obtain a matching flow balance (right in **fig. 1**). Ideally, both gears will show no losses, the movement being defined by the pump flow(s) while all internal load conditions result from external forces. Given the dynamics of a servo-motor this arrangement may obviously be a considerable step towards improved efficiency and controllability.

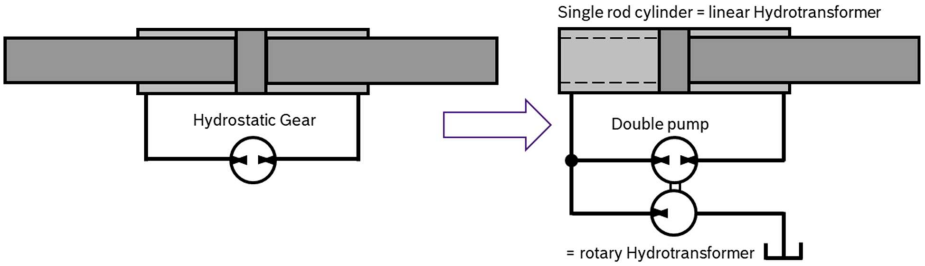


Figure 1: Basic concept of a hydrostatic gear with double- and single-rod-cylinder

Single-rod cylinders are commonly preferred to double rod-cylinders due to many good reasons (except controllability), so while the double-pump approach seems to be a consistent idea to obtain continuous 4-quadrant-ability, there are 2 issues of this kind of gear that deserve attention.

- Firstly, if pump flows do not meet the area ratio of the cylinder, the pressure level in the circuit will change during and because of a movement.[1] It also will because of leakage. The effects on the operability turn out to be considerable.
- Secondly, while pressure on rod side is higher than on piston side, which is valid for a wide range of load conditions, the power flow of the pumps will be in opposite direction, making an efficient power exchange between both units advisable, e.g. by mechanical coupling.

Both issues shall be treated in the following under various aspects.

2. TOPOLOGY OF ELECTRO-HYDROSTATIC LINEAR GEARS

2.1. Basic Concept, basic traits and basic prerequisites

Some relevant traits of hydrostatic gears show clearly in a model-based comparison with a valve-controlled cylinder drive. The scope of models may be minimal but should consider pump leakage.

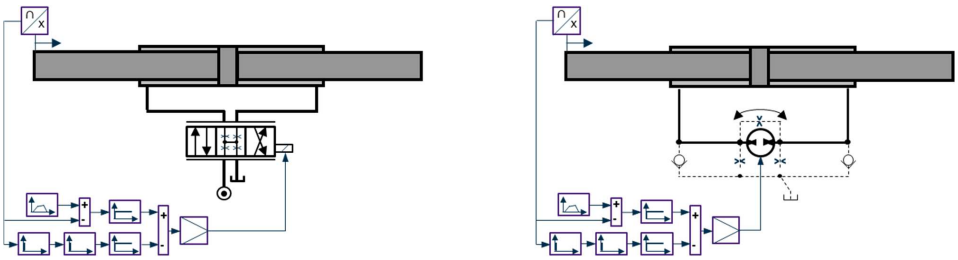


Figure 2: Schemes of simulated drive models, pump with leakage and suction valves.

Because of external leakage, pressures in the pump driven axis will tend to decrease to reservoir level in stand-still and without load. Any motion of the pump or of the cylinder will lead to a pressure increase in one and a (undesirable) pressure drop below reservoir level in the other cylinder chamber. Check valves (see **fig. 2**) are commonly used to ensure filling of the chamber and protecting components. Unfortunately, in doing so, they introduce discontinuous state changes showing in poor, unreproducible controlled performance, which shall be illustrated by **Fig. 3**. In initial state, shown in the left scheme in **fig. 3**, with pump in standstill and no load applied, chamber pressures are down to reservoir level. The check valves are closed and can therefore be neglected in the scheme. An external

[1] Example: if the additional pump covers more than the share of the rod area, pressures will increase during extraction and decrease during retraction, no matter what the load conditions are.

load displaces the piston and one chamber pressure decreases, opening the respective check-valve and connecting the chamber to reservoir, as shown in the right scheme in **fig. 3**. In this state, the axis behaves like a plunger cylinder and structurally has lost its full 4-quadrant-ability.

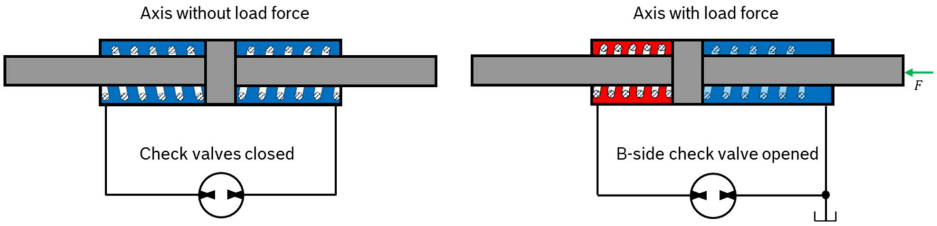


Figure 3: State of axis and check-valves with and without load force.
Schematic springs shown for illustration purposes only.

While cylinder chambers are a closed space, the fluid inside acts like a spring attached to the piston, represented by symbolic springs **fig. 3**. This may clarify that the opened check-valve keeps pressure constant and disengages one "spring" completely, thus decreasing load stiffness and dynamics of the gear considerably. Oscillations at this operation point cause valve chattering and discontinuous state-changes[2], showing in a simulation with the following main model and simulation parameters:

- Cylinder dimensions: 50/36-36/600 (in mm), moved mass 300 kg.
- Valve characteristics: Nominal flow 20 l/min @ 70 bar, 2% neg. overlap, supply 210 bar.
- Pump size 11.5 ccm, leakage coefficient 0.0015 l/min/bar, tank pressure 2.5 bar.
- Target values for position interpolation: travel distance 400 mm, velocity 400 mm/s.
- Load is applied near extracted position, resulting in about 150 bar of pressure difference.

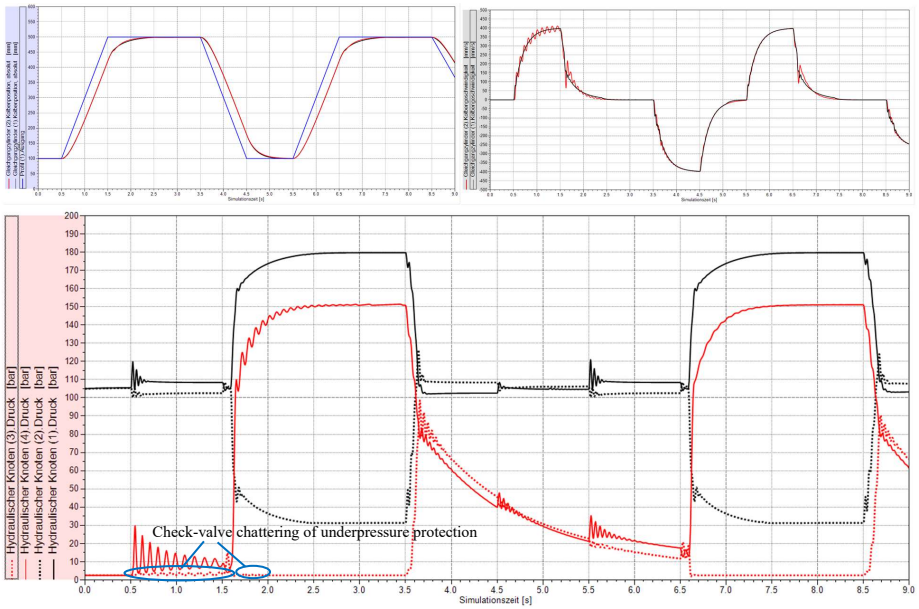


Figure 4: Simulation results for valve- and pump-controlled axes in an exemplary press process.

[2] The lift-off of the unloaded spring shown in **Fig.6** will not happen in the shown way, as the cylinder will always be oil-filled from the reservoir. Nevertheless, it helps to illustrate, that the gear changes its state in a discontinuous way.

Fig. 4 shows the simulation results in 3 consecutive plots: position (top left), velocities (top right), and pressures (bottom). Values of valve-controlled axis are black, pump-controlled axis in red, set-values are blue. Pressures of the retracting chambers are dotted. The following findings are important:
Position (top left):

- Both types of axes are controllable and reach the demanded positions, following the set-value-path with a contouring error corresponding to the desired loop gain of the position controller.

Velocity (top right):

- Progression is strongly different in the two cycles.
- If velocity were relevant for axis performance, the behaviour of the pump-driven axis would most probably be inacceptable in the first cycle, as acceleration is periodically changing sign.

Pressures (bottom):

- With valve-control chamber pressures are in average half of supply pressure (105 bar), while pressures of the pump-controlled axis are defined by the highest load pressure, resulting in a level of around 85 bar after decrease of the load force which then decreases because of leakage.
- Both types of axes show considerable oscillations when accelerating or when load is changed.
- Oscillations of the pump-controlled axis are significantly higher. One reason are the small hydraulic losses, which increase efficiency, but decrease dissipation of oscillation energy.
- Eigenfrequency with valve-control is 24 to 25 Hz, corresponding to an effective load stiffness of roughly 7.4 kN/mm, which is alike for pump control with pressures at around 85 bar.
- Eigenfrequency of the pump-controlled axis decreases with the pressure level to 18 Hz at 20 bar, indicating only 3.8 kN/mm stiffness due to dependency of bulk modulus on pressure. Additionally, it is significantly varying between the two cycles.

In sum, the valve-controlled axis shows the desired full 4-quadrant-ability, resulting from the constantly increased pressure level. The drive state is always steady in the nominal force range. The load stiffness is up to the best possible level and reproducible in every cycle.
In the pump-driven axis the actuator is structurally not able to provide energy to both chambers at the same time, so that pressures cannot be changed without changing load conditions or movement. With constant load and motion, pressures will decrease due to leakage, until at least one chamber pressure is down to reservoir level, potentially leading to undefined, unsteady state changes, where the axis switches from 2-quadrant- to 4-quadrant-ability and back.
In this configuration the axis is not on a level that could be achieved with a valve control in terms of controlled accuracy, not to mention an electro-mechanic axis. The loop-gain of the position controller, which is at its limit in this simulation, is a factor 10 to 20 lower than what could be expected using e.g., a pre-tensioned spindle. Even more, state- and force-control are very sensitive to harsh changes in drive stiffness.[3] The conclusion is, that for maxing out the full abilities of a hydrostatic gear it must be pre-tensioned by increasing pressure level. Means to achieve this and their systematic implementation, using primary servo-control only, shall be presented in chapter 2.3.

2.2. Consideration of Single-Rod Cylinder and Double Pump as Hydrostatic Transformers

In **fig. 1** the (single-rod) cylinder of the introduced hydrostatic gear is introduced as a hydraulic transformer. This interpretation may be based on the following considerations:
If there are no other forces (friction, load or inertia) are acting on the cylinder except pressure forces, then their sum must be zero, so it is $p_A \cdot A_A = p_B \cdot A_B$. Therefore, it counts for the chamber pressures:

$$\Longrightarrow p_B = \alpha \cdot p_A \quad with\ \alpha = {}^{A_A}/_{A_B} \ and\ F = 0 \tag{1}$$

[3] There are of course means to avoid the chattering state under external force, e.g. by actively opening decompression valves, but this is only suitable as long as the operating conditions always lead to load conditions that definitely allow for 2-quadrant-operation. Furthermore, precise mechanical elements like spindle nuts are commonly also pretensioned.

If the cylinder moves with the velocity v, the following displacement flows occur at the flanks of the piston: $Q_A = v \cdot A_A$ and $Q_B = v \cdot A_B$. The relation of the flows is therefore also given by the area ratio:

$$Q_A = \alpha \cdot Q_B \quad with\ \alpha = {}^{A_A}/_{A_B} \tag{2}$$

In the given ideally load-free state the power balance over the piston is:

$$p_A \cdot Q_A - p_B \cdot Q_B = 0 \tag{3}$$

As pressures and flows are both transformed by the area-ratio α between port A and B, and ingoing power at A equals outgoing power at B, the single-rod cylinder indeed *is* a hydraulic transformer.

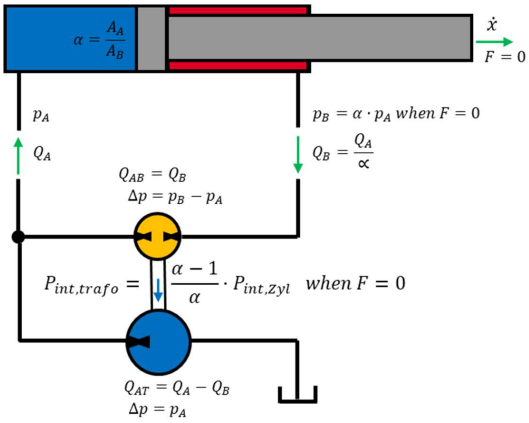


Figure 5: Power balance of the hydrostatic gear with single rod cylinder: internal power split with no external load applied.

Concerning the double pump configuration in **fig. 5** the power balance can be obtained similarly. In steady state, pump flows (marked with suffixes "AT" and "AB") must match the cylinder flows:

$$Q_{AB} \overset{\text{def}}{=} Q_B \tag{4a}$$

$$Q_{AB} + Q_{AT} \overset{\text{def}}{=} Q_A \tag{4b}$$

$$\Longrightarrow Q_{AT} = Q_A - Q_B = Q_A \cdot \frac{\alpha - 1}{\alpha} \tag{4c}$$

Two mechanically coupled pumps have identic speed, so the relation of pump sizes must be:

$$\beta \overset{\text{def}}{=} \frac{V_{AT}}{V_{AB}} = \frac{Q_{AT}}{Q_{AB}} \overset{\text{def}}{=} \alpha - 1 \tag{5}$$

Corresponding to the considerations on the cylinder, the load balance of the double-pump-unit is:

$$\Delta p_{AB} = p_B - p_A \tag{6a}$$

$$\Delta p_{AT} = p_A - p_T \tag{6b}$$

With reservoir pressure p_T and external torque M being zero, the power balance reads:

$$-\Delta p_{AT} \cdot Q_{AT} + \Delta p_{AB} \cdot Q_{AB} = 0 \tag{7}$$

Comparing equ. 3 and equ. 7 it shows that *a double-pump also acts as a hydraulic transformer and the combination with a matching single-rod cylinder results in a balanced functional hydrostatic gear.* The internal power split of the double-pump is in fixed relation to the internal cylinder power.

2.3. Basic Topologies of the primary controlled Single-Rod Cylinder

Based on the previous considerations the design of the hydraulic circuit involving double-pump and single-rod cylinder shall be generalized in the following. Basically, a hydrostatic gear always requires the control of flow between three ports: cylinder chamber A and B and reservoir T, which may accordingly involve 3 displacement-controlled flows: A-T, A-B, and B-T. **Fig. 6** gives an overview.

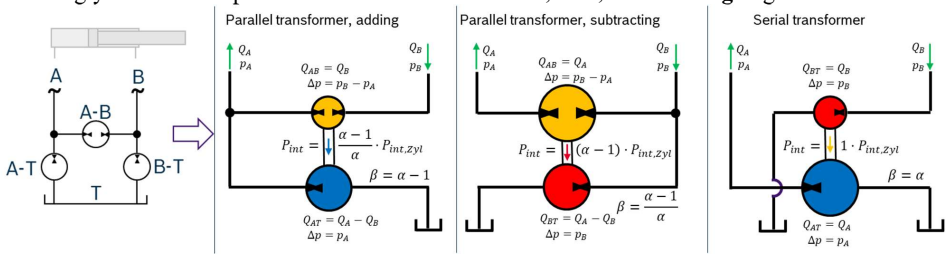


Figure 6: Systematics of suitable double-pump configurations, derived from generic function scheme.

Three suitable double-pump configurations (see **fig. 6**.) can be derived from the three basic functions. The internal power balances of these configurations and the relation β of pump sizes matching the area ratio α of the cylinder are also included in **fig. 6**. Still, these consistent hydrostatic gear configurations for single-rod cylinders require exactly balanced displacement units and cannot provide a well-defined pressure level in the gear. This problem will be addressed in the next chapter.

3. REALIZATION OF CONTINUOUS 4-QUADRANT-OPERABILITY

3.1. The internal pressure level as a separate Control Task for Cylinder Drives

The main task of a hydraulic cylinder is to realize motion and forces. Efficacious in this sense is only the sum of pressure forces of the two chambers, the value of a particular pressure being unimportant:

$$F_{p,eff} = p_A \cdot A_A - p_B \cdot A_B \tag{8}$$

In chapter 2.1 on the contrary it was shown, that for the operability of the gear an increased pressure level is crucial. The inevitable leakage of real components means an ongoing loss of pre-tension, which, however, can be compensated by an actively controlled flow from the reservoir to the gear.
It is clearly target of the presented work, to not use any valve- or throttle-control, but to stay in the concept of direct pump control. To refill the gear and keep up pressure level will therefore require:

- at least one displacement unit connected to reservoir.
- a second degree of freedom to separately control the refilling flow to chamber A and B.

The following shall show, how this additional effort can reasonably be “double used”.

3.2. Flow Matching and Pre-Loading – a Multi-variable Control Task

All combinations of double-pumps require a certain relation between the area ratio of the cylinder and the ratio of pump sizes. With real components an adequate match of geometries is unrealistic. Standardised diameters such as 100/70 (α = 1.960784…), won´t be matched by fixed displacement double-pumps. The problem can be solved by adding a third pump, driven by a separate motor, so introducing a second source of energy and a second degree of freedom, e.g. to compensate for leakage. **Fig. 7** shows an exemplary arrangement, including all relevant flows and speeds necessary for covering the displacement flows of the cylinder and also gives the power balances of the hydrostatic gear with additional motor-pump unit. Indeed, the share of power taken over by the additional unit matches with the share of B-side flow this unit covers. The additional motor pump unit may therefore be very small in comparison to the main unit. It shall be named “flow matching unit” (FMU) in the

following. Coupling the two electric drives with a DC-bus adds an electric power split, so that the complete arrangement has again no external power demand if load and losses of the gear are zero.

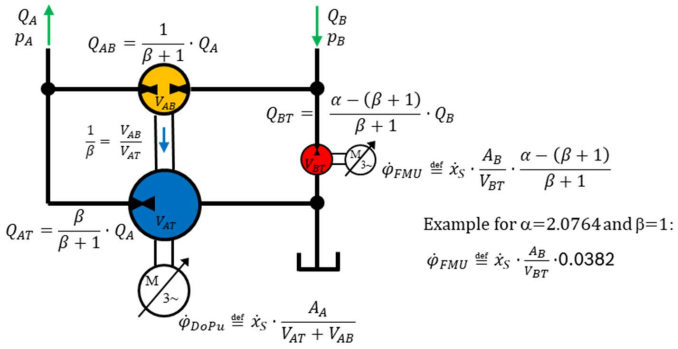


Figure 7: Adding parallel transformer with additional BT-pump as "flow matching" unit (FMU).

Fig. 8 gives an overview of possible configurations of FMU´s corresponding to the three basic types of transformers, marking their basic function as supplement or complement to the double-pump.

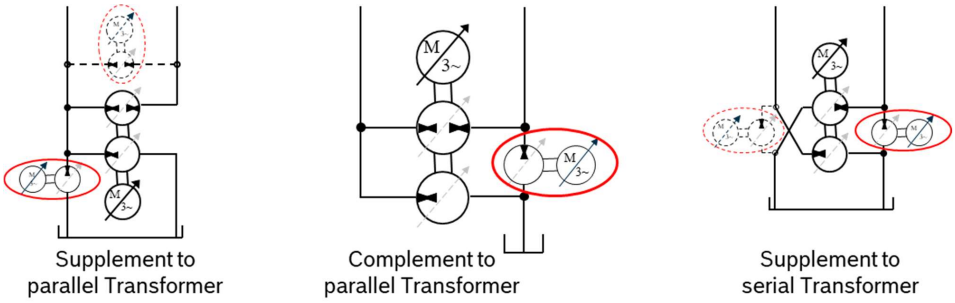


Figure 8: Combinations of "flow matching" units with different types of transformers.

3.3. Sum Pressure Control

It was shown previously that an increased pressure level is crucial for controllability and performance and that load changes should simultaneous increase and decrease pressures in the both chambers.

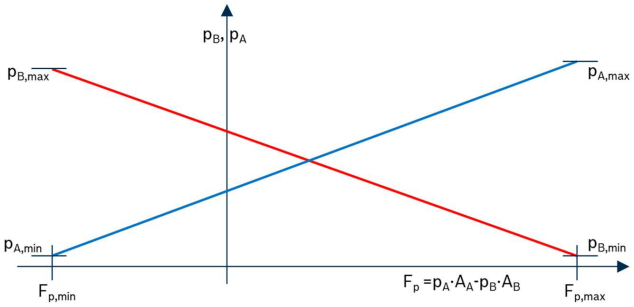


Figure 9: Chamber pressures as a function of force and pressure limits.

Maximum pressure in one and minimum pressure in the other chamber should always be reached simultaneously, thus realizing maximum pushing resp. pulling force, while pressures change contrarywise and linear with load force. This characteristic is shown in **fig. 9** and can be written as:

$$F_{press,max} = p_{A,max} \cdot A_A - p_{B,min} \cdot A_B \tag{9a}$$

$$F_{pull,max} = p_{A,min} \cdot A_A - p_{B,max} \cdot A_B \tag{9b}$$

With symmetric pressure limits $p_{A,max}=p_{B,max}=p_{max}$ and $p_{A,min}=p_{B,min}=p_{min}$ this can be achieved by:

$$p_A + p_B = p_{max} + p_{min} \tag{10}$$

Introducing a "sum pressure" p_{Sum} this can be realized by the following controller:

$$p_{Sum,Set} = p_{max} + p_{min} \tag{11a}$$

$$p_{Sum,act} = p_A + p_B \tag{11b}$$

$$Q_T = \mathrm{f}(p_{Sum,Set} - p_{Sum,act}, t) \tag{11c}$$

The controller outputs a demanded flow from the reservoir into the gear. The distribution of this flow to the two chambers must be determined separately, favourably in a way that prevents an influence of the sum pressure controller on the axis motion. The controller could be e.g. a PI-controller.

3.4. Outline for a generic model-based Multi-variable Control

The setups for hydrostatic gears considered previously show a variety that is demanding for the design of (control-)software. **Fig. 10** shows a modular software concept minimizing the adaptation efforts. The proposal for the functional architecture is based on mechatronic considerations and reflects the physical interrelationships of cylinder and drives by implementing a control-layer, the inverse [flow → linear motion] transformation layer and the inverse [rotatory motion → flow] transformation layer, representing control, cylinder and motor-pump-units. Modifications in the axis-configuration concern mainly pump arrangements, so restricting software variations to the corresponding functional layer.

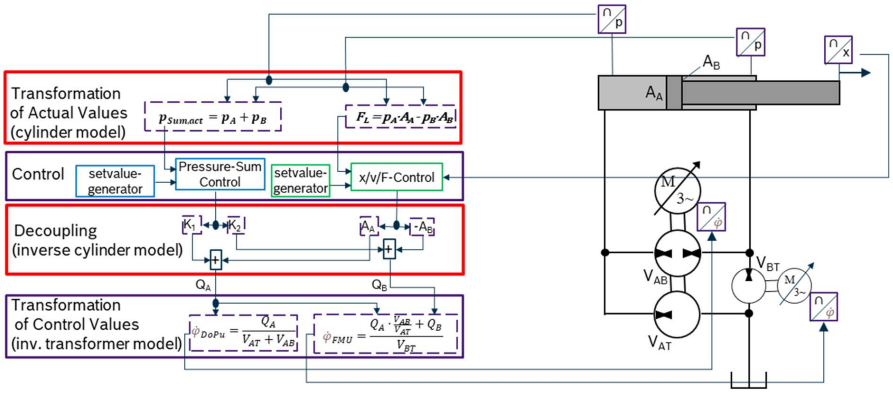


Figure 10: Functional Layers of a Model-Based Multi-variable Axis Control

4. LARGE-SCALE GRINDING MACHINE – AN APPLICATION EXAMPLE

Grinding is of crucial importance for the manufacturing of machine parts prone to sliding and rolling movement as well as for sealing surfaces. Machining of ship engines, machine tool guides etc. may give an idea of the size of grinding machines needed to this aim. **Fig. 11** shows a machine, which is an example of the "WayGrind"-series offered by german Waldrich Coburg GmbH.

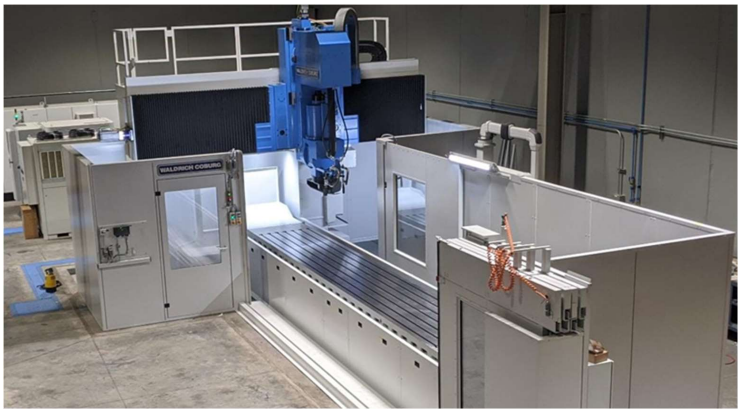


Figure 11: Grinding machine by Waldrich Coburg Company (6 m grinding length)

The size of the grinding table may give an impression of the drive stroke needed and the moved masses, even without workpiece. While processing forces are not that high in grinding, axis load is dominated by acceleration and friction. With moderate load and large stroke, bending resonance becomes a major issue for spindle drives, while hydraulic cylinders may be a cost-efficient, robust alternative. At the same time, the working conditions make the use of valve control undesirable for efficiency and thermal reasons. Altogether, an electro-hydrostatic linear drive is a superior solution.

4.1. Machine Configuration

The WayGrind series of Waldrich Coburg comprises grinding machines of various sizes in portal configuration with setup lengths of up to 15 m and clearance height and width of up to 3.5 m and 4 m. The workpiece is mounted on a grinding table, which realizes the working stroke (x-axis), while the grinding support carries and positions main drive and grinding disc in y- and z direction.



Figure 12: Table, portal and grinding support of machine chosen as application example.

A medium size WayGrind 1000 FC of this series shall be presented as an application example. The machine dimensions are 8,800 x 6,400 x 4,200 mm (LxWxH). Table mass is approx. 5.5 to, the table dimensions are 2,000x1,000 mm, carrying a workpiece with maximum weight of 4 to. Grinding width

and height are 1,200 mm x 1,000 mm. Travelling distances of the axes: 3,000 mm (X), 3,140 mm (Y) and 1,000 mm (Z). The grinding disc diameter may be up to 600 mm, the spindle power up to 22 kW. The feed velocity of the table ranges between 2 m/min and 45 m/min (750 mm/s), while the minimal infeed of 0.001 mm of y- and z-axis may give an impression of the machine precision.

4.2. Drive Configuration for the Grinding Table

For driving the machining table (x-axis) of the machine, Bosch Rexroth has developed and built a preloaded electro-hydrostatic linear drive, consisting of single-rod cylinder, serial hydrotransformer and supplementary flow matching unit as shown in **fig. 13**.
Cylinder dimensions are 80/56-3000 ($\alpha = 1.961$), the eigenfrequency of the unloaded table is around 3 Hz. The sizes of the double-pump are 100.2 ccm and 64.7 ccm respectively, so $\beta = 1.549$. The flow-matching unit size is 20.1 ccm. Motor speeds are nominal 2,260 rpm and 1,530 rpm at v = 750 mm/s. The transformer is built in serial configuration for the main reason, that the same type of internal gear pump in open circuit version can be used for all 3 displacement units. The pressure level is increased to 90 bar of sum pressure, resulting in 30.4 and 59.6 bar chamber pressure at the load free cylinder. The double-pump drive implements position control and active damping, while sum-pressure control is implemented by the FMU-drive. Cross-communication of the drives runs over SERCOS CCD.

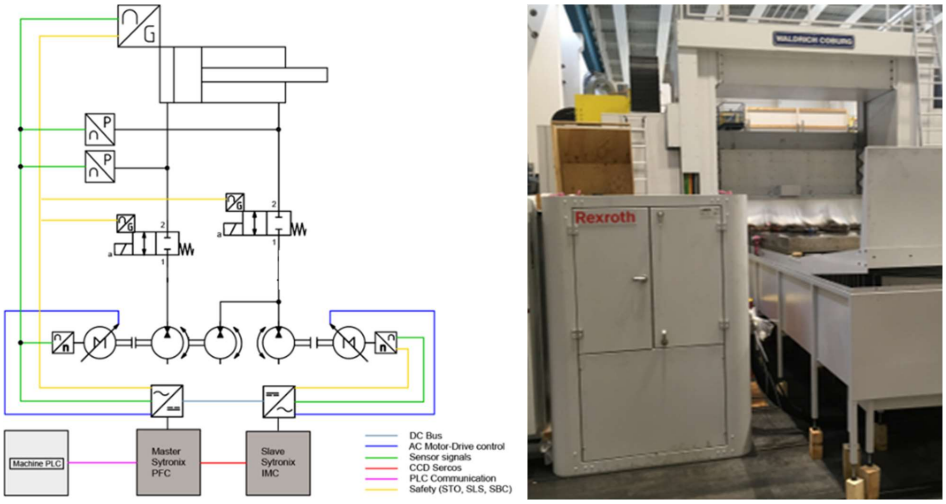


Figure 13: Principal hydraulic scheme of hydrostatic gear and hydraulic cabinet of grinding table drive.

4.3. Performance of the Hydrostatic Drive within the Grinding Cycle

During the grinding process the table with the working-piece is moved with constant velocity between two end-positions that ensure that the table has reached constant speed when the grinding wheel engages. The uniformity of the movement ensures process quality, while the distance needed to accelerate from standstill to constant velocity determines productivity and the maximal size of the working piece. In the following example the travel distance is 2.5 m The x-axis accelerates to v = 500 mm/s with a nominal acceleration of 1.0 m/s^2. **Fig. 14** shows the measured performance of the drive in a time plot of velocity and pressures. The nominal acceleration and deceleration distance given in the set-value path is 125 mm each and sufficient to reach feed and zero velocity. High static friction initially excites the axis and causes an overshoot of 6.6 % in velocity, leading to another travel

of 235 mm necessary to stabilize the axis to an accuracy < 1%.
After stabilization the velocity deviation is 0.84 mm/s in average (0.168 %) and 3.7 mm/s at maximum (0.74 %). Even given the high-quality table guides, the accuracy of the axis is remarkable. The stability shown also in pressures underlines the meaning of pre-tension for 4-quadrant-ability.

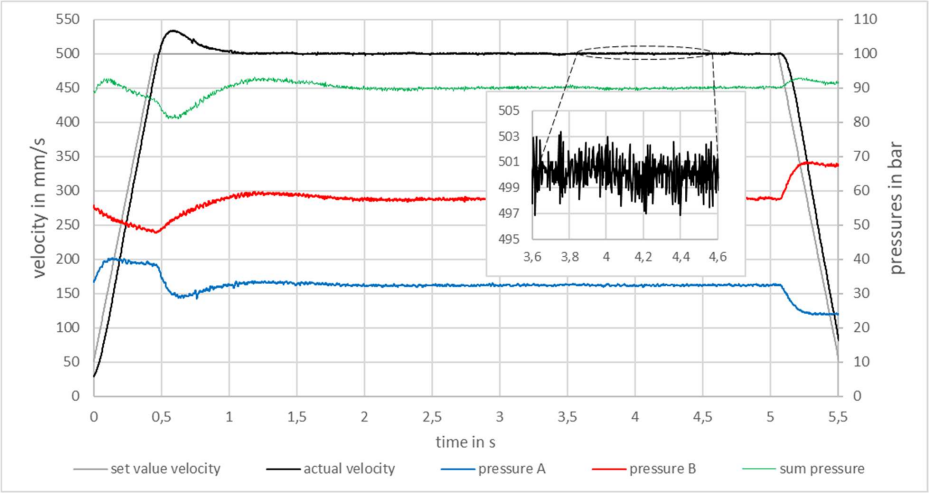


Figure 14: Velocity and pressures of grinding table drive during grinding process over 2.5 m travel.

5. CONCLUSION AND OUTLOOK

This paper concerns the design of primary controlled cylinder drives for applications that require continuous 4-quadrant ability. It focuses on the consistent use of electric power distribution and speed-variable motor-pump units. Hydraulics is used for power transmission only, but not for control. Double-pumps can be used as hydro transformer, thus being an efficient means for the build-up of hydrostatic gears involving single-rod cylinders. The basic variance of hydraulic setups is restricted to two transformer types, serial and parallel transformer. In general, hydrostatic gears are prone the loss of pressure level over time due to leakage. An additional motor-pump unit allows for actively increasing the pressure level, the control of the sum of pressures being an appropriate approach. With this, both chamber pressures may change concurrently with varying load force, thus preventing from reaching pressure limits in the complete range of operation and maximizing the stiffness of the drive. Moreover, the additional motor-pump unit allows for compensating non-matching flow balances of double-pump and cylinder. The number of circuit variations combining such a flow matching unit with a double-pump is limited to three basic configurations, two of them having a variation each. A modular approach to a generic “mechatronic software architecture” restricts software-variance due to different actuator configurations to one single layer, involving an inverse model of the pump network. A hydrostatic gear developed accordingly is able to meet all expectations in terms of functionality, accuracy and efficiency in the demanding application of an advanced feed of a large-scale grinding machine. It drives a grinding table of 5.5 to with a maximum velocity of 750 mm/s over 3 m of stroke. The shown approach can be seen structurally consistent and overall promising. Open questions concern a more detailed consideration of structural variations under the aspect of power balances, especially of the flow matching units. This should lay basis for a systematic selection and sizing of components and a systematic system design based on application demands. Moreover, a systematic examination of the implemented multi-variable control in terms of stability, robustness, and static and dynamic decoupling should contribute substantially to a successful use in demanding applications.

NOMENCLATURE

A	area	m^2, cm^2
α	area ratio	-
β	ratio of displacement volumes of double pump	-
f_0	eigenfrequency	Hz
F	force	N, kN
m	mass, weight	kg, to
M	Torque	Nm
n	speed	rpm
p	pressure	N/m^2, bar
P	power	W
Q	flow	m^3/s, l/min
$\dot{\varphi}$	angular speed	rad/s
v, $\dot{x}$	velocity	m/s, mm/s
V	volume	m^3, l, ccm
x	position	m, mm
ω_0	natural frequency	1/s

Cyl	cylinder
DoPu	double-pump
FMU	flow matching unit
A,B,T	Indices for hydraulic ports and related values
act	Abbreviation for “actual” in Indices
int	Abbreviation for “internal” in Indices

REFERENCES

[1] Lodewyks J (1994) Der Differentialzylinder im geschlossenen hydrostatischen Getriebe. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen

[2] The Business Research Company (2023), Hydraulic Cylinder Global Market Report. Website of The Business Research Company, Hyderabad, accessed 15th Dec. 2023, https://www.thebusinessresearchcompany.com/report/hydraulic-cylinder-global-market-report

[3] The Business Research Company (2023), Actuator Global Market Report. Website of The Business Research Company, Hyderabad, accessed 15th Dec. 2023, https://www.thebusinessresearchcompany.com/report/actuator-global-market-report

NOVEL ENGINEERING AND PRODUCT SOLUTIONS TOWARDS DIGITALIZATION AND SUSTAINABILITY IN VACUUM HANDLING AUTOMATION

Maik Fiedler [1,*], Walter Schaaf [2]

[1] *J. Schmalz GmbH, Business Unit Vacuum Automation, Johannes-Schmalz-Str. 1, 72293 Glatten, Germany*
2 J. Schmalz GmbH, Pre-Development, Johannes-Schmalz-Str. 1, 72293 Glatten, Germany
* Corresponding author: Tel.: +49 7443 2403- 7423; E-mail address: maik.fiedler@schmalz.de

ABSTRACT

This paper outlines new digital services for vacuum handling automation as well as new fluidic system concepts and products. Due to the large number of applications and different objects to be gripped, as well as the large product portfolio, the design process is heavily based on experience and testing. A digital engineering platform for vacuum gripping systems can help to rapidly provide optimal solutions. With a view to greater sustainability, new or previously little-used fluidic system concepts and innovative components are presented. These include, in particular, system concepts with controllable pumps, grippers with venting function, or the reduction of energy consumption. More efficient cups seal better to the surface of the object being gripped and therefore have less leakage. Simulation and test results are used to demonstrate the potential of the solutions in terms of energy efficiency.

Keywords: Digitalization, Energy Efficiency, Vacuum Handling, Vacuum Grippers

1. INTRODUCTION

Current trends such as demographic change, the shortage of skilled workers, digitalization, fragile supply chains, the need for greater sustainability and, last but not least, increased competitive pressure require new solutions and products in and for industrial production technology. Digital support for the product development process, modular systems and energy-efficient components and system concepts can help here. This article presents suitable concepts and products from the field of vacuum handling technology.

2. VACUUM HANDLING TECHNOLOGY AND KEY ISSUES

As a world market leader in vacuum handling technology, the J. Schmalz GmbH covers a product spectrum ranging from small suction cups with a diameter of 1 mm to layer grippers for intralogistics with a load capacity of 300 kg and vacuum lifting systems for the wind power industry with a load capacity of up to 40 tons [1]. In the area of vacuum components, there are more than 10,500 active sales items, of which over 4,000 are suction cups and over 1,500 are vacuum generators. Many advantages of vacuum handling technology are the reason for its widespread use: coverage of a wide range of requirements and applications, gentle handling and good accessibility from above, high process safety and flexibility, high process speed, energy efficient components and concepts, digitalization based on integrated sensors and therefore high IIOT readiness. The range of applications can be extended even further by combining several gripping principles [2, 3].

Vacuum gripping systems are made up of suitable support structures and vacuum components from six main groups: 1. vacuum generators, 2. connectors, distributors and filters, 3. fastening elements, 4. suction cups or suction modules, 5. valve technology, 6. sensors and switches for system monitoring [1, 2]. The product development process for vacuum gripping systems begins with a force calculation that takes into account the object properties and the handling process [1, 5]. The attainable suction force, which must be greater than the theoretical holding force with a safety factor, depends on the actual vacuum achieved. This in turn depends on the vacuum generator used and the properties of the object, which can lead to leakage. The leakage can be caused by the object-suction pairing, the air permeability of the object or by system defects. The calculation of the required holding force and the achievable vacuum results in the required suction area, which in turn is used to determine the number and type of suction cups or suction modules, taking into account the object geometry and properties. Vacuum generators and the valve technology interacting with them must be matched to the required cycle time.

The calculation of forces and the design of vacuum components when using suction cups is based on a proven scheme [1]. It contains degrees of freedom due to the large solution portfolio, which are designed in consultation with the customers or with knowledge of their wishes and objectives by the employees in sales and engineering for customer-optimised solutions. This is where the respective experience of the people involved comes into play. Such objectives or design priorities may be a fast process, high energy efficiency, little noise, cost effectiveness or short delivery time. The design process by experienced and competent personnel can therefore lead to different solutions due to the existing degrees of freedom. In view of demographic change and the shortage of skilled labour, it will become more difficult in future to provide customer-optimised solutions in this way. The product development process is therefore characterised by an iterative design calculation and a subsequent engineering phase in which the remaining degrees of freedom of the solution are designed depending on the personnel involved.

The energy efficiency and therefore the sustainability of vacuum gripping systems is significantly influenced by three system features and characteristics:

- operating energy: pneumatic or electric,
- fluidic concept of the vacuum supply: centralised or decentralised,
- type of object release: blow-off or ventilation.

Modular product kits can shorten the development time through the use of digital configurators, according to the motto "configuring instead of designing", and contribute to the circular economy and thus to sustainability through good feasibility of so-called R-strategies [6]. In addition to defined vacuum components, modular kits for vacuum gripping systems include in particular the elements for building the mechanical support structure. Due to the large number of applications for vacuum handling technology, it is advantageous to define gripper kits on an application-specific basis. Examples of product kits for gripping systems are the SLG lightweight gripper with a 3D-printed support structure, the PXT and the SXT gripper construction kits with a support structure made of tubular elements, the VacuMaster and SSP-HD construction kits for higher object weights, and the SPZ-M-C layer gripper with a combination of several gripping principles [1, 2]. Some digital configurators are already available online, planned or in use within the company as a preliminary stage of configurators.

Pure product configurators are characterised by the fact that the customers subsequently selects the design features of the solution, guided by the software and with knowledge of their application, and thus reach the definition of the overall solution. With pure application configurators, the system solution is automatically generated without user input based on the requirements from the application.

Currently available configurators for vacuum gripping systems can be mixed forms, e.g. if the geometry is uploaded as a data file in a product configurator or can at least be defined by manual input, such as in the configurator for the SLG gripping system.

3. DIGITAL ENGINEERING PLATFORM FOR VACUUM GRIPPING SYSTEMS

The visionary concept of a digital engineering platform for vacuum gripping systems as an application configurator shows **Figure 1**. The input variables are requirements that can be more or less structured. Possible mechanical configurations and fluidic solutions are generated and evaluated based on these requirements. The result should be directly usable files, e.g. the CAD data, the parts list of the gripping system, the dynamic behavior model, possibly as part of a digital twin, the offer to the customer or, for example, a production order. Standardized exchange formats for behavior models and digital twins are for example the Functional Mock-up Interface FMI and the Asset Administration Shell Format AASX, i.e. as digital components Functional Mock-up Units FMU and Asset Administration Shells AAS [7, 8] respectively.

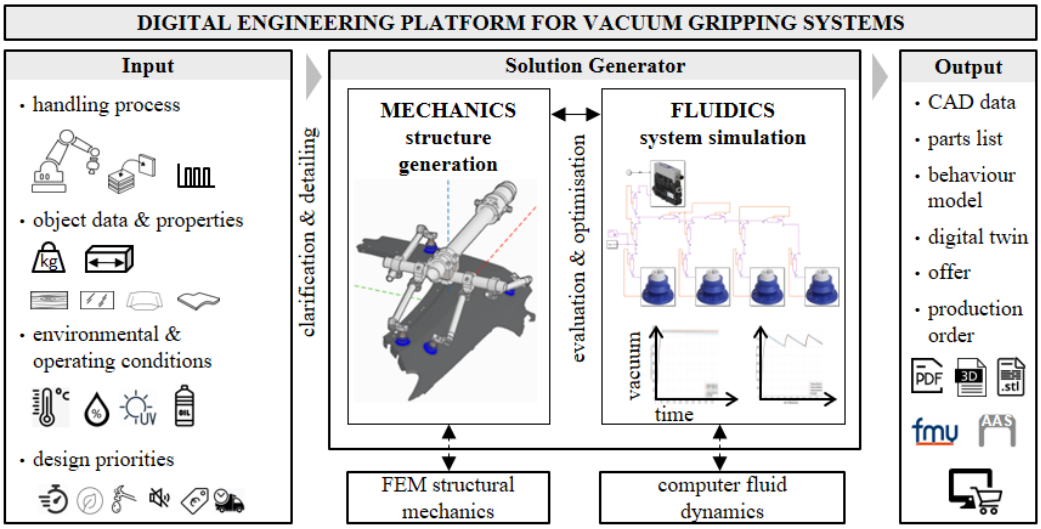


Figure 1: Visionary concept of a Digital Engineering Platform for vacuum gripping systems

Scientific results show that digital behavior models for vacuum generators and suction cups can be created and that they provide sufficiently accurate results in system models [9, 10, 11]. The modelling depth can be taken into account for balancing accuracy and computing time [12]. For the efficient creation of system models, so-called graphs can be used [13], which serve as a transfer interface between the mechanical and fluid subsystems, e.g. for hose lengths. An FEM calculation of the structural mechanics or a flow simulation can be coupled to evaluate the mechanical and fluidic partial solutions. For example, it is known that FEM analyses can be used to determine optimum gripping points on thin objects that are prone to deflection [14].

Such a digital engineering platform was implemented as a prototype on a workstation for the SXT gripper construction kit and initial tests were carried out for shaped air-tight objects of sheet metal which are typical for the automotive industry [15]. For the orchestration of the platform and the generation of the mechanical structure from modular elements the software Synera is used [16]. A system simulation realised with Matlab Simulink and the Simscape fluid library is used to calculate

the the fluid system [17]. The implemented process of generating the mechanical support structure is shown in **Figure 2**.

Based on the geometry data, a surface analysis is carried out to select suitable areas on the sheet metal for the placement of suction cups. At the same time, suitable sets of suction cups are generated based on the dimensions and other geometric features of the gripping object as well as the required total suction surface determined in the force calculation. A suction cup set contains a number of suction cups, possibly of different dimensions and types, which fulfil the requirements. Gripping points are then defined within the selected areas. Based on this, the skeleton of the mechanical support structure is then generated, taking into account the bar elements available in the respective modular system, and the 3D CAD model is created. This is followed by the generation of the fluid distributors and hoses. The result is the complete CAD model, which consists of valid components from the modular gripper system.

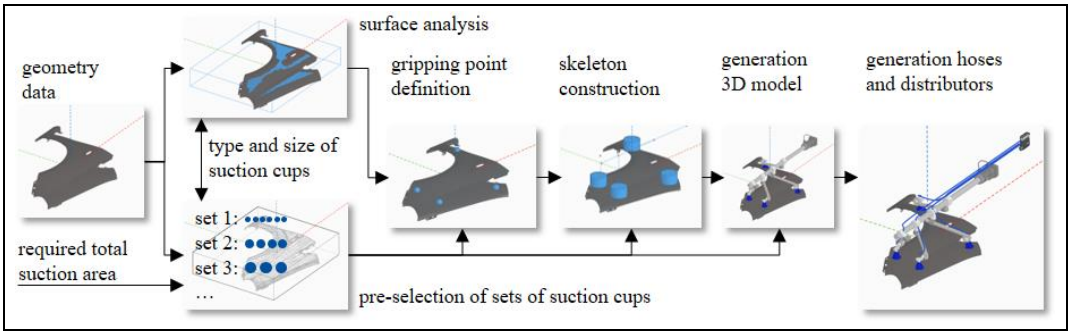
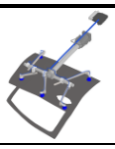
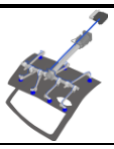
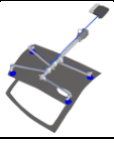
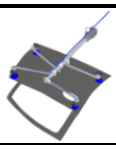


Figure 2: Procedure to generate the mechanical structure of the SXT gripper construction kit

Table 1 shows four different gripper solutions and the calculation results for the "side door" gripping object. Three different sets of suction cups are supplied centrally with a controllable SCPSi15 compact ejector and the third set of suction cups with four controllable decentralised ejectors SEAC10RP. The calculation times were less than one minute for the variants with a central ejector and slightly over 10 minutes for the variant with four decentralised ejectors. The operating pressure was set to 5 bar,rel. The blow-off time was determined by calculating the release time using the system simulation plus a safety reserve of 20 %. The internal volumes of the hoses were taken into account in the simulation. The hoses to the suction cups have an inner diameter of 6 mm and lengths in the range of 700-780 mm. They are combined to a hose with an inner diameter of 9 mm and a length of 1120 mm, neglecting a distributor.

The variants with the different number of suction cups show that the price increases with the number of suction cups. This is due to the fact that the number of mechanical components required also increases. As expected, the evacuation and release time is shortest for the system with the decentralised ejectors. However, the system costs are then the highest because four ejectors used. The energy consumption for evacuation in this system is within the range of the other variants, whereas the energy consumption for blow-off is significantly higher here, although blow-off takes place directly at the suction point. This is due to the fact that the simulation takes into account the hoses that have to be filled with pressurized air. If decentralised ejectors with a venting function were used, the energy for blow-off would be omitted and the total energy required would then be the lowest in this comparison. In this case, 44.4 % of the energy could be saved by venting instead of blow-off. The displayed variants can be presented to the user for selection or subjected to an automatic selection process.

Table 1: Comparison of different gripping systems based on mechanical and fluidic characteristics

alternative vacuum gripping system				
computing time	42 s	56 s	35 s	10 min 15 s
suction cup set	5x SAXB 60	7x SAXB 60	4x SAXB 80	4x SAXB 80
dimensions [mm]	1159 x 841 x 273	1147 x 864 x 273	1077 x 972 x 213	1095 x 991 x 213
price mechanical system	100.0 %	117.1 %	71.5 %	71.9 %
ejector and fluid concept vacuum supply	1x SCPSi 15 central	1x SCPSi 15 central	1x SCPSi 15 central	4x SEAC 10 RP decentralised
vacuum	650-750 mbar	650-750 mbar	650-750 mbar	650-750 mbar
threshold “part presence”	600 mbar	600 mbar	600 mbar	600 mbar
suction force (safety factor)	918.6 N (1.0)	1285.8 N (1.5)	1307.4 N (1.5)	1307.4 N (1.5)
evacuation time to 600 mbar	358 ms	473 ms	433 ms	158 ms
evacuation time to 750 mbar	640 ms	828 ms	639 ms	199 ms
blow-off time / release time	185 ms /154 ms	244 ms / 203 ms	197 ms / 164 ms	42 ms / 35 ms
energy consumption per cycle for evacuation to 750 mbar	645.3 J	835.4 J	644.3 J	771.2 J
energy consumption per cycle for blow-off	153.0 J	202.3 J	163.0 J	615.0 J
price fluidic system	100.0 %	108.6%	96.6%	197.7%
energy consumption per cycle: evacuation + blow-off	798.3 J	1037.7 J	807.3 J	1386.2 J
gripper weight	14.97 kg	16.68 kg	7.28 kg	8.14 kg
price mech. + fluidic system	100.0 %	115.4 %	76.5 %	96.8 %

4. NEW ENERGY-EFFICIENT VACUUM GENERATORS AND FLUID CONCEPTS

Industry accounts for around 44 % of electricity consumption in Germany, of which 7 % is used to generate compressed air [18]. It is estimated that around 20 % of the compressed air produced is used for pneumatic drives [18] and around 5-20 % for pneumatic vacuum generation with ejectors [19]. Compressed air is often referred to as the "most expensive form of industrial energy" for several reasons. Up to 30 % of the compressed air produced is lost through leakage [18]. Due to conversion and distribution losses, only 15-18 % of the energy is available at the pneumatic component in compressed air applications [20, 21]. In a typical compressed air application, approx. 75 % alone is necessary for energy provision and distribution [22]. Although the losses in compressed air distribution are ideally less than 10 %, they can be up to 50 % in industrial practice [22]. Added to this is the pressure drop in the distribution system, which is often up to 0.5 bar [22]. The costs of compressed air have risen from 0.015-0.02 €/sm³ about a decade ago to 0.03-0.054 €/sm³ nowadays, depending on the electricity price used as a basis [18, 23]. Due to the disadvantages, the scenario of the "compressed air-free factory" was established [24]. Nevertheless, pneumatic components have important advantages: robustness, long-life operation, compact design due to high power density, comparatively low price and, particularly important in terms of energy savings, the ability to switch on and off very quickly and thus realise two-point control.

There are a number of approaches to increase energy efficiency in vacuum handling technology. For drive technology, it has been shown that up to 93 % of energy can be saved by electrifying pneumatic solutions [24]. In the evacuation process in vacuum handling technology, savings of 82-99 % are possible for typical applications with air-tight objects, depending on the internal volume and vacuum level, by using electrically controllable pumps [25]. Electric vacuum generators can also score points with a higher efficiency compared to ejector nozzles, which was stated in a study to be max. 18.2 % [26]. For the energy comparison of pneumatic and electric solutions, 0.12 kWh/sm³ can be assumed for well-designed large compressor stations with optimum compressed air distribution without leaks [18]. In the following, however, the distribution with leaks is taken into account using a higher conversion factor of 0.15 kWh/sm³, i.e. 9 W/(sl/min).

However, vacuum-based handling processes with pneumatic ejectors can also be designed to be energy-efficient. In particular, controlled compact ejectors with an air-saving function save up to 90 % energy compared to basic ejectors when handling air-tight objects. Multi-stage nozzles in compact ejectors reduce the evacuation time for air-tight objects and therefore also the air consumption. If operating points of ejector nozzles have to be adapted to specific conditions, the nozzle geometry can be optimised by means of flow simulation, for example at the transition from motive nozzle to receiver nozzle [4]. Using models to predict suction cup-object leakage and grip stability, compressed air savings of up to 30 % are estimated to be feasible [27]. Internal analyses and feedback from customers revealed that when using compact ejectors, an average of 50 % of the energy is used for blowing off. This percentage can therefore be saved by atmospheric ventilation. The principle of demand-based power provision realised with the air-saving function in compact ejectors cannot simply be transferred to electric pumps. For example, the rotary vane pumps often used have a very low permissible switch-on frequency of typically 10/hour and are therefore used uncontrolled in constant operation [28]. Until some time ago, no controllable electric vacuum generators were available below the typical sizes of rotary vane pumps [24].

Important approaches for increasing energy efficiency in vacuum handling technology are therefore: electric vacuum generators that can be controlled over the entire speed range, enabling the "passive vacuum" principle in automation, lightweight, fully electric ventilation valves with large nominal sizes and the use of system simulation for alternative fluid concepts in order to analyse their energy efficiency before they are actually installed. Accordingly, **Table 2** shows new energy-efficient controllable vacuum generators with venting function and **Table 3** venting valves for the use in vacuum handling technology.

Table 2: New energy-efficient controllable vacuum generators with ventilation function

				
product	GCPi50	ECBPi	ECBPMi	EcoGripper
dimensions	220 x 236 x 102 mm	Ø152 x 89 mm	Ø63 x 61 mm	Ø75 x 40 mm
energy type „primary“ vacuum	active	active	active	passive
weight	3220 g	775 g	230 g	282 g
max. vacuum	800 mbar	750 mbar	600 mbar	400 mbar
suction rate	46 sl/min	12 sl/min	1.6 sl/min	0.33 sl/min
rated electrical power	75 W	13 W	7,2 W	0,3 W
venting diameter	3.5 mm	2.5 mm	0.8 mm	0.8 mm

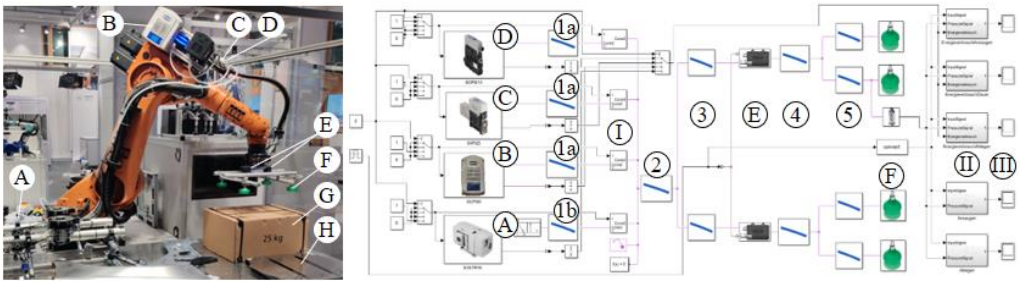
The new venting valve LQE is in the same range as the EMV10 and EMVO12 in terms of venting volume flow of 300-350 l/min. However, it stands out in this class due to its low weight and low power consumption. Of the electric valves shown, it has the best ratio of weight to venting capacity. The valve LQE and the Inline Valve IV are particularly suitable for electric gripper systems due to their low weight and the associated possibility of positioning them on the gripper as close as possible to the suction cups.

Table 3: Venting valves for use in vacuum handling technology

						
product	EMV10 24V	EMVO12 24V	LQE	Inline Valve IV	SEV	SEAC10 ECO
actuation energy	electrical	electrical	electrical	electrical	pneumatical	pneumatical
type	3/2 way valve	3/2 way valve	3/2 way valve	3/2 way valve	fast venting valve	ejector with venting
weight	2510 g	1200 g	399 g	70 g	24 g	95 g
rated electrical power	32 W	18.3 W	3.2 W	4.5 W	-	-
valve nominal diameter	Ø 10 mm	Ø 12 mm	Ø 8 mm	Ø 3 mm	Ø 4.6 mm	Ø 6 mm
venting rate	333 sl/min	350 sl/min	300 sl/min	21 sl/min	82 sl/min	39 sl/min
ratio weight/venting	7.5 g/(sl/min)	3.4 g/(sl/min)	1.3 g/(sl/min)	3.3 g/(sl/min)	0.3 g/(sl/min)	2.4 g/(sl/min)

To investigate different concepts for vacuum supply in various application scenarios, a robot cell and a digital behavior model for system simulation with defined parameters were set up, shown in **Figure 3**. The results of the system simulation for the evacuation and release process in a handling cycle in terms of energy consumption, evacuation and release time for the air-tight sheet metal without or with two LQE valves installed close to the gripper are shown in **Figure 4** and are in line with expectations. It is confirmed that around half of the energy consumption can be saved if object release is realised by venting instead of blowing off. As expected, the electrically controllable vacuum generator GCPi50 with venting function has the lowest energy consumption. This can be reduced even further by using the LQE valves close to the suction cups. The constantly operated electric rotary vane pump requires the most energy. It can be seen that the controlled pneumatic compact ejectors are significantly more energy-efficient in this case of handling air-tight parts.

The energy consumption of the variant with a rotary vane pump is slightly higher in conjunction with the two LQEs and the aim of achieving the shortest possible evacuation time by pre-tensioning the vacuum system, because the pump has to work longer against a higher vacuum. This shows that the simulation model works correctly with the stored characteristic curve of the pump. In the "sealed object and no leakage" scenario, the vacuum generators that can be switched off do not absorb any energy on the return stroke in a cycle. In contrast, the constantly running rotary vane pump also absorbs energy on the return stroke in a cycle, at least in idle mode or against vacuum, depending on whether the return stroke takes place with the venting valve open or closed. The evacuation time is shortest with the SXMPi25 compact ejector. Overall, the evacuation time was reduced for all four concepts when using the LQE valves. The deposition time is longer for the electrical concepts with integrated ventilation function than for the ejectors with blow-off function. Here too, the LQEs installed on the gripper can bring about a significant reduction.



Vacuum components and test objects: (A) EVE-TR10 + EMV10 24V, (B) GCPi50, (C) SXMPi25, (D) SCPSi10, (E) 2x LQE, (F) 4x suction cups SPB1 80, (G) air-permeable cardboard box, (H) air-tight sheet metal. **Hoses:** (1a) VSL12-9/L500, (1b) VSL12-9/L4000, (2) VSL12-9/L1100, (3) 2x VSL8-6/L180, (4) 2x VSL8-6/L200, (5) 4x VSL8-6/L250. **Vacuum generators**: air consumption suction SCPSi10 46 sl/min, SXMPi25 290 sl/min; conversion factor 9 W/(sl/min); electrical power consumption: GCPi50 0-74.4 W depending on speed; EVE-TR10 370-430 W depending on vacuum; max. suction capacity SCPSi10 36 sl/min, SXMPi25 185 sl/min, GCPi50 50 sl/min, EVE-TR10 167 sl/min; max. vacuum SCPSi10 850 mbar, SXMPi25 850 mbar, GCPi50 800 mbar:, EVE-TR10 850 mbar; blow-off flow rate SCPSi10 120 sl/min, SXMPi25 320 sl/min; operating mode: SCPSi10 / SXMPi25 air saving function two-point control, GCPi50 speed-controlled, EVE-TR10 uncontrolled. **Building blocks for simulation control:** (I) selection valves for vacuum generators (II) calculations (III) outputs. **Boundary conditions:** pressure sensor in each vacuum generator; connectors and push-in fittings neglected; suction cups pressed down before evacuation by 8 mm; switching threshold vacuum part presence 400 mbar, air saving function two-point control vacuum 450-550 mbar; release criteria: rest vacuum ≤ 10 mbar and max. release time 0.5 s; cycle: vacuum active 2s, no return stroke.

Figure 3: Robot cell and simulation model for analysing different fluid concepts

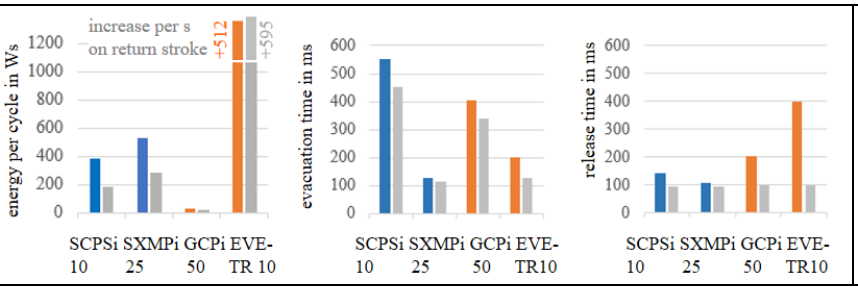


Figure 4: Results of the system simulation for the application scenario with air-tight sheet metal

5. NEW ENERGY-EFFICIENT CONCEPTS FOR SUCTION CUPS

Suction cups can contribute to the energy efficiency of vacuum gripping systems by having a low internal volume and good sealing to the gripped object. This enables shorter evacuation times and therefore shorter switch-on times for controllable vacuum generators and reduces leakage [29]. Both lead to lower energy consumption for the vacuum generators and enable concepts with passive vacuum without active vacuum generation.

Various concepts for passive vacuum with diaphragm suction cups are known, which have a diaphragm that is closed towards the object to be gripped. In addition to the effect of the weight of the object during lifting, the lifting of this diaphragm can also be achieved by actuators: mechanically by levers, fluidically by a vacuum [30], by an actuator made of shape memory alloy [31], by a preloaded spring [32] or by a magnetic coil. This lifting can be equipped with a "lifting reserve" to compensate for any leakage that occurs for a certain period of time. However, this does not ensure process safety, e.g. if a gripper has to wait for an event for an indefinite period of time while the

object is gripped. Due to the requirement for process safety in automation, a passively generated vacuum without an active vacuum generator is therefore only suitable to a limited extent. For this reason, suction cups should be used for this purpose that enable suction to compensate for leaks. A scientific study has also shown that objects can be held for up to several hours with standard suction cups, depending on the level of the starting vacuum, the surface roughness and the weight of the gripped object and the Shore hardness of the suction cup material [33]. To reduce leakage between the suction cup and the gripping object, two basic approaches can be derived from scientific research [34]: Increasing the sealing surface, i.e. widening the sealing lip, and the use of soft sealing lip materials. The approaches for achieving process safety and leakage reduction were realised in the EcoGripper shown in **Table 2**. It has already been shown that soft sealing lips bonded to various standard suction cups can improve the sealing behavior on differently structured surfaces [29, 35].

In order to rule out the influence of different suction cup types and sealing lip widths on the sealing behavior, additional silicone sealing rings were applied to standard suction cups FSGA 53 SI-55 by casting and their Shore hardness was measured. For the test, the suction cups were lightly pressed as in an automated process, suction was applied with ejectors of different nozzle sizes and thus different air consumption in continuous suction mode and the vacuum value achieved was determined. The suction cups and samples used as well as the test results are shown in **Figure 5**.

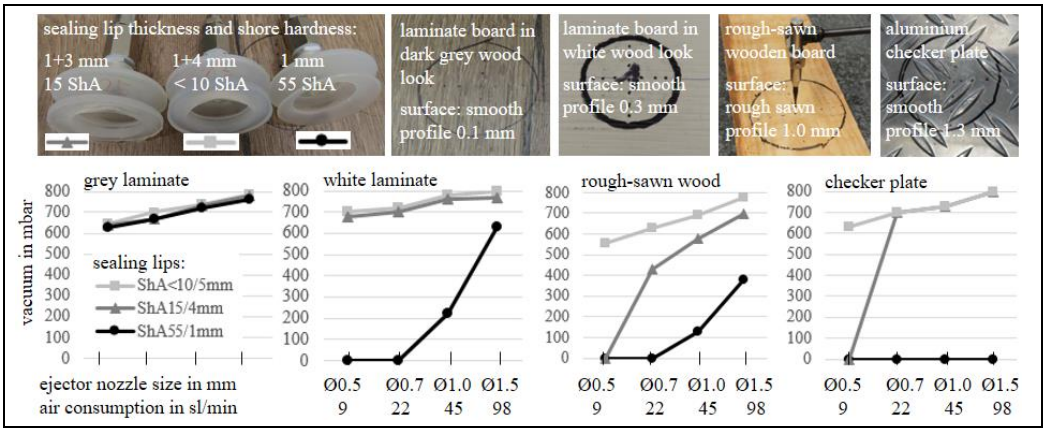


Figure 5: Results of the investigation of the sealing behavior of soft sealing lips

First of all, the known correlation is confirmed that with increasing nozzle size and the associated higher suction volume flow, leakage that occurs can be better compensated and therefore the available vacuum increases. Furthermore, the expectation that sealing lips seal better the softer they are is confirmed for all samples. The coarser the surface structure of the object to be gripped, the more soft materials contribute to the gripping success. The checker plate can only be gripped with the two softer materials. So compared to harder materials, softer sealing lip materials achieve a desired target vacuum level with a smaller ejector nozzle and therefore with less operating energy. This shows that soft sealing materials can contribute to the energy efficiency of vacuum gripping systems.

6. SUMMARY AND OUTLOOK

New contributions towards digitalization and sustainability for vacuum handling technology were presented. The visionary concept of a digital engineering platform for vacuum gripping systems was outlined in order to shorten development times and provide customer-optimised solutions. An initial implementation has shown that it is possible to automatically create the mechanical gripper structure from predefined mechanical structural elements using generative engineering in conjunction with fluidic system simulation. For the case of air-tight gripping objects, new components and concepts were described that increase energy efficiency and thus sustainability. It was shown that controllable vacuum generators, in particular electrical ones, as well as releasing by ventilation have great saving potential. System simulation can be used to analyse the energy efficiency of different fluid concepts in advance. Another contribution to sustainability is to reduce the surface leakage that occurs by using soft suction cup sealing lips.

With regard to the visionary concept of a digital engineering platform for vacuum gripping systems, further work must be carried out to determine how incomplete requirement data can be clarified and detailed automatically, e.g. by using rule-based methods or machine learning approaches. An extensive taxonomy for vacuum handling technology can form the basis for this. Time- and resource-efficient algorithms are required to generate optimised system concepts. For the system simulation of the fluid concepts, behavior models of all relevant vacuum components are required with sufficient accuracy to build up a complete library. Particularly in the case of hose interfaces and distributors, the modelling should not only take into account the geometric cross-sections, but also the fluidic inflow effects which were already investigated in the field of hydraulics [36]. The system simulation could be integrated into simulation environments that model the dynamics in order to take masses and torsional torques into account.

Energy efficiency in vacuum handling technology can be further increased. With regard to the contribution of vacuum generators and fluid concepts, the solution space of only three basic concept decisions - electrical versus pneumatic, centralised versus decentralised and blow-off versus ventilation - already leads to an optimisation problem that should be scientifically investigated with simulation for representative application scenarios, i.e. also objects with different air permeability. The utilisation of process-safe passive vacuum for automation should be further advanced. The sealing behavior of soft sealing lips of suction cups should be further investigated and economical processes for their manufacture developed. At the same time, the wear behavior must also be quantified in order to estimate the break-even points of reduced energy consumption during operation due to reduced leakage and earlier replacement of the component due to increased wear.

ACKNOWLEDGEMENTS

Parts of this work were developed in joint research projects. The authors acknowledge funding by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the project number 03ET1559 “Increased energy efficiency in vacuum handling by bionic principles and reduction of dead volumes (BiVaS)”. Also, the authors acknowledge funding by the the German Federal Ministry for Economic Affairs and Climate Action (BMWK) under the project number 03EN4050C “Data-driven system monitoring and design to secure the ecological efficiency in operation“ (OekEffDatA).

REFERENCES

[1] J. Schmalz GmbH: https://www.schmalz.com/. Accessed 8 Dec 2023

[2] Schaaf, W.; Mäder, K.-H.: Vacuum Layer Grippers: Applications, system concepts and key issues. ISR 2006, 37th International, Symposium on Robotics

[3] Schaaf, W.; Chisu, R.: Szenario: Hochflexible, RFID-gesteuerte Handhabung von Stückgut. In: Willibald Günthner, Michael Hompel (Herausgeber) Internet der Dinge in der Intralogistik. 978-3-642-04895-1 (ISBN). Springer Berlin (Verlag) 2010, pp. 319-333

[4] Kuolt, H.; Gauß, J.; Schaaf, W.; Winter, A.: Optimization of pneumatic vacuum generators – heading for energy-efficient handling processes. In: Proceedings of the IFK2016, Dresden, S. 267-280

[5] Hesse, S.: Greiferanwendung. Blue Digest on Automation. 1997

[6] Schuh, G.: Framework for Circular Production. In: Proceedings of the AWK'23 Empower Green Production, pp. 304-334. 2023

[7] Modelica Association: https://fmi-standard.org/. Accessed 8 Dec 2023

[8] Industrial Digital Twin Association e. V.: https://industrialdigitaltwin.org/. Accessed 8 Dec 2023

[9] Stegmaier, V.; Schaaf, W.; Jazdi, N.; Weyrich, M.: Simulation Model for Digital Twins of Pneumatic Vacuum Ejectors. Chemical Engineering Technology 2023, 46, No. 1, 71–79

[10] Stegmaier, V.; Eberhardt, T.; Schaaf, W.; Jazdi, N.; Weyrich, M.: A behavior model for Digital Twins of vacuum suction cups. 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '22, Italy. Procedia CIRP 118 (2023) 958–963

[11] Stegmaier, V.; Schaaf, W.; Jazdi, N.; Weyrich, W.: Anwendungsfälle und Ansatz zur Erstellung des Digitalen Zwillings aus Sicht eines Komponentenherstellers. In book: Automation 2022 (pp.5-18). DOI:10.51202/9783181023990-5. 2022.

[12] Stegmaier, V.; Dittler, D.; Jazdi, N.; Weyrich, M.: A Structure of Modelling Depths in Behavior Models for Digital Twins. 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA) | 978-1-6654-9996-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ETFA52439.2022.9921645

[13] Stegmaier, V.; Schaaf, W.; Jazdi, N.; Weyrich, W.: Efficient Creation of Behavior Models for Digital Twins Exemplified for Vacuum Gripping Systems. 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA) | 978-1-6654-9996-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ETFA52439.2022.9921487

[14] Tuleja, P.; Šidlovská, L.: „Analysis of Power Relations in the Design of Effector for Unilateral Gripping with Active Vacuum Suction Cup". In: Applied Mechanics and Materials 613 (2014), S. 286–291. doi: 10.4028/www.scientific.net/AMM.613.286.

[15] Birkert, A.; Haage, S.; Straub, M.: Umformtechnische Herstellung komplexer Karosserieteile. Auslegung von Ziehanlagen. Springer Vieweg. ISBN 978-3-642-34669-9. Berlin Heidelberg 2013

[16] Synera GmbH: https://de.synera.io/platform. Accessed 8 Dec 2023

[17] The MathWorks Inc.: https://de.mathworks.com/products/simscape.html. Accessed 8 Dec 2023

[18] N.N.: Energieeffizienz in der Produktion im Bereich Antriebs- und Handhabungstechnik, Brochure of BMWi Project EneffAH, 2012.

[19] Kuolt, H.; Kampowski, T.; Poppinga, S.; Speck, T.; Tautenhahn, R.; Moosavi, A.; Weber, J.; Gabriel, F.; Perri, E.; Dröder, K.: Increase of energy efficiency in vacuum handling systems based on biomimetic principles In: Proceedings of the IFK2020, Dresden, S. 17-26. https://doi.org/10.25368/2020.89

[20] Gloor, R.: Druckluftsysteme - Kennzahlen und Informationen über Energiesparmöglichkeiten bei Druckluftanlagen. https://energie.ch/druckluft/. Accessed 8 Dec 2023

[21] Krichel, S.; Sawodny, O. Hülsmann, S.; Hirzel, S.; Elsland, R.: Exergy flow diagrams as novel

approach to discuss the efficiency of compressed air systems. 8th International Fluid Power Conference, IFK 2012. Vol.3 Conference Proceedings

[22] Krichel, S.: Komponentenmodellierung und Strukturoptimierung in industriellen Druckluftnetzen. Shaker-Verlag, 2012. Zgl. Dissertation Universität Stuttgart.

[23] Atlas Copco Kompressoren und Drucklufttechnik GmbH: https://info.atlascopco-kompressoren.de/blog/kostencheck-was-kostet-druckluft-pro-kubikmeter. Accessed 8 Dec 2023

[24] Pohl, C.: Druckluftlose Fabrik. ISBN: 978-3-86219-521-3. 2013. Zgl. Dissertation Universität Kassel

[25] Straub, D.; Schaaf, W.: Experimental and Theoretical Investigation of Lightweight Pumps and Fluid Reservoirs for Electrically Driven Vacuum Systems in Automated Handling Processes. 11. Internationales Fluidtechnisches Kolloqium 2018. Proceedings p. 435-445

[26] Fritz, F.; von Grabe, C.; Kuolt, H.; Murrenhoff, H.: Benchmark of Existing Energy Conversion Efficiency Definitions for Pneumatic Vacuum Generators. In: Re-engineering Manufacturing for Sustainability. Proceedings of the 20th CIRP International Conference on Life Cycle Engineering, Singapore 17-19 Abril, 2013, pp 135-140

[27] Gabriel, F.: Methodenentwicklung zur Energieeffizienzsteigerung in der vakuumbasierten Handhabung. In: Schriftenreihe des Instituts für Werkzeugmaschinen und Fertigungstechnik. Essen: Vulkan, 2023. Online unter: https://nbn-resolving.org/urn:nbn:de:gbv:084-2023031710368.

[28] Fritz, F.: Ein Modell zur energetischen Analyse und Bewertung von Vakuum-Greifsystemen. Dissertation. Stuttgart University, Shaker, 2017

[29] Kuolt, H.; Kampowski, T.; Poppinga, S.; Speck, T.; Tautenhahn, R.; Moosavi, A.; Weber, J.; Gabriel, F.; Perri, E.; Dröder, K.: Biomimetic Suction Cups for energy-efficient industrial applications. In: Zukunftstechnologien für den multifunktionalen Leichtbau. Springer Vieweg, Berlin, Heidelberg 2001. https://doi.org/10.1007/978-3-662-62924-6_15

[30] Griebel, S.: Entwicklung und Charakterisierung fluidmechanischer nachgiebiger Aktuatoren am Beispiel eines multifunktionalen Sauggreifers. Universitätsverlag Ilmenau, 2021. (Berichte der Ilmenauer Mechanismentechnik ; Band 6) Technische Universität Ilmenau, Dissertation 2020. https://doi.org/10.22032/dbt.46923

[31] Motzki, P.; Kunze, J.; York, A.; Seelecke, S.: Energy-efficient SMA Vacuum Gripper System. 2016. doi: 10.13140/RG.2.2.25486.97609.

[32] N.N.: Vorbild Kugelschreiber – In einer Studie entwickelte Greiferlösungen kommen ohne externe Energiezufuhr aus. In: Der Konstrukteur 04/13, p.2-3, Mainz, Vereinigte Fachverlage, 2013

[33] Straub, D.: Methode zur technischen Auslegung von Vakuumgreifsystemen mit einer Mindesthaltedauer auf Basis fluidischer Untersuchungen. Stuttgart: Fraunhofer Verlag 2020. Stuttgarter Beiträge zur Produktionsforschung 118. Zgl. Dissertation Universität Stuttgart 2020. 978-3-8396-1666-6

[32] Tiwari, A.; Persson, B.N.J.: Physics of suction cups in air and in water. arXiv:1905.09042 [cond-mat.soft] 2019

[35] Ditsche, P. & Summers, A.P. (2019). Learning from Northern clingfish (Gobiesox maeandricus): bioinspired suction cups attach to rough surfaces. Phil. Trans. R. Soc. B, 374: 20190204.

[36] Idel'chik, I.E.: Handbook of Hydraulic Resistance. Coefficients of Local Resistance and of Friction. Springfield, 1960.

HANDS-ON APPROACH ON DEVELOPING A DEEP LEARNING ALGORITHM FOR STATE CLASSIFICATION OF A HYDRAULIC ACCUMULATOR

Oliver Mehl [1]*, Faras Brumand-Poor [2], Peter Kloft [1], Katharina Schmitz [2]

[1] *Hydac Technology GmbH, Germany*
[2] *RWTH Aachen University, Institute for Fluid Power Drives and Systems (ifas)*
* Corresponding author: Tel.: +49 151 64421032; E-mail address: oliver.mehl@hydac.com

ABSTRACT

Hydro-pneumatic accumulators are essential components in fluid power systems, serving various purposes like dampening pulsations, stabilizing flow, and ensuring safety. Monitoring their energy state is challenging due to gas leakage, requiring knowledge of gas pressure, temperature, and volume. Real-time measurements of gas temperature and volume are difficult due to transient changes during operation. This paper introduces a novel approach to classify gas mass in bladder pressure accumulators using Deep-Learning, particularly long short-term memory (LSTM) networks. The study aims to classify load situations with minimal sensor data, providing insights into the workflow for classifying multivariate time-series data with deep neural networks. Therefore, a bladder accumulator is equipped with sensors, its dynamic behaviour is recorded and used to train and validate the LSTM network's ability to classify the gas amount inside the accumulator. This method offers a promising way to determine the pre-charge pressure, a crucial parameter for assessing the accumulator's energy state.

Keywords: Deep-Learning, Hydraulic Accumulator, Multivariate-Time-Series-Classification, Condition-Monitoring

1. INTRODUCTION

With the ongoing increase in computing power, machine learning (ML) is developing faster than ever. The field of Deep-Learning as a subcategory of ML is profiting from this trend.
In contrast to classical machine learning, deep learning algorithms are more versatile in their application, albeit more complex. Due to the computing power available in today's PCs and the continuously improved algorithms behind such deep neural networks, they can be applied to an increasing array of problems. [1]

In this work, a deep neural network is developed and trained to classify the pre-charge pressure of an active bladder pressure accumulator. Since the amount of gas contained in the accumulator is directly related to the amount of energy that can be stored by the accumulator, it is important to monitor this property to ensure the safe operation of the connected hydraulic system. The so-called pre-charge pressure p_0 describes this gas amount by indicating the pressure inside the accumulator at 20°C while the bladder fills the whole body (Figure 1). Therefore, the mass of gas can be determined by using the gas equation (1) that describes the relationship of a hypothetical ideal gas, where P is pressure, V is volume, m is the mass of the gas, R is the ideal gas constant, and T is the gas temperature. As soon as the accumulator is active, the determination of the preload pressure is not trivial anymore. Due to the high dynamics with which fluid flows into and out of the

accumulator, the volumetric flow rate and thus the volume of the gas bubble cannot be determined precisely. In addition, the temperature within the gas bubble changes just as quickly and inhomogeneously due to the rapid compression. Sufficiently accurate recording of the variables required to determine the pre-charge pressure is too complex and costly to digitize a bladder accumulator in an economically viable manner. The gas pressure can be easily and cost-effectively recorded in real-time at the valve of the bladder and the oil connection. Since the relationship between preload pressure and the associated temperature behavior is influenced by the temperature measurement's inertia, it can only be inferred from the temporal evolution of the data rather than a snapshot of the data. Hence, a single measurement at a specific time does not provide any information about the prevailing preload pressure.

$$pV = mRT \tag{1}$$

In this work, the easily acquired metrics of temperature and pressure at different points of the accumulator (Chapter 2.1) are used to train a neural network that can classify the pre-charge pressure within the reservoir with sufficient accuracy based on the characteristics of these data records considered in combination. As input for the network training the raw sensor dataSufficiently accurate in this case means to arrange the pressure in intervals of 10 bar, as this corresponds to the acceptable limit before re-charging the accumulator with gas.

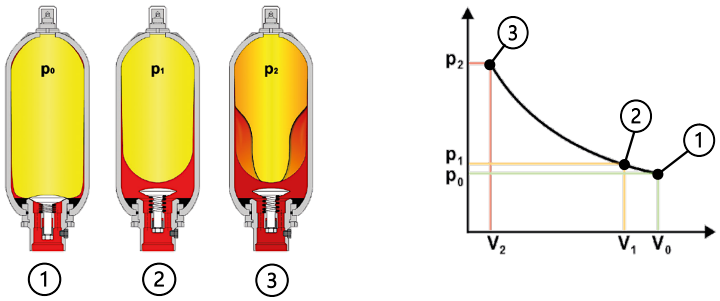


Figure 1: Overview of the states of a bladder accumulator: 1.) inactive; 2.) low operating pressure/unloaded; 3.) high operating pressure/loaded

2. METHODOLOGY

ML is finding increasing applications in the field of hydraulics, particularly in areas such as condition monitoring and predictive maintenance. Deep-Learning algorithms are increasingly being employed for these purposes. In most cases, machine learning is used to monitor entire hydraulic systems [2, 3]. Occasionally, individual components, such as hydraulic pumps, are also monitored [4].

For condition monitoring of hydraulic accumulators, ML has primarily been used to detect anomalies in operational behavior based on pressure measurements [5]. It is a novel approach in current research to equip a bladder-type hydraulic accumulator with sensors to use the data from these sensors to detect its condition using deep learning algorithms.

To accomplish tasks using ML, a structure for a data science process as shown in Figure 2 has proven useful. This process is the basis for many projects in ML. It can vary depending on the task, but the basic steps are usually the same. Since this work is also based on this procedure, the phases of such a process are presented below. [6, 7]

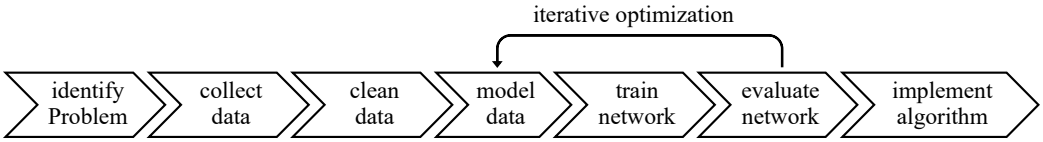


Figure 2: Steps of a typical data science process according to [6] and [7]

At the beginning of the process, the problem is identified by examining the characteristics of the data and defining the task that needs to be accomplished. Based on this analysis, decisions are made regarding the architecture and training of the neural network throughout the rest of the process. The goal is to tailor the network's structure and training approach to best address the specific problem and dataset at hand. To train the network, data is needed that represents the behavior of the pressure accumulator in different load situations at different load cycles.

2.1. Test bench setup

Therefore, a database is being recorded on the hydraulic test bench at Hydac Technology GmbH using a total of 14 Sensors mounted to a 50-liter bladder accumulator (Type SB330- 50A1/112U-330A). Figure 3 provides a more detailed overview of the kind and positioning of the used sensors. In general, the characteristic of the load being put on the accumulator at the test bench is inspired by the load on accumulators used in hydraulic plastic injection molding machines.

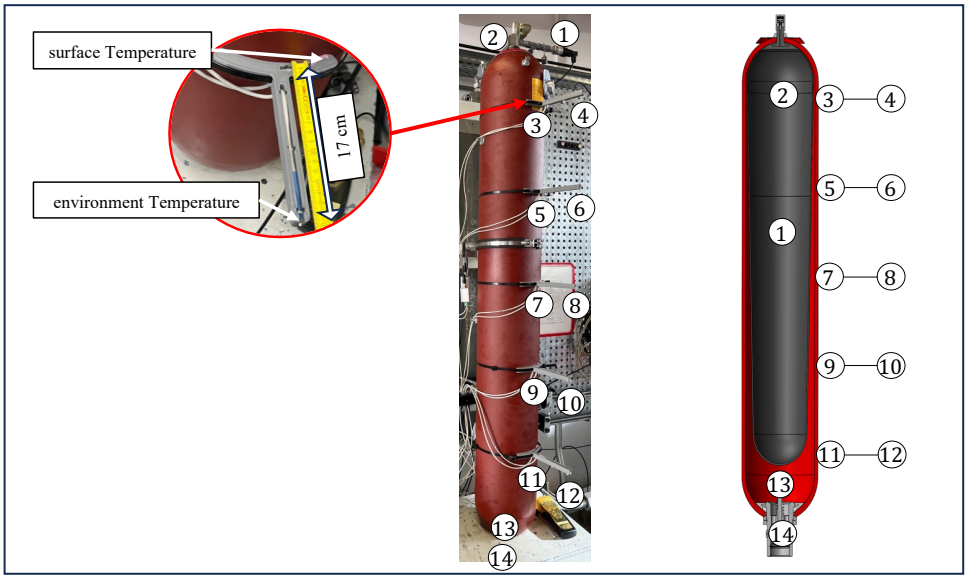


Figure 3: Detailed view of temperature sensor mount (left); picture of test bench with marked sensors (middle); schematic view of sensor positioning (right)

The sensors with the numbers 1 and 13 in Figure 3 are pressure sensors of type HDA4800 from Hydac Electronics GmbH. Number 2 and 14 are sheathed thermocouples type T5 from Conatex GmbH where the thermocouple on the gas side (nr. 2) has a nominal diameter of 0.5 mm and the one on the oil side of the accumulator (nr. 14) 2 mm. The sensors 3-12 are PT1000 temperature sensors where number 3,5,7,9 and 11 are reading the surface temperature of the accumulator and the sensors 4,6,8,10 and 12 read the corresponding environment temperature.

These temperature sensors are positioned in pairs along the accumulator. Longitudinal positioning is essential because hydraulic bladder accumulators exhibit a characteristic longitudinal heat distribution during usage. This distribution results from the bladder's varying contact with the pressure tank of the accumulator, depending on the loading state. The arrangement in pairs is intended to measure the temperature gradient at specific points on the accumulator, which is proportional to the heat flow from the surface to the surrounding environment at those locations.

It's worth noting that other sensors have been considered for data recording, such as vibration sensors to detect changes in resonance behavior, radar sensors to track the shape and behavior of the gas bladder, or an infrared pyrometer to measure the gas temperature. Ultimately, these sensors were not utilized due to challenges in recording and managing the substantial amounts of data generated by vibration recording, as well as high hardware costs.

2.2. Recorded data

For the neural network to learn the system behavior of the pressure accumulator, the accumulator is exposed to a total of three different pressure profiles with different variations. For the variations in the pressure profiles, different holding times are set at various points of the profiles, or the pressure at which the curve runs is varied. A total of 11 different pressure cycles were generated through these variations. Each of these cycles was recorded over several hours at various pre-set preload pressures. Figure 4 provides an overview of the recorded database with examples of the contained cycles. The shown table of the database indicates which of the 11 cycles has been recorded at which preload pressure. All 14 sensors were recorded at a sampling rate of 100 Hz by two CMU1000 devices from Hydac Electronics GmbH that have been synced in time. As part of the data cleaning step (Figure 2), faulty measurement outliers were corrected, and missing measurements were supplemented through interpolation. In addition, two types of labels were added to the data: Firstly, the data is labeled with the preload pressure that prevailed during the recording, and secondly, a status label is added indicating whether the accumulator was 'inactive', 'good', or 'heating' at the time of recording. With this recorded, preprocessed, and labeled database, the next step is to train the neural network.

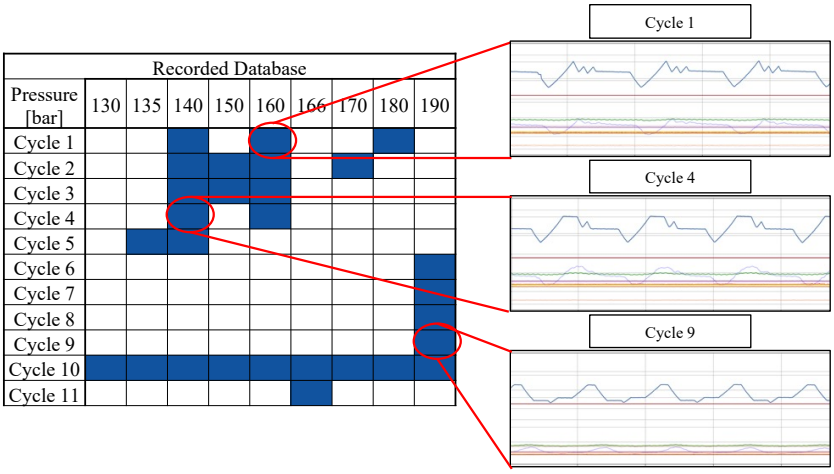


Figure 4: Overview of recorded database (left) with examples of different cycles and their different pressure profiles (right)

2.3. Network training

The basic structure of this training consists of a data pipeline that processes and models the data from the database and feeds it into the network. Figure 5 provides an overview of the recorded data of one day with the labels that have been given during pre-processing. The data is prepared in such a manner that the network can effectively recognize and learn relationships between the recorded sensor data and the preload pressure provided as the solution. Both the data pipeline and the neural network can be configured with a variety of parameters. The following subsections explain the structure of the data pipeline, the connected network, and their configurations. Subsequently, the approach to finding an ideal configuration will be explained.

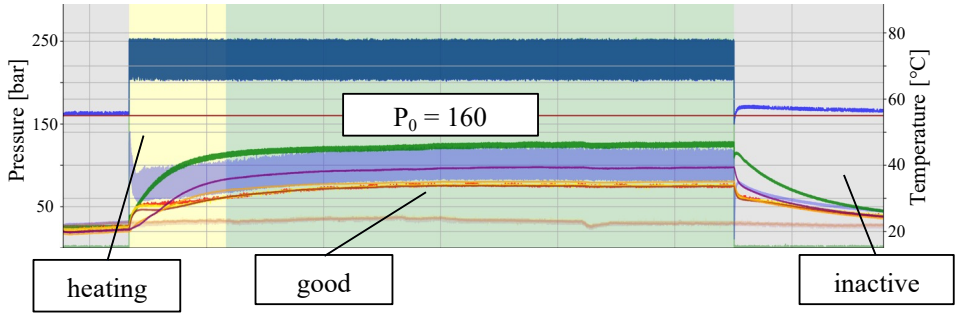


Figure 5: Preprocessed and labeled data of one day. Each cycle at each set pre-load pressure has been recorded for one day.

Data-Pipeline

The pipeline processes the data from the database in eight consecutive steps, allowing the network to recognize and learn the relationships between the sensor data patterns and the preload pressure.

1. The data to be used for training is read from the database and loaded into Pandas DataFrames [8].
2. The data is divided into time windows and organized into batches. This division into time windows is a common practice in processing time-series data to provide the network with the information contained in the temporal evolution [9].
3. Faulty batches, undesired time ranges, and values that should not be considered are filtered.
4. The data sampling rate is reduced to reduce computational load, by considering only every n-th data point for network training. Where n is an adjustable parameter to configure the desired sampling rate that is processed by the network.
5. The batches are normalized based on predefined threshold values to ensure consistent processing. These threshold values are derived from the product data sheet of the used accumulator.
6. The batches are divided into training, validation, and testing sets while preserving the chronological order.
7. The preload pressure is categorized into intervals of 10 bar steps to enable classification.
8. Finally, the data is transformed into the required format for the network of Numpy-Arrays [10].

In this framework, several variable factors influence how the network learns from the data. In this work, the number of sensors, the window size, and the sampling rate are examined in more detail.

2.4. Machine Learning Model

As mentioned in the introduction, a deep neural network is trained to autonomously recognize and learn features from the data. This entails connecting multiple layers of neurons. The type of neurons in these layers and how they are interconnected depends on the characteristics of the data and the task at hand. For the layers in this study, a neural network consisting of LSTM layers is constructed. These models are specifically designed to capture relationships over multiple time steps without consuming excessive hardware memory resources [11].

The neural network trained in this study consists of three consecutive LSTM layers, each with 64 neurons, followed by three fully connected dense layers. The subsequent dense layers have 64, 32, and, 8 neurons, which is a common approach to reduce the features identified and processed by the LSTM layers to match the number of categories to be classified. The number of units in the last layer must match the number of possible solutions. This described structure has been identified and tested in a heuristic approach. Table 2 provides a summary of the basic configuration of the network. In addition to the basic architecture, several other settings influence the learning behavior of the neural network.

Table 2: Default parameters of data pipeline and neural network that are tested during network training.

Data Pipeline	
Parameter	Value
Considered sample rate [Hz]	50
Time window size [s]	60
Neural network	
Learning rate	0.001
Dropout rate	0
Number of processed batches	128

Some settings, such as the activation functions of the Dense-Layers (set to Rectified Linear Unit) and the "return sequences" setting of the LSTM layers (set to “true”), are fixed because they have proven effective for time data classification tasks. However, settings like the dropout rate, batch size, and learning rate are randomly chosen (or set to their default values) and will be analyzed during the hyperparameter tuning process to improve training performance.

2.5. Training Process

The training of the network is designed to reveal which information, in what shape, is best suited for the intended classification task. At the same time, it aims to provide a clear understanding of how well the network learns with the chosen data pipeline and network settings. After a basic configuration has been successfully tested, achieving a classification accuracy of 99% for cycles that are included in the training data (known cycles), the focus shifts to evaluating how well the network can classify new cycles that are not contained in the training data. Considering the amount and diversity of recorded data, it is unlikely that the network, after being trained on this data, can classify significantly divergent cycles effectively. Furthermore, it is interesting to identify which sensors are relevant for a reliable classification of the accumulator's state.
To determine the optimal configuration of the pipeline and network, the training process is divided into three steps:

Firstly, in the hyperparameter tuning phase, the configuration parameters defined as variables are examined for their impact on training. This involves exploring different values for these parameters to find the combinations that yield the best results. The training process for this first step is as follows: For each parameter to be tested, neural networks are trained with different-sized datasets. These datasets contain training-, validation- and test-data. Therefore, the trained networks are already tested on the cycles they have been trained on. Each of these trained networks is then tested on the remaining, yet for the network unknown, data from the database. This unknown data contains all the varied cycles with differing characteristics compared to the training dataset.
This involves a form of cross-validation, with the particularity that, after training, the performance is tested on a dataset derived from the training data, and the final performance evaluation is separately conducted using a dataset distinct from the training data. By gradually expanding the dataset, a correct interaction between data pipeline and neural network can be assured, since the learning behaviour is expected to improve with an increasing dataset. This approach allows for a systematic exploration of how different configurations and data sizes impact the network's performance. It helps in understanding which parameter settings led to better generalization and robustness, as the network is tested on unseen data, thus providing insights into the model's ability to handle new and diverse samples. The goal is to find the optimal combination of parameters that maximizes the network's classification accuracy and generalization ability.
Secondly, in the network optimization step, the combination of the most promising parameters from the hyperparameter tuning phase is fine-tuned. Since the parameters, that have been individually tested in the hyperparameter tuning, can influence each other, this step aims to find the best combined setting of parameters. By breaking down the training process into these two steps, the study aims to systematically explore the impact of different configurations and identify the most effective settings to achieve the desired classification performance.
In the final third step, after evaluating the importance of each sensor for classification, the best configuration is tested with the data of fewer sensors. This aims to determine the number of sensors that contain significant information about the state of the accumulator and therefore are needed for a robust classification. Since more sensors also mean more data to be processed it is desirable to have as few sensors as possible to reduce the cost for hardware regarding not only the sensors themselves but also the necessary hardware for processing and storing the data.

2.6. Evaluation

The evaluation of the trained networks is conducted by analyzing the training metrics (progression of loss and accuracy) of the network training and the test performance on classifying unseen data. When considering the training metrics, an assessment of the training behavior’s stability is performed. The evaluation of test performance provides insights into how effectively the analyzed network can classify data that differs from the training data and therefore how general the network can model the real world behaviour of the accumulator based on the given sensor data.

In the analysis of training metrics, the loss and accuracy values are observed across the epochs of the conducted training. For this work, the categorical-crossentropy-function was employed. Observing these graphs can determine whether the network has reached an optimum and, if so, how rapid and reliable [12]. Ideally, in this scenario, these metrics approach their optimum values (accuracy towards one and loss towards zero) quickly at the beginning and then level off towards the end of training. Furthermore, validation performance should closely follow the training values. If the trajectory of these metrics is highly volatile or fails to converge to an optimum, this indicates that the network might struggle to learn the relationships between the input data and the class labels it is meant to predict (as illustrated in Figure 6).

To evaluate the test performance, the network to be analyzed must classify all cycles that were not used for training. For this classification, the data is processed by the data pipeline in the same manner as during training. The test data is fed to the trained network and the predicted results are then compared with the corresponding labels. The percentage of correctly classified batches is used as a metric to measure transfer performance.

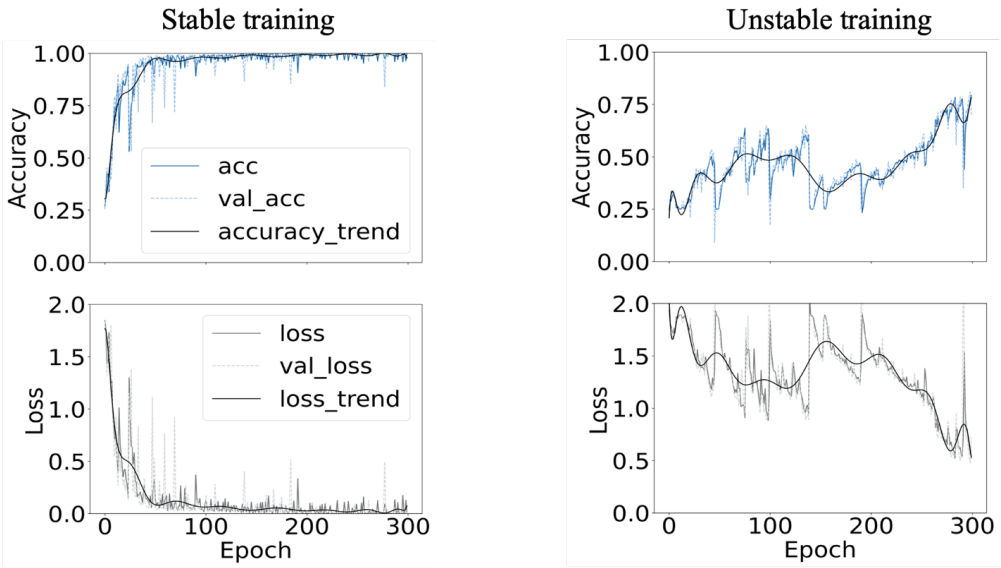


Figure 6: Comparison of examples of stable training metrics (left) and unstable training metrics (right).

As mentioned, every sensor's importance for the classification should be analyzed. It is then tested by training neural networks without the sensors that provide the least amount of information according to the before-determined importance. Since the data of each sensor is being processed as one feature by the pipeline as well as the first network layer, a feature importance analysis must be performed. One way to accomplish this is the feature permutation analysis [14], which is also being used in this work. The concept behind this permutation analysis is to permute the data of one feature (in this case one sensor) to delete information that is contained in the pattern of the sensor data and let the trained network classify the biased data. The drop in accuracy of the classification after the permutation indicates how important the information of that sensor is for the classification. In this work, the permutation is being performed in two different ways. One method is to take the data of one sensor and replace the data of each batch with the average value of that batch. That way the information contained in the pattern is replaced by a flat line. The second method is to shuffle the values of one sensor in each batch that is classified so that the pattern does not follow any characteristic curve anymore. Although both approaches yield similar results, in this work, the second (shuffle) method was used for evaluation, as the neural network appears to respond more sensitively to it.

3. RESULTS AND DISCUSSION

In the hyperparameter tuning phase, starting from the successfully tested base configuration, the parameters summarized in Table 2 were individually varied and analyzed. As a result, a total of 12 neural networks were trained, each using increasingly diverse and varying dataset sizes. These networks were then compared based on the presented metrics.

As expected, the learning behavior of the network becomes increasingly stable with larger and more diverse datasets. This behavior is characteristic of deep neural networks and indicates a successful synergy between the data pipeline and the neural architecture [15]. Furthermore, it can be observed that nearly every tested configuration is capable of correctly classifying known cycles with an accuracy of 99%. This consistency in achieving high classification accuracy highlights the effectiveness of the network configurations in capturing the underlying patterns in the data. Based on these findings, the determination of which settings to use for each of the 5 adjustable parameters in the final configuration is primarily guided by the described test performance metric of the networks trained on the largest dataset. However, considerations such as available computational capacity and the stability of the learning behavior also contribute to the selection of parameters like the sampling rate, window size, and batch size. The settings for each parameter that reached the best results are then put together into a compiled configuration. The network trained with this compiled configuration reaches an accuracy on the test data of 53%.

This compiled configuration is then tested for additional improvement potential during the network optimization phase. Since individual parameters can potentially influence each other, this iterative process aims to fine-tune the network by exploring possible synergies between different parameter settings. The goal here is to achieve the highest possible classification accuracy and robustness while effectively utilizing available resources. The tested variations with their results of the optimization phase can be found in Table 3. Subsequently, in this step, the test accuracy could be increased up to 70%. That means, that 70% of batches of new cycles have been classified correctly by the neural network.

Table 3: Overview of the trained neural networks

Window size [s]	Sample rate [Hz]	Dropout rate [%]	Nr. of processed batches	Learning rate	Sensor nr. according to Figure 3	Test-accuracy [%]
Base configuration before hyperparameter tuning						
60	50	0	128	0.001	all	**36**
Combined configuration after hyperparameter tuning						
15	50	40	200	0.001	all	**53**
Tested configurations during network optimization						
15	16	40	200	0.005	all	**47**
15	16	40	200	0.001	all	**59**
15	50	40	200	0.005	all	**70**
Results of training with a reduced number of sensors						
15	50	40	200	0.005	1,2,3,5,6,9,13,14	**65**
15	50	40	200	0.005	1,2,3,10,13,14	**56**

In the last step, this final network is used to perform the described feature importance analysis to identify important sensors. This analysis indicates that most of the environment temperature sensors do not contribute a lot of information for the classification task, while the surface temperature sensors as well as the pressure and temperature sensors in oil and gas side of the accumulator seem to deliver relevant data. To validate this outcome, two additional neural networks are trained. One uses data from eight sensors, and another utilizes data from six sensors. The results of these two network trainings are summarized in Table 3. The network trained with data from eight sensors managed to achieve a transfer performance of 65% accuracy. This is only a 5% decrease compared to the network trained using data from all 14 sensors. On the other hand, training with the specified six sensors resulted in a transfer performance accuracy of 56%. When additionally comparing the training metrics of the three neural networks trained with data from all sensors, eight sensors, and six sensors (Figure 7), it also becomes evident that the network's learning behavior is getting more uncertain as fewer sensor data are available. The observation of generalization capabilities and training metrics imply that the sensors used provide information about the energetic state of the accumulator and that the temperature distribution along the accumulator's surface provides significant value for predicting pre-load pressure. Hence, for the final tests, the configuration using the mentioned 8 sensors will be employed. This configuration strikes a favorable balance between a reduced number of sensors and delivering strong results in classifying unknown cycles.

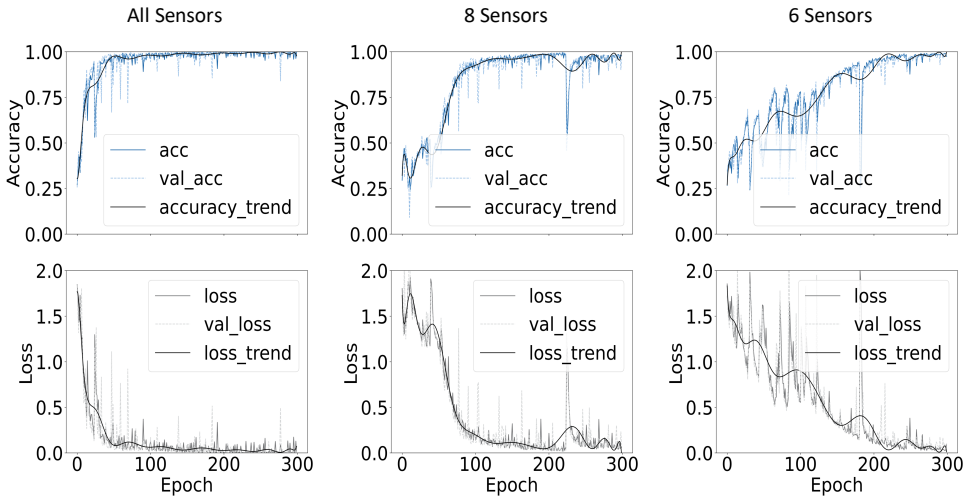


Figure 7: Comparision of the training metrics with decreasing number of features/sensor readings.

4. TESTING THE FINAL CONFIGURATION

In the previous stages of this work, around 50% of the data contained in the dataset was used not for training but for subsequent evaluation of test performance. The final assessment of the potential of this network structure will involve training a neural network using as much of the available data as possible. During the hyperparameter tuning phase, it has become evident that both the stability of learning behavior and the transfer performance of the trained networks notably increase with a growing training dataset. Therefore, during this final test, it is expected that the network's transfer performance will further improve, as the dataset becomes significantly larger and encompasses a greater diversity of different pressure profiles. This final training aims to best replicate the scenario where the pressure accumulator is connected in a new environment, and the operational data of the first day, including the target preload pressure, is recorded and contained in the training data. The

network is trained on all recorded cycles with all preload pressures except for cycle 3, of which only the data at 160 bar preload pressure is included. This network is then tested using the remaining data from cycle 3, which involves preload pressures different from the target. This test assesses how well the network can classify a preload pressure deviating from the intended one.
Due to the larger dataset, the algorithm in the final configuration is trained for 500 epochs. The evaluation demonstrates that the network training progresses very stably. Furthermore, the network managed to correctly classify about 98% of the previously unseen data from cycle 3, signifying an excellent test performance. These results suggest that the trained network is adept at generalizing its learned patterns to scenarios beyond its training data. This provides a strong indication of the network's capability to effectively handle various real-world situations involving differing preload pressures.

Finally, the network is tasked with classifying the data from an entirely new cycle. The pressure profile follows a sawtooth pattern with a labeled preload pressure of 160 bar, significantly differing from all the cycles present in the training dataset. In predicting the preload pressure of this unknown profile, the neural network correctly classifies approximately 71% of the segmented time windows.

5. CONCLUSION

In summary, it can be concluded that the developed neural network can classify preload pressure conditions it has learned with high accuracy and confidence. The closer the pressure profile being classified aligns with the learned training data, the more reliable the classification becomes. The hyperparameter tuning carried out for network development has shown promising results in enhancing the network's performance. Due to the multitude of settings available in data preprocessing and network configuration, this work did not achieve a global optimum in training configuration, indicating further optimization potential. Furthermore, it is evident that prediction quality and the ability to classify unfamiliar conditions significantly improve when more data from diverse load scenarios are available for training. This underscores the need for additional, diverse operational data on the pressure accumulator.

The conducted tests of the finally trained neural network also reveal additional potentials to further enhance the robustness of predictions. During the evaluation of feature importance, it is evident that the network can withstand the failure of individual sensors and still classify the state of the accumulator accurately. The state classification can be further optimized by considering the classification of the last assessed time windows. In the presented model, only the current time window is classified in isolation. As the system's condition changes gradually, multiple consecutive time windows can be collectively considered for state assessment. By implementing simple logical operations at the software level using these methods, the robustness and accuracy of predictions can be significantly improved without the need to train a new neural network.

This enhanced network performance in conjunction with a more diverse and expansive dataset reflects the benefit of using deep neural networks (referred to as deep learning). Another advantage of this network structure lies in the fact that the combination of data pre-processing and the neural network is fundamentally suited for processing such sensor data and subsequently automating the recognition of relationships between data and labels. This stems from the capability of such neural networks to autonomously identify connections between data and the categories they are meant to classify.

NOMENCLATURE

p	Pressure	bar
V	Volume	m^3
m	Mass	kg
R	Gas constant	$J \cdot mol^{-1} \cdot K^{-1}$
T	Temperature	K
p_0	Pre-charge pressure at 20°C	bar

REFERENCES

[1] I. Goodfellw, Deep learning. eng. Adaptive computation and machine learning series, Massachusetts: MIT Press, 2016.

[2] N. Helwig, E. Pignanelli and A. Schütze, "Condition monitoring of a complex hydraulic system using multivariate statistics," 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2015.

[3] K. Huang, S. Wu, F. Li, C. Yang and W. Gui, "Fault Diagnosis of Hydraulic Systems Based on Deep Learning Model With Multirate Data Samples," IEEE Transactions on Neural Networks and Learning Systems, no. 33, 2022.

[4] S. Tang, Y. Zhu and S. Yuan, "An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal," Engineering Failure Analysis, 01.08.2022.

[5] M.-H. Park, S. Chakraborty, Q. D. Vuong, D.-H. Noh, J.-W. Lee, J.-U. Lee, J.-H. Choi and W.-J. Lee, "Anomaly Detection Based on Time Series Data of Hydraulic Accumulator," Sensors, 01 2022.

[6] A. Kishore, "A Layman's Guide to Data Science Workflow," 2023. [Online]. Available: https://www.knowledgehut.com/blog/data- science/data-science-workflow. [Accessed 25.10.2023].

[7] S. Salimon, "Data Science Workflow: How to Create and Structure it Simplified 101," 2023. [Online]. Available: https: //hevodata.com/learn/data-science-workflows/. [Accessed 25.10.2023].

[8] W. McKinney, "Documentation Pandas," 2023. [Online]. Available: https://pandas.pydata.org/docs/. [Accessed 25.10.2023].

[9] M. M. Rahman, Multivariate Time Series Classification of Sensor Data from an Industrial Drying Hopper: A Deep Learning Approach, West Virginia: West Virginia University Libraries, 2021.

[10] "Documentation Numpy," [Online]. Available: https://numpy.org/doc/stable/index.html. [Accessed 25.10.2023].

[11] J. Schmidhuber and S. Hochreiter, "Long Short-term Memory," Neural computation, no. 9, 1997.

[12] " Artificial Intelligence Wiki," 2020. [Online]. Available: https://machine-learning.paperspace.com/wiki/accuracy-and-loss. [Accessed 25.10.2023].

[13] "Keras documentation," [Online]. Available: https://keras.io/about/. [Accessed 25.10.2023].

[14] S. Raschka, "Youtube - Feature Permutation Importance," 2022. [Online]. Available: https://www.youtube.com/watch?v=VUvShOEFdQo. [Accessed 25.10.2023].

[15] D. Sarkar, R. Bali and T. Sharma, Practical Machine Learning with Python, Berkeley, CA: Springer, 2018.

ACOUSTIC OPTIMIZATION OF A SERVO-HYDRAULIC PUMP UNIT AND AI EVALUATION OF THE SUBJECTIVE SOUND PERCEPTION

Stefan-Georg Backhaus[1]

[1]*Bosch Rexroth AG, Partensteiner Straße 23, 97816 Lohr am Main*

ABSTRACT

As part of the development of a new type of primary-controlled press drives (direct mounted servo-hydraulic pump unit on the manifold), various concepts were designed and acoustically optimized with the help of the finite element method. In this study, sound measurements were carried out on a hydraulic press. The aim was to analyse the influence of different designs of the pump unit on the sound power of the overall system. In addition to this state-of-the-art procedure, the recorded sounds were played back to participants in listening tests (> 1000) and the subjective perception of the sounds was assessed and evaluated. Psychoacoustic metrics were also analysed. With the aim of replacing time-consuming jury tests in ongoing developments, an artificial intelligence (AI) was developed and trained with the help of the data obtained. This AI helps to accelerate the development process. The results are compared based on the different investigated development stages of the servo-hydraulic pump unit.

Keywords: acoustics, nvh, servo-hydraulic pump, fem, sound power, listening tests, subjective sound perception, psychoacoustics, artificial intelligence, ai

1. OBJECTIVE OF THE STUDY

1.1. Primary-Controlled Press Drive with Servo-Hydraulic Pumps

To develop a space-saving, energy-efficient primary-controlled press drive with high power density and excellent robustness, a classical hydraulic approach with a proportional directional control valve is replaced by a servo-hydraulic pump (SHP) [1]. A SHP is a combination of synchronous servomotor and an axial piston pump and generates the volume flow required for the axis movements adjusting the two control variables: rotational speed and swivel angle [2].

As part of the development of the primary-controlled press drive, three different axial piston pumps were compared with each other as direct drives for the upper and lower pistons on a PA200HE powder press of the company Frey, as shown in **Figure 1**. Pump A and pump B are standard axial piston pumps. The A4VZA with a displacement of 71 ccm is a newly developed axial piston pump, designed for direct block mounting.

1.2. Acoustic Measures and Investigations

To ensure, that the development of the A4VZA71 will meet the acoustic requirements the following steps were taken:

1. Acoustic optimisation of the servo-hydraulic pump using finite element analysis techniques.
2. Measurements of the sound power generated by the servo-hydraulic pumps.
3. Calculation and evaluation of the psychoacoustic parameters characterizing the sound produced by the pump.
4. Determination of the human sound perception and training of an artificial intelligence model.

These objectives guide the research and analysis presented in this paper, allowing for a comprehensive understanding of the various aspects related to the servo-hydraulic pump and its acoustic characteristics.

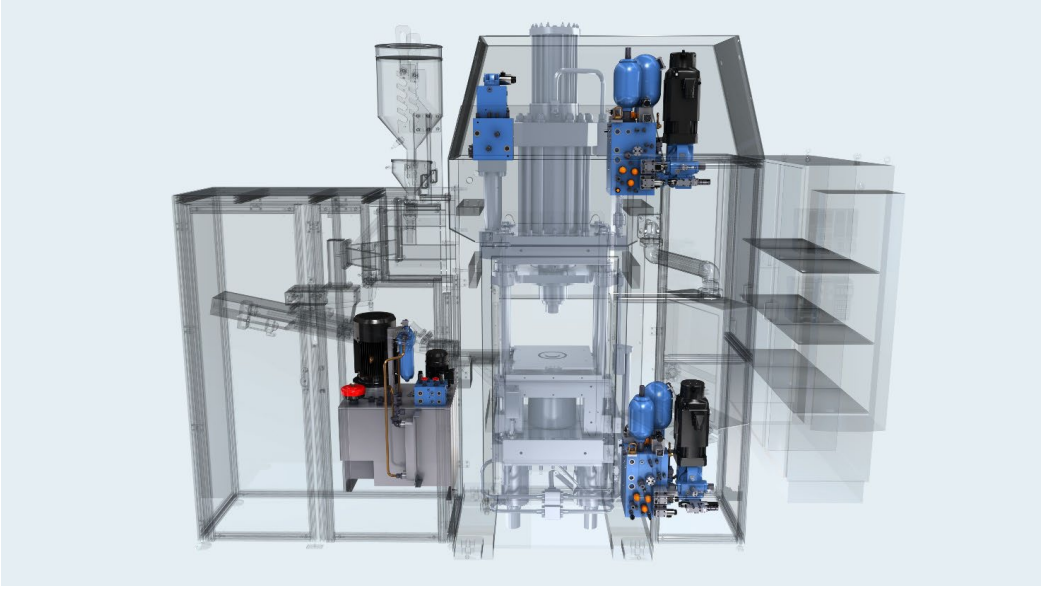


Figure 1: Schematic representation of the powder press with servo-hydraulic upper and lower piston drive

2. METHODOLOGY

2.1. Finite Element Optimisation of the Servo-Hydraulic Pump

After it had been shown that the primarily-controlled drive concept for the press meets the requirements of the market, the development of a pump especially designed for direct block mounting was started.

To meet the acoustic requirements of the market even better, the pump was investigated and optimized during the design phase by means of acoustic finite element simulations. In addition to the airborne sound power of the servo-hydraulic pump, the structure-borne sound excitation at the interface to the press was evaluated in order to diminish both quantities significantly.

Figure 2 illustrates the application of the finite element method in the acoustic optimization of the servo-hydraulic pump.

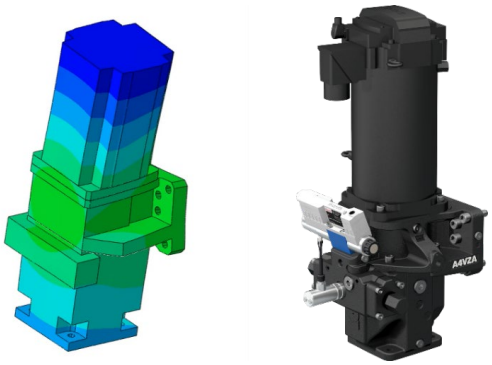


Figure 2: Finite element model (left); SHP with A4VZA71 and servomotor (right)

2.2. Sound Power Measurements

In this study, sound power measurements were carried out on a hydraulic press. The aim was to analyse the influence of different designs of the pump unit on the sound power of the overall system.

The travel speed of the cylinder piston was set as the operating parameter, and the operating modes *rapid traverse* and *press traverse* were investigated. Additionally, the different servo-hydraulic pumps were all operated with the same operating parameters. This means that a pump with a lower displacement is operated at significantly higher speeds than a pump with a larger displacement at the same cylinder piston speed.

To evaluate the influence of the servo-hydraulic pumps on the sound radiation of the entire press, all sound power measurements were conducted according to DIN 45635-1 [3]. The microphones were arranged on a parallelepiped measurement surface with a distance of 1 m from the reference box of the press, matching accuracy class 3, as shown in **Figure 3**.

During the investigations, the press was located in a test hall and could not be set up in a qualified hemi-anechoic test room. For this reason, various measures have been taken to ensure that the installation conditions affect the investigations as little as possible.

To minimize influences from the installation environment, the background noise correction K_1 was determined before each series of measurements. The environmental correction K_2 was determined using a reference sound source based on DIN 45635-1 [3]. Additionally, care was taken in all measurements to minimize disturbances from the test hall.

The diagram in **Figure 4** shows the sound power of the press equipped with different pumps. The distance between the horizontal lines in the diagram is 5 dB.

In the case of pump A, and pump B, which were not developed for the "direct block mounting" application, the sound power starts at low values, but increases significantly at different cylinder piston speeds. The sound power of the press equipped with the A4VZA71 shows a rather flat, continuously increasing trend. This was achieved through the acoustically optimized design of the SHP using the finite element method. As a result, both the structure-borne sound excitation of the press and the airborne sound radiation of the SHP could be significantly reduced. In several operating points, the sound power of the press equipped with the A4VZA71 is 10-12 dB lower. On average, the use of the A4VZA71, which has been especially developed for the "direct block mounting" application, results in a 5 dB lower sound power of the press. Since the sound power is already halved with a reduction of 3 dB, this can be considered a significant reduction. The fact that the sound power fluctuates significantly less over the increasing cylinder speed is also regarded as a sign of quality.

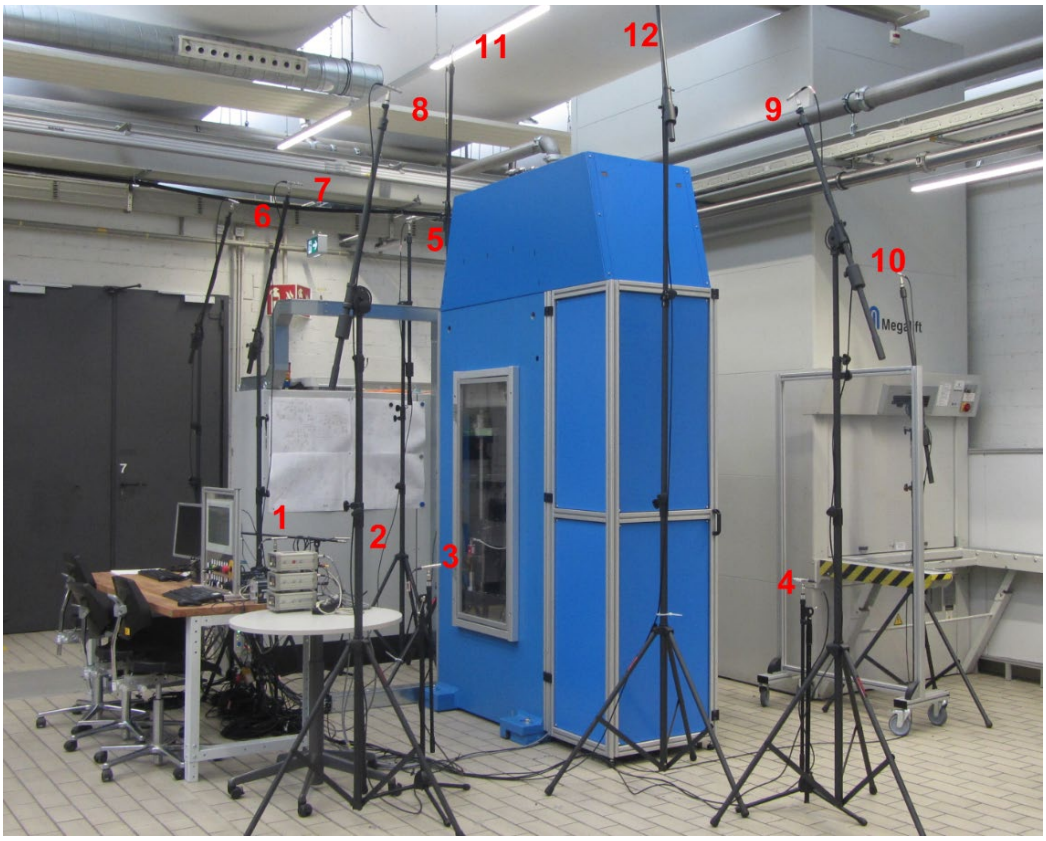


Figure 3: Parallelepiped microphone array and the powder press

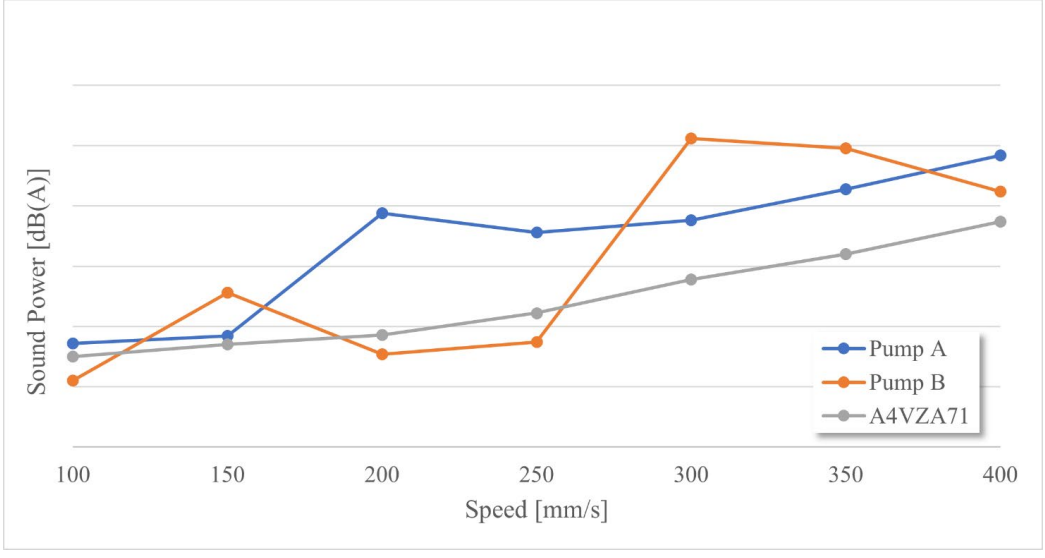


Figure 4: Sound power of the press, equipped with different pumps

2.3. Psychoacoustic Parameters

In addition to the sound power also the psychoacoustic parameters of the airborne sound emitted by the press was investigated. Since the influence of the servo-hydraulic pumps on the sound of the press is of special interest in this study, a suitable microphone position within the microphone array was selected. This position is determined by finding a balance between placing the microphone as close as possible to the operator's ear position and in a position suited to detecting the influence of the SHP. Using the sound pressure time signals recorded at this microphone position, the psychoacoustic parameters of loudness [4], sharpness [5] and tonality [6] were evaluated and compared for the different SHPs.

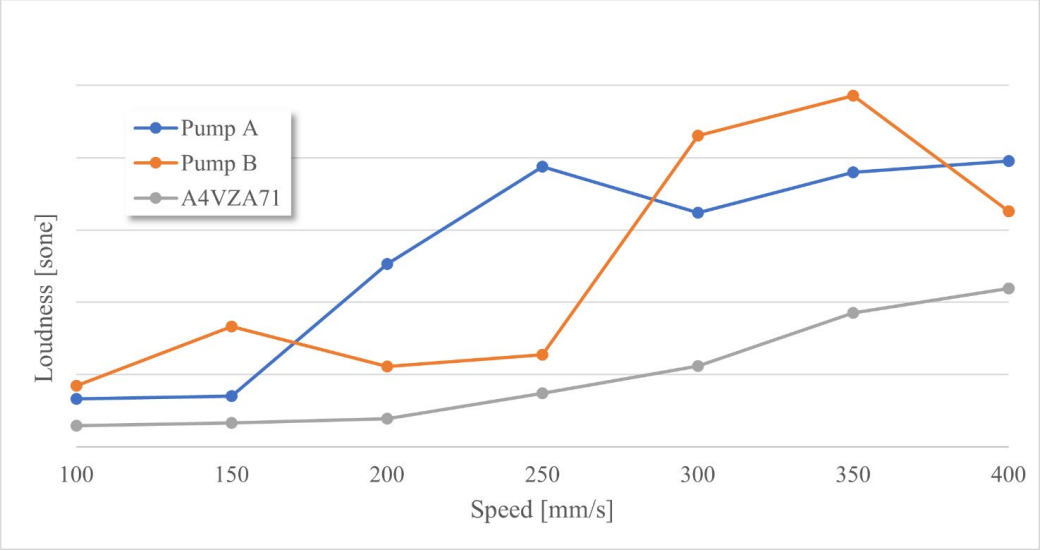


Figure 5: Loudness at microphone 4 at the press, equipped with different pumps

Loudness is the “perceived magnitude of a sound, which depends on the acoustic properties of the sound and the specific listening conditions, as estimated by that the average human listener with normal hearing” [6].

Figure 5 shows the loudness in relation to the cylinder piston speed, evaluated at microphone position 4, recorded on the press equipped with various pumps. The distance between the horizontal lines represents 5 sone.

Similar to the sound power in **Figure 4**, the loudness of pump A and pump B, which were not especially designed for the "direct block mounting" application, starts at low values and increases significantly at different cylinder piston speeds. Operating the A4VZA71 the loudness shows a rather flat, continuously increasing trend. The difference between the highest and lowest amplitude for the A4VZA71 is more than 50% less than for pump B.

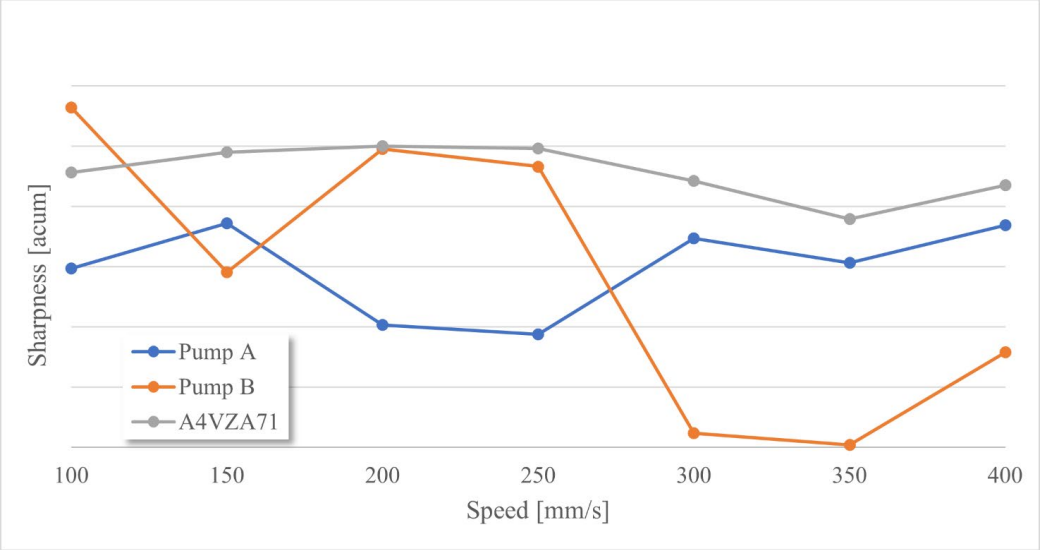


Figure 6: Sharpness at microphone 4 at the press, equipped with different pumps

Figure 6 shows the sharpness in relation to the cylinder piston speed, evaluated at microphone position 4, recorded on the press equipped with different pumps. The distance between the horizontal lines represents 0,1 acum. “Sharpness is the ratio of the loudness of the higher-frequency spectral components to the total loudness” [6].

In terms of sharpness, the behaviour is different compared to loudness and sound power. While the sharpness changes only slightly when operating the A4VZA71 over wide operating ranges, it fluctuates much more when operating the other two pumps.

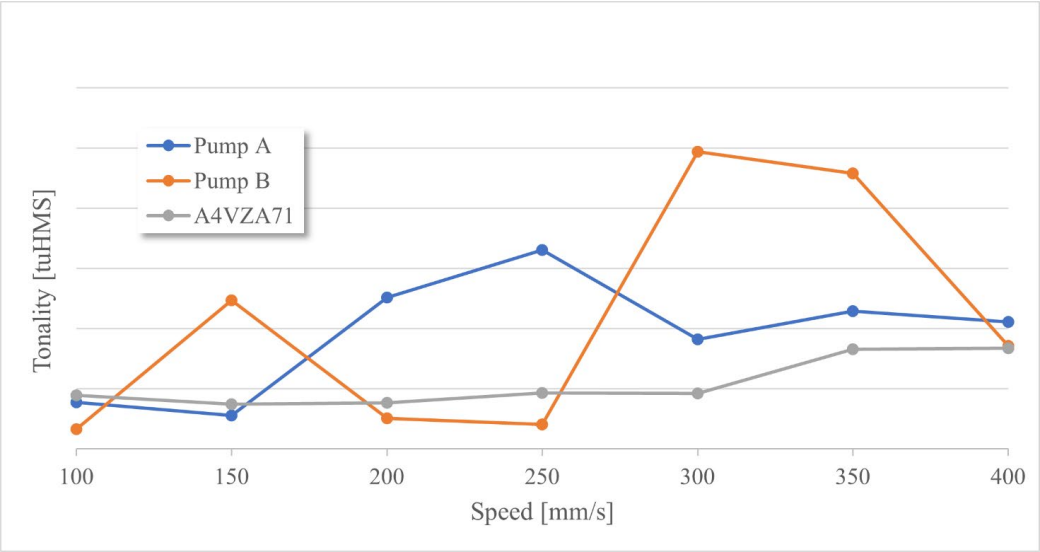


Figure 7: Tonality at microphone 4 at the press, equipped with different pumps

Figure 7 shows the tonality in relation to the cylinder piston speed, evaluated at microphone position 4, recorded on the press equipped with various pumps. The distance between the horizontal lines

represents 0,5 tuHMS. The tonality is “a characteristic of sound containing a single-frequency component or narrow-band components that emerge audibly from the total sound” [7].

The evaluation of the tonality shows that the values change only slightly during the operation of the A4VZA71 over the entire operating range and only increase slightly at 300 m/s cylinder piston speed. For the other two pumps, major changes are caused by the increase of the cylinder piston speed and higher values are reached.

2.4. Human Sound Perception and AI Training

In addition to this state-of-the-art procedure, the recorded sounds were played back to participants in listening tests (jury tests) and the subjective sound perception was assessed and evaluated. For this purpose, the sounds played back to the participants were the same as already used for the analysis of the objective psychoacoustic parameters (sounds recorded on the microphone 4). The sounds used are thus as close as possible to product sounds as is customary for listening tests [8].

Since jury testing is time-consuming and requires the involvement of a larger group of people, a solution had to be found that would save time and effort in future. With the objective of replacing time-consuming jury tests in the future and in ongoing developments, an artificial intelligence (AI) model was built and trained using the data obtained. Therefore, the number of listening test (> 1000) was chosen to be significantly larger than usual. The sounds were collected over the entire operating range of the pumps and not only focused on particularly noticeable operating points. The training and test data were split in a ratio of 80:20, as is often the case [9].

This AI model helps to accelerate the development process [10, 11].

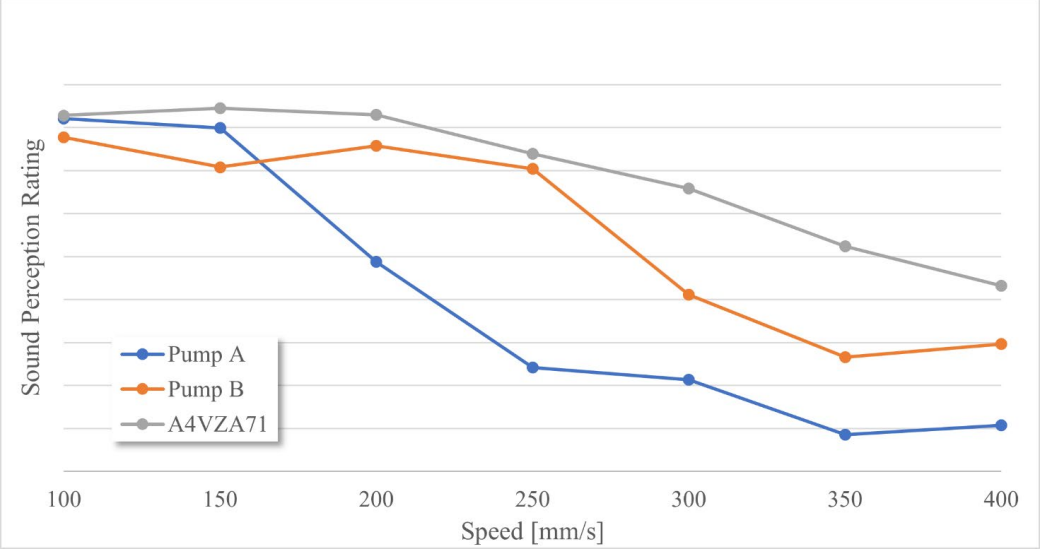


Figure 8: Sound perception rating at microphone 4 at the press, equipped with different pumps

Figure 8 shows the sound perception rating in relation to the cylinder piston speed. The values shown are the results of the AI trained to evaluate the sound signals based on human sound evaluation. The higher the values, the higher the agreement with the sounds, indicating a more positive evaluation.

It can be clearly seen that the sound perception rating remains at a high level up to medium cylinder

speeds when using the A4VZA71. The rating also drops less significantly at very high cylinder piston speeds than the rating of pumps that were not developed for the direct block mounting. Since the human perception of the machine sound is an important quality criterion, the approach described leads to an acoustical optimization of the system.

3. CONCLUSION

By developing a new servo-hydraulic pump for direct mounting on a valve block and optimising its acoustics by means of the finite element method, the acoustics of a press application could be significantly improved compared to standard axial piston units. In addition to the significant reduction in sound power, a noticeable improvement of the psychoacoustic parameters loudness and tonality could also be achieved.

Besides these objective quantities, the human perception of the sounds was also investigated. Comprehensive listening tests show that also the subjective sound perception of the machine was significantly improved due to using an acoustically optimised servo-hydraulic pump.

In order to reduce the effort of future developments noticeably, an AI was trained using the recorded data, which reflects the subjective evaluation of the test persons. The trained AI reduces the number of necessary listening tests for further development steps in the future significantly.

REFERENCES

[1] RE 08148 (2022) Technical Information | Drive Solutions based on modular Servo Hydraulic Pump Units (SHP), 2022, Bosch Rexroth AG, Lohr am Main, Germany

[2] Kübert T (2021) Richtungsweisende Effizienz durch Servohydraulik, O + P Fluidtechnik, 2021/03, 34-35, Mainz, Germany

[3] DIN 45635-1 (1984) Geräuschmessung an Maschinen; Luftschallemission, Hüllflächen-Verfahren; Rahmenverfahren für 3 Genauigkeitsklassen

[4] DIN 45631 (1991) Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum – Verfahren nach E. Zwicker

[5] DIN 45692 (2009) Messtechnische Simulation der Hörempfindung Schärfe

[6] Genuit K (2010) Sound-Engineering im Automobilbereich, Springer-Verlag, 2010, Berlin Heidelberg

[7] ECMA-418-2 (2020) Psychoacoustic metrics for ITT equipment — Part 2 (models based on human perception), 1st Edition, 2020

[8] Otto N, Amman S, Eaton C, Lake S (2001) Guidelines for jury evaluations of automotive sounds, S.V. Sound and Vibration, 04/2001, Volume 35, Issue 4

[9] Heesen B (2023) Künstliche Intelligenz und Machine Learning mit R, Springer Gabler, Wiesbaden, 2023

[10] Henke A, Herman M (2020) KI unterstützt bei Akustik-Herausforderungen im Elektrofahrzeug, all-electronics, https://www.all-electronics.de/e-mobility/ki-unterstuetzt-bei-akustik-herausforderungen-im-elektrofahrzeug-449.html

[11] Henke A (2022) If you don´t like how it sounds, you won´t buy it, Bosch, https://www.bosch.com/stories/sound-design/

Chapter 21

Digital Construction

DIGITAL ASSISTED COLLISION AVOIDANCE FOR MOBILE MACHINERY

Simon Köhler[1]*, Julian Taesch[2], Bejamin Beck[2], Denis Ritz[1], Jürgen Weber[1], Manuel Boes[3], Herbert Pfab[3]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden, Deutschland*
[2]*Construction Future Lab gGmbH, 01062 Dresden, Deutschland*
[3]*Liebherr-Werk Bischofshofen GmbH, Dr.-Hans-Liebherr-Straße 4, Postfach 49, 5500 Bischofshofen, Österreich*

* Corresponding author: Tel.: +49 351 463-34558; E-mail address: simon.koehler1@tu-dresden.de

ABSTRACT

Despite high safety standards and numerous preventive measures, both employees and objects on the construction site are exposed to a high risk of accidents. These scenarios are often caused in the handling of construction machinery, such as excavators or wheel loaders. Today there are different technologies and technological solutions for avoiding such scenarios including mobile machinery, provided by OEMs or third-party developers. This paper presents on the one hand an overview of existing technologies and technological solutions, classifying them into perception sensor, transponder-based and digital twin approaches. Furthermore, various available sensor modalities with their pros and cons for obstacle detection and classification as well as complete solutions are discussed more deeply. On the other hand, this paper introduces a dynamic collision avoidance system for mobile excavators using the cross-platform game engine unity, which combines advantages of digital twin and transponder based approaches. The system collects machine data based on so-called 3D-systems on a construction site and connects them with planning data as well as regulations and guidelines concerning protection zones. By using multiple colliders for static and dynamic objects, a reliable collision prediction on the construction site is realized. The system interacts with the excavator respectively the machine control by limiting actuator speeds, wherefore a certain minimum distance between mobile machinery and objects can always be ensured. For demonstrating the applicability of the solution, different tests on a real construction site environment are presented. The investigated scenarios are oriented towards typical track construction activities.

Keywords: Digitalization, Safety, Mobile Machinery, Conncetivity, Construction Site

1. INTRODUCTION

The rules for occupational health and safety on construction sites, RAB for short, are defined and adapted to current developments by the Committee for Safety and Health Protection on Construction Sites (ASGB), which is affiliated to the German Social Accident Insurance Institution for the construction industry. These rules are regarded as the state of the art for safety on construction sites and are binding for the client during construction. It is therefore the client's responsibility to ensure that all measures to maintain the rules are organised and enforced [1].
The "risk of construction sites" is generally known. Due to the uneven terrain, the many machines and tools moving around there, construction sites conceal a particularly large number of risks. Safety represents an elementary basic human need, whether this concerns the working environment, the

private household or the public space. According to the definition of the Federal Ministry for the Environment, Nature Conservation, Building and Risk Safety, objective safety can be considered to be given if real danger scenarios are prevented in the best possible way. Safety is guaranteed when damage occurs and the extent is reduced as far as possible. Technical safety requirements are part of product development and are demanded by customers in order to avoid accidents [2].

Despite legal requirements and preventive measures, employees in the construction industry are exposed to a particularly high risk of accidents and health hazards. The frequency of accidents on construction sites in Germany is on average more than twice as high than in the commercial sector. This is also confirmed by statistics from the German Social Accident Insurance (DGUV). In 2021, 49.84 notifiable occupational accidents per 1000 full-time workers were recorded in the construction industry. In the commercial economy as a whole, however, the rate was only 22.83 per 1000 fulltime workers. In 2021, 103,970 reportable accidents occurred on German construction sites, 85 of which were fatal [3].

One general issue of mobile earth-moving machinery is the often repeating, monotonous cycle-type workflows. In many cases, wheel loaders and excavators cyclically pick up material at one location and dump it at another. Dump trucks drive between loading and unloading points over-and-over again in short distances. Cycle times vary between around 20 seconds up to multiple minutes, but repeat often over a full work shift. Depending on the machine type, vehicles can travel up to 50% of their time in the backwards direction (e.g. a wheel loader in a short Y-cycle).

Multiple risk factors can come into play:

- very monotonous work cycles can lead to a gradual reduction of attention of the operator after long periods of time, especially in uneventful scenarios
- in stressful scenarios (e.g. a recycling yard) with multiple other vehicles and pedestrians around the machine, high productivity targets and the requirement to monitoring the environment 360° around the machine over a full shift can be very exhausting
- night-time operation can lead to accelerated fatigue and inattention

Today there are different technologies and technological solutions for avoiding accidents including mobile machinery on construction sites, provided by OEMs or third-party developers. This paper presents in section 2 an overview of existing technologies and technological solutions. In order to reduce the number of accidents and ensure safety on the digitalised construction sites of the future, different solutions are presented in this paper. Section 3 discusses solutions for mobile machines based on on-board perception sensors. In section 4, a dynamic collision avoidance system for mobile excavators using the cross-platform game engine unity is introduced.

2. STATE OF THE ART

As mentioned before, there are various technologies and technological solutions, which are already being used today for increasing safety on construction sites. Regarding the safety of mobile machinery, these solutions can be clustered into two general approaches:

- Perception Sensor Approach (Camera, lidar, radar, ultrasonic, …)
- Transponder-Based Approach (UWB, Bluetooth,…)
- Digital Twin Approach (position sensors, e.g. IMU, GNSS tracker, angle sensor)

The perception sensor approach aims to increase site safety by using local sensors and corresponding electronics on the mobile machine. For example using cameras or radar sensors for detecting potential obstacles or humans. Within this approach, the priority is always to improve site safety based on sensor data, independent of external information sources. Transponder-based approaches collect tracker information of mobile machinery and construction workers for avoiding accidents. The digital

twin approach aims to improve the safety by combining different information sources, which means sensor data and a digital environment as well as machine model. In the following, the state of art for these approaches is explained. Each solution has their pros and cons, which will be discussed further in the next chapters.
One general requirement for any solution is a high tolerance towards harsh environmental conditions. Mobile machines experience high levels of shock and vibration, dust, fog, snow, high heat and extreme cold weather conditions, uneven terrain and are also typically cleaned with high-pressure washers. Typical requirements for any electronic device exposed to the outside are therefore a temperature range of -40 … +85°C, an IP69K protection class and a multitude of environmental certifications such as salt-spray tests, chemical resistance, radiation resistance etc.

2.1. Perception Sensor Approaches

The main benefit of equipping a mobile machine with on-board perception sensors for obstacle/pedestrian detection and avoidance is the fact that this machine can be operated anywhere without the need for special organizational measures, special infrastructure or equipment on other vehicles or pedestrians.
Suitable sensors can be classified into passive or active sensors. Passive sensors only passively receive information from the environment (e.g. camera, stereo-camera), while active sensors send out signals and react based on the return of those signals (e.g. ultrasonic, radar, lidar or various time-of-flight solutions). In general, both types are suitable for the task, but with active sensors additional regulations need to be considered such as limitations in transmittable signal power, regulatory approval for certain emitted frequencies etc.
One important aspect to consider is that many mobile machines operate in tight environments and need to travel close to stationary obstacles regularly – often even in every consecutive work cycle. This brings us to the topic of obstacle classification.
Depending on the sensor modality there are various possibilities to tune the reaction of the vehicle (i.e. if a warning is given, how the warning is given, if the machine is stopping automatically) depending on the type of obstacle that is present. We describe two extreme scenarios on different ends of the spectrum on what is possible today.

a) Simple System (Obstacle recognition)
A simple system could be realized by using an affordable sensor modality like an ultrasonic sensor that only provides range data to an obstacle. Depending on the travel direction of the vehicle, an alarm can be sounded in the operator’s cabin when an obstacle comes closer than a predefined threshold value. This system is easy to obtain and tune and might work well in rather open environments, but fails to gain the operators acceptance in confined spaces due to the constant sounding of unnecessary warnings. The subsequent "alarm fatigue", i.e. the operator becoming accustomed to constant warnings, leads to warnings being ignored, even if they could be relevant in exceptional cases.

b) Smart System (Obstacle classification)
In an intelligent system scenario, the type of obstacle approaching the machine is classified (e.g. into classes such as pedestrians, cyclists, wheeled machines, vegetation, etc.). This can typically be done using AI-based approaches in combination with different sensor types such as cameras, radar and lidar. In addition to knowing the type of obstacle, the relative movement of the obstacle is also tracked. Together with the known ego-movement of the vehicle, the basis for decision-making for warnings and movement reduction is very broad.
With such a system, the challenge of driving in a confined construction environment and the goal of not overwhelming the human driver with unnecessary warnings can be achieved. A machine equipped with this technology warns the driver or intervenes in the vehicle's movements only when it is really

necessary - for example, when pedestrians are present and there is a risk of them being involved in collisions. Avoiding 'alarm fatigue' and ensuring that the operator does not become accustomed to constantly appearing unnecessary warnings in a confined working environment offers clear practical benefits.

2.2. Transponder-Based Approaches

Another approach is to equip both the machines as well as the pedestrians around the machine with tracking devices or transponders. With technologies such as UWB (ultra-wideband), these devices can be tracked relative to one another using vehicle-fixed antenna arrays. A control system of a machine will therefore always know where the specially equipped pedestrians are in relation to the vehicle. A clear benefit against on-board perception sensor approaches is that this technology is independent of pedestrian poses (standing, kneeling, lying on the ground), can localize pedestrians around corners or even through certain obstacles and is very tolerant against environmental impact such as dust or fog.

On the other hand, this approach relies heavily on organizational measures to provide site safety. Every pedestrian that can potentially get close to a machine always needs to be equipped with a tracking unit, the battery always needs to be charged and access to the work area without passing through a dedicated checkpoint must be prohibited. There are many approaches such as integrating the tracking devices into high-vis wests or helmets to try to make sure people do not forget to carry them. This is the main reason that these types of solutions are limited to very few scenarios such as underground mining and certain industrial plants today. For typical construction sites, the applicability is very limited with often multiple companies working on one site and no complete separation against external pedestrians. Scaling of costs is another challenge, since every additional potential person at the site needs another tracking device. Another issue is that this technology can potentially create unnecessary warnings or machines braking even if pedestrians are not actually in the danger zone. A typical example is a trench construction scenario in which the machine works on one side of the trench and the person works on the other.

One such system to mention is the PLINX [4] of Mesafox. This system is based on tags/ transponders, which are attached to machines, objects and people. These tags are equipped with a variety of sensors, including GNSS, proximity sensors, accelerometers, which send their data to local networks via various interfaces such as Bluetooth or UWB. If two tags draw near to each other, an event is triggered, and therefore acoustic, haptic and visual signals are activated to warn the people affected. Furthermore, individual events, which trigger corresponding warning signals, can be defined with the help of the associated software. Such individual events can be SOS signals, dead man’s switches and the definition of danger zones.

2.3. Digital Twin Approaches

The implementation of 3D machine control systems in mobile machinery is becoming increasingly common. This technological development is revolutionising the precise control of excavators and other heavy machinery through the integration of tilt/ rotation sensors, RTK-GNSS system and digital, geo-referenced construction site models. Regarding achieving higher accuracy, efficiency and cost-effectiveness, more and more companies are using these control systems. There are few advanced collision avoidance systems based on such existing 3D control systems. These systems [5] [6] offer generating virtual barriers in the digital construction site environment. These virtual barriers prevent the excavator from crossing certain areas. If the mobile machine draws near such a virtual fence, the speed is automatically slowed down or the movement is completely stopped. These measures ensure a precise and immediate response to potentially dangerous situations.

3. SOLUTIONS BASED ON ON-BOARD PERCEPTION SENSORS

Solutions that are established on the market and commercially available for the use on mobile machines today can be classified based on the protection goal (all obstacles or specifically pedestrians) and the sensor modality used.

Basic systems use ultrasonic sensors that measure a distance from the sensor to a possible obstacle in a short, cone-shaped field-of view. These types of sensors are active sensors that typically use a diaphragm actuated by piezo crystals. The returning wave is measured back using the same diaphragm and converted back into an electric signal by the piezo. The distance is calculated based on the time-of-flight principle. Typical measuring ranges are between ~0.1 m up to ~6 m, the lower limit given by the fact that the diaphragm needs some time to come to a standstill after transmitting before it can receive an incoming wave. The measuring ranges are very practical for many scenarios, however smaller obstacles such as cables or very thin posts might not be detectable. Typically multiple ultrasonic sensors are combined to cover for example the rear of a machine. If enough overlap is provided between neighbouring sensors, the position of the obstacle can be calculated based on multiple range measurements from neighbouring sensors using trilateration. This can be used for more advanced visualizations and warnings for the operator.

Radar-based systems in comparison typically allow for much greater measuring ranges (up to several hundred meters) and higher field-of-views, meaning a single sensor can easily cover one face of a vehicle. Most types of radar also not only report a range measurement but provide a relative speed output based on the Doppler principle, which can be used for filtering and smarter alarm strategies (e.g. ignore dynamic obstacles, that are close but moving away fast). Radar signals often include multiple returns that can be used to classify the object that is being measured (e.g. by AI methods), allowing for even more advanced vehicle integration strategies and solutions avoiding alarm fatigue (e.g. ignoring all obstacles that are not classified as pedestrians).

Camera-based solutions can be classified into the used spectrum (visible spectrum vs. infrared spectrum for pedestrian detection) and the amount of cameras. While most types of digital cameras allow for AI-based obstacle classification especially stereo-cameras have the added benefit of being able to calculate the distance of objects based on the stereo-vision principle (at the cost of heavy computing power needs on the vehicle).

To mitigate this cost and scalability issue many approaches have been developed also for the mono-camera to provide distance information, the structure-from-motion algorithm or AI mono-depth-models being named as examples.

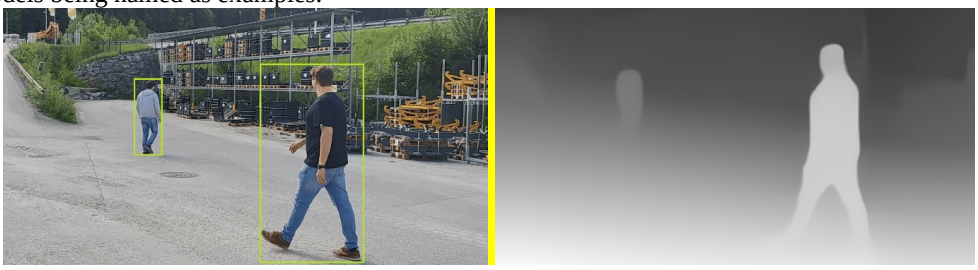


Figure 1: Example of 2D RGB camera with AI-based object classification and AI mono-depth estimation

Lidar sensors are positioned on the upper end of the spectrum of on-board perception sensors when it comes to capabilities (and costs). Depending on the model up to multiple hundreds of thousands of points are measured in a 3D space per second, allowing for a very detailed point-cloud of the environment with sub-cm range measurement precision. This data can be used in a variety of applications from simple 3D obstacle avoidance (using live-data) to 3D site mapping (aggregated data), GNSS-independent vehicle localization and object classification to name a few.

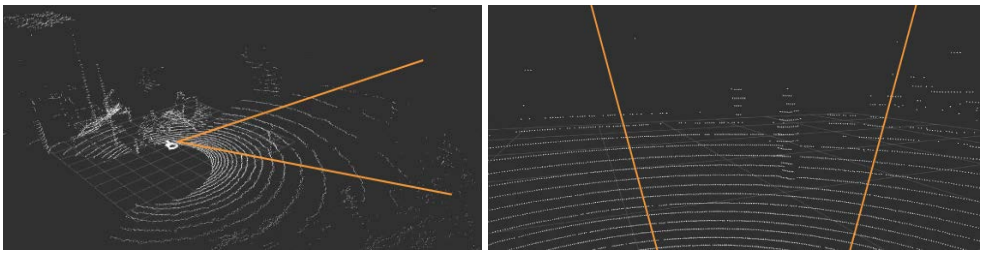


Figure 2: 3D lidar point cloud of the same scene as above with the approximate camera field-of-view highlighted.

Various other types of sensors exist that can be used for the task, such as 3D PMD cameras.
One of the most advanced commercially available solution today is [7]. Using a stereo-vision camera at the rear of the machine the system is able to both classify the detected obstacles into “pedestrians” and “other obstacles” and measure the relative position between the obstacle and the machine. Subsequently warnings for the operators are tailored by a) two different warning zone sizes (pedestrians up to ~7m, obstacles only ~2.5m) and b) different display visualizations and audible warnings depending on obstacle type. With that approach a high level of operator acceptance is achieved, since the vehicle can be driven close to obstacles in tight confines without unnecessary warnings while early warnings are given when pedestrians are present.
Since this particular system is not a retrofit option but a fully integrated OEM solution automatic speed reduction can be realized instead of just providing warnings to the operator. The clear benefit lies in even shorter stopping distances and more practical collision scenarios avoided since the reaction time of the operator is skipped and the machine brought to a stop earlier. At a typical reverse speed of 10 km/h and an estimated reaction time of the operator of 1 s the machine comes to a standstill 2.78 m sooner than with a pure active warning system.
Aside from the machine operator and the pedestrian in the danger zone immediately benefitting from this system in a potential collision scenario it is assumed that organizational deficiencies at the work site have to occur in order for the pedestrian to appear in the danger zone in the first place. This is also addressed by the system in that a data package with time stamp, vehicle speed, GNSS-position and other related data is sent by the machine into the OEMs IoT system. The end-customer can then use this information to bring transparency into the workforce (e.g. inform safety-responsible personnel about practical risk areas) and subsequently improve work processes by organizational measures in order to avoid pedestrian to vehicle convergence in the first place. Both individual incident markers as well as a site heat map representing pedestrian detection occurrence rates can be accessed via web interface.



Figure 3: Left: Wheel loader with smart pedestrian detection (approximate warning zones for obstacles and pedestrians highlighted). Right: Online incident map

4. COLLISION AVOIDANCE SYSTEM FOR EXCAVATORS BASED ON DIGITAL TWINS

The developed dynamic collision avoidance system for mobile excavators combines advantages of digital twin and transponder based approaches. In general, the system is able to collect data of the presented 3D-systems from different machines on a vendor independent platform in a local network on the construction site and connects them with planning data as well as site specific guidelines and regulations concerning protection zones. In contrast to the solutions mentioned in section 2.3, also accidents between different mobile machines can be prevented in this way.

4.1. System Architecture

The system architecture of the collision avoidance system can be seen in **Figure 4**. The core of this architecture is the so-called safety control station, which is a variation and simplification of the Site Execution System (SES) developed in the joint research project “Bauen4.0” [8] and is mainly based on the Site Information System introduced in ISO 15143-1 [9]. The safety control station is a micro-service-based software environment, which runs on a local Edge Cloud Server. Basic micro-services are located in a virtual machine environment with defined access rights.

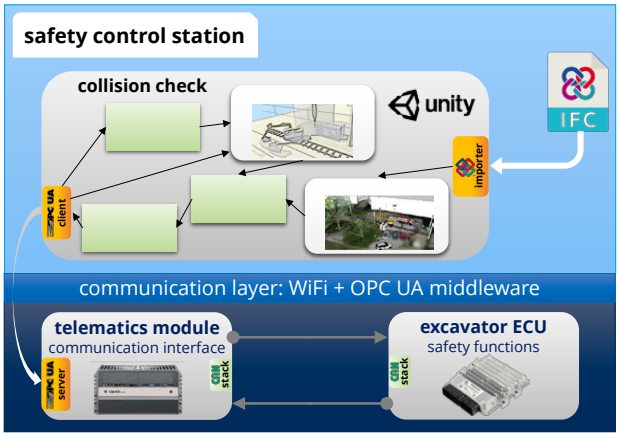


Figure 4: System Architecture of the proposed collision avoidance system

The core micro service is the so-called collision check, which represents the digital twin of a construction site. This digital twin, running on the cross-platform game engine unity, includes all static and dynamic information for predicting potential accidents, respectively collisions between mobile machines and their working environment. The construction site facilities including static collision bodies (based on German safety regulations and guidelines, e.g. DIN 18920) for facilities worthy of protection, e.g. trees, can be imported to unity via a plugin for BIM-files. The creation of the construction site layout and the static collision bodies in form of a BIM-file can be done in typical 3D design review packages like Navisworks. The investigated mobile machine, a hydraulic excavator, was directly built in unity in the aforementioned research project “Bauen 4.0” (see [10]) and extended by dynamic collision bodies. The function of the collision bodies and the collision prediction itself is explained more detailed in section 4.3. For displaying the mobile machine in real-time as well as calculating the collision bodies, unity includes an OPC UA Client, which collects the data of the real-world excavator respectively its 3D machine control systems. So the excavator can be displayed in the digital environment based on the GNSS position and the IMU sensor values. The OPC UA Client is also used for sending speed limits or emergency stops to the excavator in case of predicted collisions. The communication concept using the middleware OPC UA in combination with WIFI is

explained more detailed in following section 4.2. The planned but not implemented workflow actually envisages the mobile machinery to login to the control station and transmit an *.xml file in the urdf schema via OPC UA, which contains the complete kinematic description of the machine including the collision bodies.

4.2. Communication Concept

A middleware approach using OPC Unified Architecture (OPC UA) has been chosen for the communication on the construction site. This decision is based on the promising results of the aforementioned research project “Bauen 4.0”. Using the middleware OPC UA enables interoperable communication on the construction site between different participants, i.e. all kinds of mobile machinery, as well as various applications, e.g. the micro services of the safety control station. As stated by the OPC Foundation, OPC UA is “a platform independent service-oriented architecture that integrates all the functionality of the individual OPC Classic specifications into one extensible framework [11].” OPC UA typically includes protocol specifications from the third to the seventh OSI layer and can be combined with Ethernet (IEEE 802.3) or WIFI (IEEE 802.11).
The concept of “Bauen4.0” provides for each mobile machine to offer its information and services via an OPC UA server, which is hosted on a machine's own telematics module. The central control station on the construction site uses a so-called aggregating server, which aggregates the data from all mobile machines via OPC UA clients and acts as an interface to the individual applications. A 5G campus network was used in the research project to enable an IP-based network on the construction site. The offered information of the mobile machinery include mainly telematics data defined in ISO 15143-3 [12], as well as the description of the kinematic tree for monitoring mobile machinery during operation and hydraulic condition monitoring data. These data sets had been formalized into information models based on existing Companion Specifications. For further information regarding the concept as well as developed information models, have a look at [13], [14].

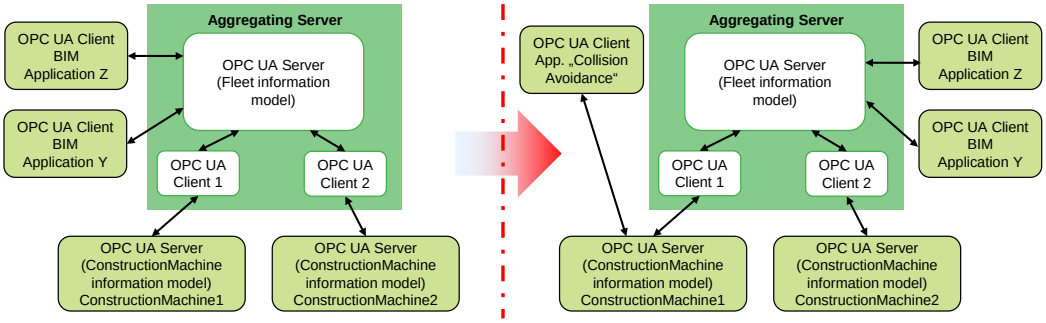


Figure 5: Adaption of the Bauen 4.0 communication concept; left: Bauen 4.0 concept; right: communication concept for the safety control station

For reasons of efficiency and decreasing communication delays, the concept was simplified. The adapted version is shown in **Figure 5**. The aggregating-server is bypassed by connecting the OPC UA client of the application, hence the collision check, directly to the OPC UA server of the mobile machine. In addition, the campus network was dispensed and WIFI 6 (IEEE 802.11ax) was used instead. The application requires the angles of the attachment and the global position of the excavator by reading corresponding OPC UA variables. On the other side, the application has to set the speed limits as well as the status regarding the actuators by writing corresponding OPC UA variables. These speed limits are sent afterwards via CAN by the machines telematics module to the Excavator ECU, which controls the excavator and takes them accordingly into account (compare **Figure 4**).
While the desired information is already covered by existing information models and therefore

available by the OPC UA server on the used excavator, speed limits and status information had to been integrated into the information models. Therefore a new *ObjectType*, called *TaskControlTypeExtension*, had been developed, which compromises this information, compare **Figure 6**. The Type is based on the *TaskControlType* of the OPC UA Companion Specification for Robotics [15] and adapted to the use-case.

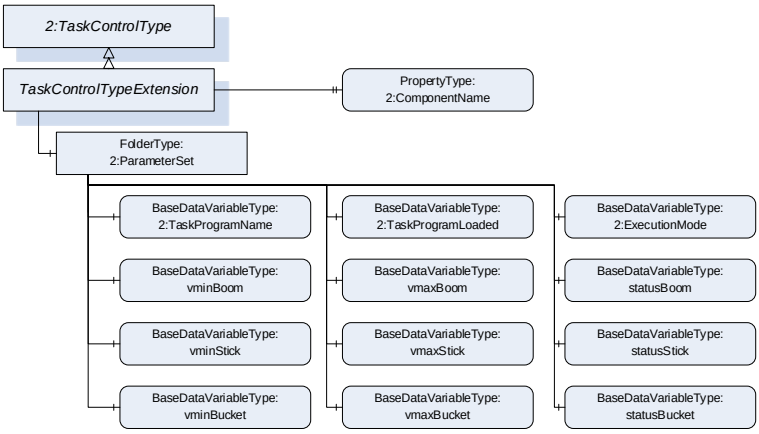


Figure 6: Visualisation of the ObjectType TaskControlExtensionType

Reading and writing OPC UA variables by the client has been realised by using standard OPC UA read and write requests of the implementation open62541. In general, each read or write request of every variable causes two TCP Acknowledgements (TCP ACK). For reducing network traffic and therefore potential time delay, all read requests have been integrated into one request (Batch Request) and all write request vice versa. The communication loop of the OPC UA client is shown in the sequence chart in **Figure 7**. Read and write requests are executed one after the other as described, whereas acknowledgements follow each request and response.

Furthermore, concerning the client/server pattern, it would have been possible to switch to *Subscriptions* by specifying desired nodes in form of so-called *MonitoredItems*. However, an additional wait cycle, the publish interval, is integrated into the communication cycle within this method. Therefore, this variant is not suitable for low-delay communication and has not been considered in this investigation. Another option would have been to switch to publish/subscribe patterns. This would eliminate the TCP ACK, which would presumably lead to performance advantages. However, this approach was not pursued further because of the good results regarding batch requests in section 5.1.

In this investigation, as mentioned, WIFI 6 (IEEE 802.11ax) was used to provide an IP-based network without the need of physical wiring. For evaluating the communication concept and the network, different tests with various test setups have been carried out, see section 5.1.

In general, it has to be stated, that in case of using the collision avoidance system completely independent from the operator as an automation function, the communication concept does not meet required safety standards. Such a use would require safety functions implemented on the excavator ECU, which set the machine into a safe-state in an event of disconnection or other errors. There are two prerequisites for this, which cannot be met currently. On the one hand, the whole communication has to be in accordance with the requirements of DIN EN IEC 61784-3 [16] and DIN EN IEC 61508 [17]. In fact, the OPC UA Core Specification “OPC 10000-15: UA Part 15: Safety” [18] defines mechanisms for the transmission of safety-relevant messages among participants within a network using OPC UA technology fulfilling these stated standards. Yet, there are no OPC UA implementations available, which fulfil this core specification. Reliable information concerning the

publication of such stacks are not accessible, compare [19]. On the other hand, the OPC UA server of the mobile machine, which is an integral part of the excavator control loop, has to be implemented on a device, which is able to fulfil the applications specific required Safety Integrity Level (SIL, see IEC 61508). Simply using a safe communication is not sufficient to qualify a safety device, compare [20]. Commercially available telematics modules for mobile machinery are not able to fulfil any SIL. Commercially available ECUs, which are able to fulfil SIL-2, cannot be combined with typical OPC UA stacks. In this investigation, the OPC UA server was therefore implemented on an industrial PC, spectra powerbox 410-I5, using Ubuntu 22.04 LTS as operating system.

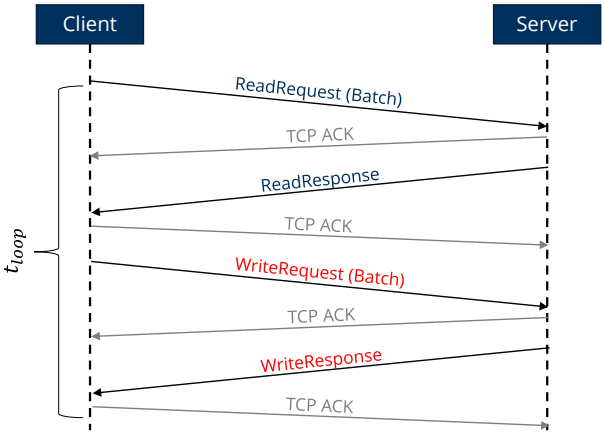


Figure 7: Communication loop OPC UA client

4.3. Collision Detection

As mentioned in section 4.1, the collision check is realized in unity, which offers mechanisms for recognising collisions between bodies using so-called ‘colliders’. In order to utilise these functionalities and prevent collisions in the real world, static and dynamic bodies, so-called collision bodies, have to be generated. The intention is to use these elements for predicting potential collisions between mobile machinery and the construction site facilities and deriving permissible machine control signals in order to ensure collision avoidance. In general, the collision check is able to consider the rotating movement of the upper carriage and all possible movements of the working attachment of a hydraulic excavator. Driving is currently not taken into account and has to be implemented in future for extending the functionality of the collision avoidance system.

As already mentioned, a distinction is made between static and dynamic collision bodies. This differentiation is necessary for depicting static and moving objects. Static bodies, as shown on the left side in **Figure 8**, represent protection zones and include all fixed objects of the construction site as well as existing infrastructure and vegetation. The critical areas are defined in accordance with the latest standards and guidelines. An overview of the existing regulations is given in **Table 1.** The zones ensure the protection and integrity of existing structures, mobile machines are not allowed to harm.

The static collision bodies are created manually and placed in the georeferenced digital construction site environment. Simple geometric bodies such as cuboids, cylinders and spheres in accordance with the applicable specifications represent the shape of these bodies. Therefore, an automatic workflow for generating these bodies based on a database including the standards, should be easy to implement in future.

In contrast to static, non-adaptive collision bodies, dynamic bodies are variable in time and represent the permitted motion space of dynamic objects. Their main purpose is to monitor the motion space of

mobile machinery and detect potential collisions, so that measures can be taken to prevent accidents. For creating dynamic collision bodies, the first step is to approximate the machine geometry using simple surfaces. In case of the chosen demonstrator, a hydraulic excavator, the geometry is approximated by rectangular surfaces. It is important that these surfaces include all components of the machine, even small attachments, such as working lights. In simple terms, dynamic collision bodies represent the breaking distance of mobile machines respectively their components. They adapt to the actual movement of the machine in order to achieve a good balance between safety, operability and productivity. This ensures that the machine is only slowed down when necessary, minimising 'false positive' errors. By analysing the dynamical behaviour of the real machine, parameters were derived for calculating the breaking distance and stored in the digital twin.

Table 1: Regulations for the definition of protection zones on static bodies concerning infrastructure sites

Regulations	Description
Eisenbahn-Bau-Betriebsordnung (EBO)	EBO defines construction and operation railways in Germany.
DIN 18920	DIN 18920 describe the protection of trees and vegetation areas.
ASR A5.2	ASR A5.2 describes requirements for work-places and traffic routes on construction sites.

The dynamical behaviour of the machine was comprehensively analysed experimentally, focusing on the individual actuators, which means the hydraulic cylinders. Several measurements were carried out in which the actuators were subjected to square-wave signals, so that the cylinders were retracted and extended in various load situations. By evaluating these measurements, it could be shown, that the dynamical behaviour of all the actuators can be approximated by simple linear decelerating respectively accelerating equations based on the actual cylinder velocity $\dot{x}_{cyl,i}$, compare following equations:

$$\Delta x_{break,i} = \frac{a_{break,i}}{2} \cdot {t_{break,i}}^2 + \dot{x}_{cyl,i} \cdot t_{break,i} \tag{1}$$

$$\Delta x_{break,i} = \frac{3}{2} \cdot \frac{{\dot{x}_{cyl,i}}^2}{a_{break,i}} \tag{2}$$

The breaking distance $\Delta x_{break,i}$ of an actuator of the digital twin, is calculated with the help of the experimental determined breaking acceleration factors $a_{break,i}$. As each actuator is attached to the working attachments, these translatoric cylinder movements are transferred to rotational movements of the members of the working attachment, such as boom, adjust boom or stick. The relationship between the translatoric cylinder position $x_{cyl,i}$ and the rotatory angle of the transmission element φ_i can be derived as follows for each cylinder, respectively transmission element:

$$G_i = \frac{\dot{\varphi}_i}{\dot{x}_{cyl,i}} \approx \frac{\Delta \varphi_i}{\Delta x_{cyl,i}} \tag{3}$$

In consequence, the necessary breaking angle $\Delta \varphi_{break,i}$ can be calculated for each component of working attachments according to equation 4:

$$\Delta \varphi_{break,i} = \frac{3}{2} \cdot \frac{{\dot{x}_{cyl,i}}^2}{a_{break,i}} \cdot G_i \tag{4}$$

This breaking angle is ultimately used for discharging the dynamic collision bodies in the direction of the movement. **Figure 8** shows an example of the dynamic collision bodies of the excavator model

on the right side, which are calculated using the described method.

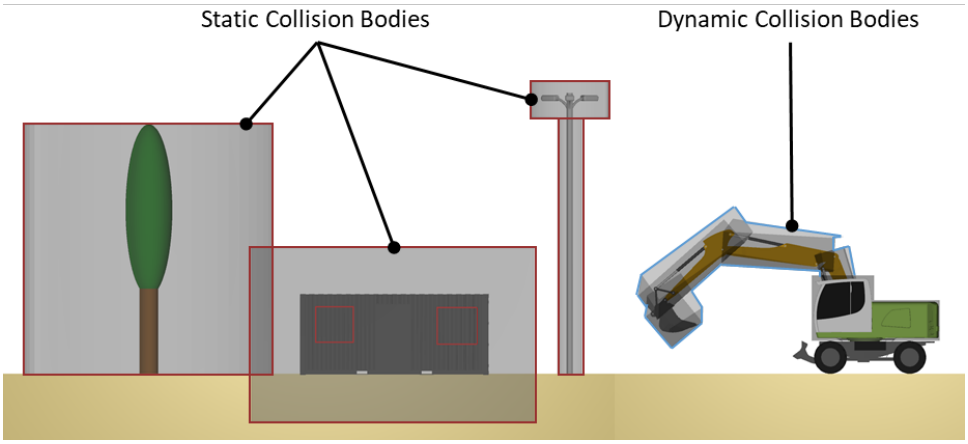


Figure 8: Static Collision Bodies (Left) and Dynamic Collision Bodies (Right)

As already mentioned, the development environment uses so-called collider components for collision detection between physical bodies. These components define the geometric shape and size of a specific body so that the integrated physics engine can recognise collisions based on the collider geometries. Collision detection is performed by cyclic checks of the collider positions and shapes in each unity frame cycle. The tests in chapter 5 were due to the use of storage and computing intensive BIM-files for the virtual construction environment performed with fixed frame cycle times t=20 ms. From a safety point of view, fixed cycle times < 10 milliseconds should be strived for. As soon as an overlap or contact between two collider geometries is detected, a collision is recognised by the engine, which leads to recording and storage of data such as the collision status and the collision points. By using so-called collision masks, it is possible to individually define collision pairs, which means two colliders, where contacts shall be detected. By default, all other collisions are ignored. In this specific context, only collisions between static and dynamic collision bodies are handled. All other possible collision pairs, such as collisions between static colliders, are deliberately excluded in order to utilise the computing power of the engine in a targeted and efficient manner by avoiding unnecessary calculations.

4.4. Intervention in the machine control

The described collision results in the previous section are the basis for determining permissible interventions of the machine control system. The primary aim is reducing the speed of the excavator actuators respectively stopping them in case of critical movements. In addition, in the event of a collision, the system must always allow the operator to move the attachment out of the static collision zones if a collision occurs again. The permissible control signals of the mobile machine are generally determined according to a predefined sequence diagram, which is shown in **Figure 9**.
In the first step, various pieces of information, such as collision status of each dynamic collision body and current control signals are aggregated. Afterwards these information are checked, whether collisions have occurred. In case of no recognised collision, no further measures are taken. If a collision is detected, the system limits the permissible speed of each actuator, which can lead to the collision. If, for example, a possible collision of the excavator bucket with an overhead line is detected by the upward movement of the excavator arm, all actuators that could lead to a collision, i.e. boom,

stick and bucket cylinders, are limited.

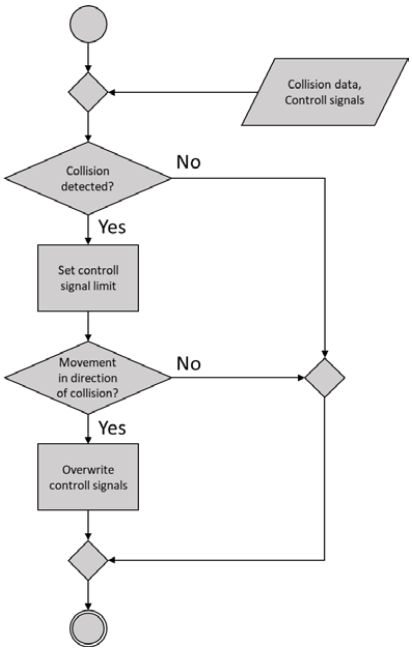


Figure 9: Sequence diagram of machine control intervention

Thereby the system takes the direction of the actuator movements in account. The actuators are not completely blocked for operation, yet only the retraction or extension speeds of the cylinders. At the end of the sequence diagram, the permissible control signals are stored in a data struct, which the OPC UA client thread can access and send them to the mobile machinery, compare section 4.2.
Furthermore, in case of an actual collision, so that all actuator movements (retraction and extension of a cylinder) lead to further collision, a so-called emergency motion state is triggered after half a minute. As mentioned above, this shall allow the operator to move the attachment back out of the static collision zones again.
In contrast to the collision avoidance in section 2.3, the selected procedure does not deactivate the entire machine control system when the machine approaches a safety zone, but the operator is still able to effectively counteract the dangerous movement of the machine, thus increasing operating comfort, safety, acceptance by the driver and productivity.

5. SENSITIVITY AND ROBUSTNESS ANALYSIS

For evaluating the sensitivity and the robustness of the developed collision avoidance system, two kinds of tests were performed. On the one hand, the communication concept is evaluated by measuring the minimum achievable cycle times of the OPC UA Client. On the other hand, the accuracy of the whole system is evaluated by measuring the distances between excavator bucket and protected facilities in two different scenarios. All tests were done with the OPC UA Server running on an industrial PC, spectra powerbox 410-I5 (i3-8100T, 8 GB RAM, Intel I210, Intel Wi-Fi AX200, Ubuntu 22.04.3 LTS) inside the mobile excavator, and the OPC UA Client running on a PC (i7-1355U CPU, 16 GB RAM; Realtek USB GbE Family Controller, Windows 11), which represents the control station for these tests. The OPC UA Client is integrated in unity and performed in a separate thread as fast as possible. For wireless connection an Ubiquiti Dream Router was used with WIFI 6

(IEEE 802.11ax).

5.1. Communication Tests

The general test layout is shown in **Figure 10**. The excavator is located approx. 9.5 m apart from the control centre and approx. 3.6 m to the WIFI router, which is standing on a platform (3 m). The following setups were selected:

- **Ethernet:** Mobile machine and control station are connected directly via Ethernet (IEEE 802.3).
- **WIFI:** Industrial PC of excavator is connected via WIFI (IEEE 802.11ax) to control station.

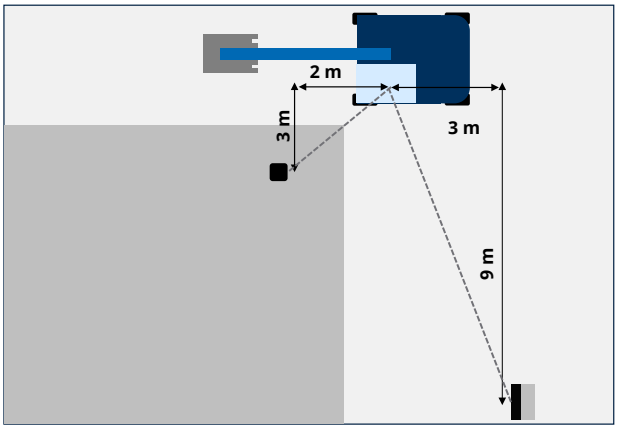


Figure 10: General test layout.

The following tests were performed:

- **Standard:** Reading and writing (sample size: 1000) the in total 31 variables by the OPC UA client with standard read and write methods of open62541.
- **Batch:** Reading and writing (sample size: 1000) the in total 31 variables by the OPC UA client with batch read and write methods of open62541.

The results of the performed tests can be seen in **Table 2** and **Figure 11**. In general the experiments show, that the cycle time can be significantly reduced by using the single batch request instead of standard write and read request, independent from the physical layer. Furthermore it is clear, that reliable communication (reliable cycle times<<<10ms) is only possible with the use of Ethernet at the moment, compare especially the standard deviation of Ethernet and WIFI. These experiments also show that the variating communication time with WIFI is likely to be a major influencing factor on the measured distances in the following section 5.2. In future, it would be very interesting to carry out these tests again with secure OPC UA communication according to OPC UA Core Specification “OPC 10000-15: UA Part 15: Safety” [18], compare section 4.2.

Table 2: Results of the communication tests

nr	setup	com. pattern	mean [ms]	max [ms]	min [ms]	std [ms]
1	Ethernet	Standard	19.6	55.8	9.50	4.1
2	Ethernet	Batch	1.1	3.6	<0.1	0.5
3	WIFI	Standard	87.9	874.6	53.2	61.9
4	WIFI	Batch	5.9	256.4	2.5	11.4

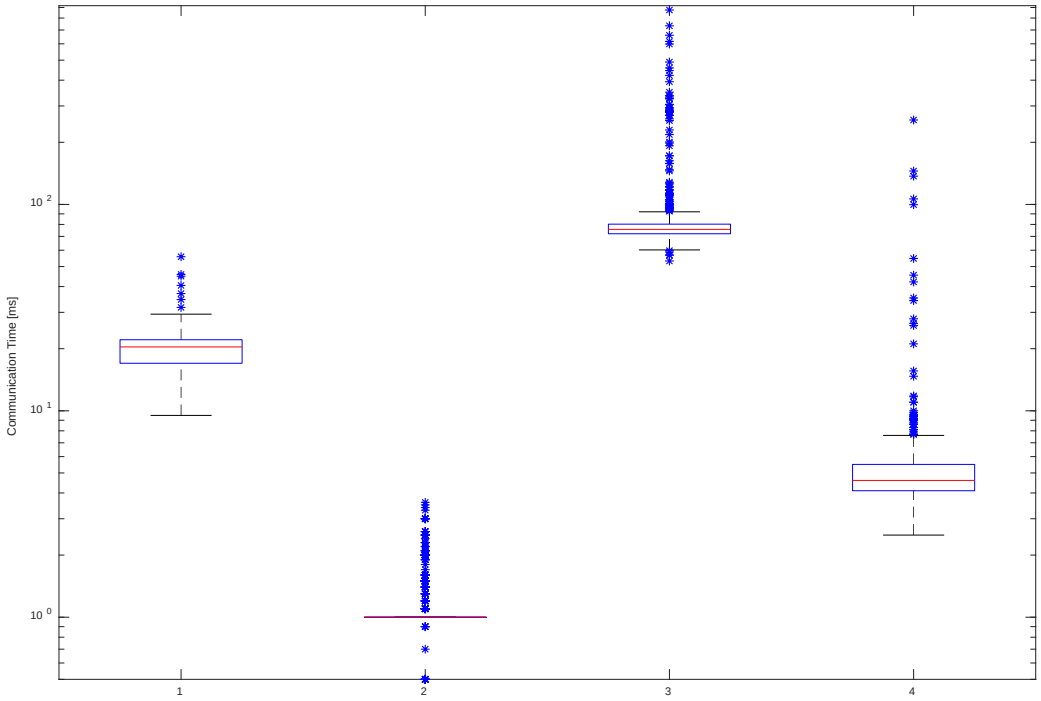


Figure 11: Box plot of the measured communication times for the configuration 1-4 displayed in **Table 2**.

For additional classification of the results standard ping test (sample size: 100) for measuring the round-trip time (RTT) for messages between host (control station) and destination computer (excavator) were performed, compare **Table 3**. The results show also, that using WIFI, there is a high risk that single communication cycles exceed the desired cycle time vastly and may lead to large deviations with regard to the predicted machine stopping points.

Table 3: Results of the ping tests

setup	test	mean [ms]	max [ms]	min [ms]	std [ms]
Ethernet	Ping	0.2	4	<0.1	0.5
WIFI	Ping	233.5	484	4	134.4

5.2. Collision Tests

In order to evaluate the accuracy of the overall system, collision tests are carried out focussing on a practical track construction scenario with the use of two-way excavators. These machines are mainly used on track construction sites or in track construction, where overhead lines are usually also present. Despite the presence of protective equipment, collisions between the machines and the overhead lines occur repeatedly, often with health and financial consequences. In the selected scenario, the boom cylinder of the excavator is extended so far that the excavator bucket or its dynamic collision element collides with the static collision element of the overhead line of a railway line (in Germany at a height of 5.5 m), as shown in **Figure 12**. A platform was set up next to the real machine, from which the minimum distance between the excavator bucket and the overhead line was measured at a height of 5 m using a laser rangefinder (Bosch DLE 50). According to legal regulations, the safety distance between the machine and the overhead line must be 0.6 m, which is defined as the target value. The measured values are compared with this target value in order to assess the effectiveness of the system in preventing collisions. The general communication structure corresponds to the previous section 5.1

(compare **Figure 7**), whereby WIFI and the batch method are used for communication. The test is repeated 25 times (n = 25) in order to be able to reliably record the various mostly stochastic influencing factors on the collision test statistically. In general, the following disturbance variables must be taken into account:

- positioning error of the GNSS system
- sensor errors of the working attachment
- varying dynamics of the mobile machine
- varying cycle times of the communication loop

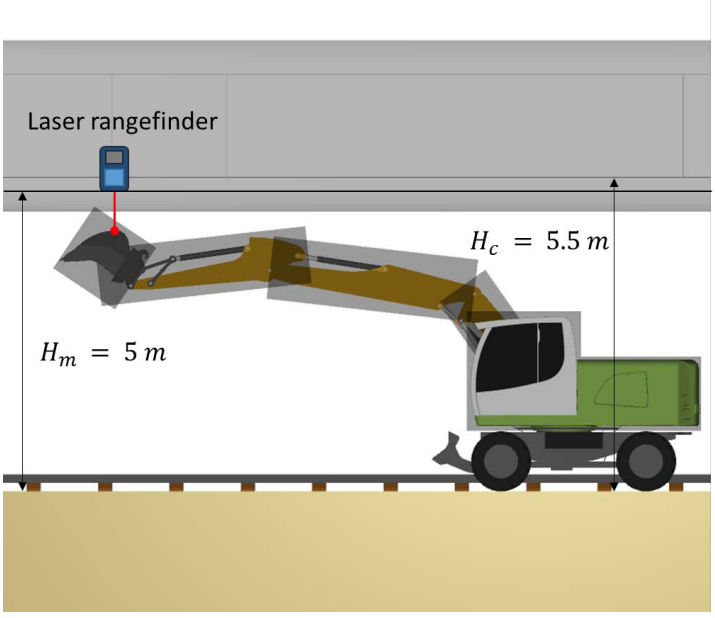


Figure 12: Test setup collision tests

Regarding the specific scenario, the positioning error of the GNSS system can be neglected due to the fact that the excavator is always positioned on the surface in unity and therefore the height deviation of the GNSS system does not disturb the calculation. In the context of a thesis at the TU Dresden [21] the accuracy of the used position sensors (IMUs) has been analysed. Over various poses of the excavator max. deviations regarding the TCP of the bucket in one spatial direction of 88 mm could be detected. Regarding the varying dynamics of the machine, no reliable conclusions can be drawn in respect to its influence on the breaking distance of the machine. For estimating the influence of varying cycle times, the communication loop and the frame times were measured during the tests. **Table 4** shows the results.

Table 4: Results of the collision tests

	mean [m]	max [m]	min [m]	std [m]
distance	0.9094	1.0680	0.5930	0.1243

The measurement results of the tests carried out show that the achieved average distance to the overhead line is 0.9094 m, which is approx. 0.3 m above the legal requirements. In terms of occupational safety, the required minimum safety distance can always be maintained. However, the system restricts the productivity of the machine by reducing the usable working space by approx. 20 %. This means, that the used parameters for the calculation of the collision bodies are at the moment way to conservative. Especially concerning the improved communication loop, safety factors

concerning the calculating of braking distances can therefore be reduced.

6. SUMMARY AND OUTLOOK

This paper presented solutions for collision avoidance between objects, respectively pedestrians on the construction site and construction machinery to increase site safety. Solutions with on-board perception sensors are already being used for detecting people. Also tracker-based solutions are used today to prevent collisions between mobile machinery and workers. The use of digital twins and reliable radio technology opens up new possibilities for collision detection and avoidance. While conventional collision avoidance systems based on 3D machine control systems only focus on collisions between mobile machinery and static objects of a digital construction site, the presented solution allows in general the integration of other machines or construction workers. Using a track construction scenario, it was shown that a safety control station, where the digital twin of a two-way excavator moves almost in real time in a georeferenced environment imported from BIM tools, allows safety distances to be maintained depending on the situation and in accordance with standards. This offers a real alternative to solutions with environmental sensors. As all mentioned technologies have their advantages and disadvantages, these are summarized and compared in the following table.

Table 5: Pros and cons of the different approaches

	on-board perception sensor solutions	**tracker-based solutions**	**conventional digital twin solutions**	**new digital twin approach**
pedestrian detection	✓ independent of infrastructure or organizational measures / setup ✓ affordable and proven technology ✓ no reliance on GNSS (fully functional in and near buildings, etc.) ✓ high operator acceptance (few unnecessary warnings) - affected by harsh environmental conditions (dirt, fog, dust, etc.)	✓ affordable and proven technology ✓ no absolute reliance on GNSS (functional in and near buildings, etc.) ✓ barely affected by harsh environmental conditions (dirt, fog, dust, etc.) - organizational measures (pedestrians have to be hardware-equipped) necessary	- not supported at the moment	✓ barely affected by harsh environmental conditions (dirt, fog, dust, etc.) ✓ reliable object classification - organizational measures (pedestrians have to be hardware-equipped) and infrastructure (site network, etc.) necessary - reliance on RTK-GNSS
collision avoidance (static & dynamic obstacles)	✓ independent of infrastructure or organizational measures / setup	✓ barely affected by harsh environmental conditions (dirt, fog, dust, etc.)	✓ barely affected by harsh environmental conditions (dirt, fog, dust, etc.)	✓ barely affected by harsh environmental conditions (dirt, fog, dust, etc.)

	✓ affordable and proven technology ✓ no reliance on GNSS (fully functional in and near buildings, etc.) - expensive onboard sensor setup necessary (e.g. lidar for wire detection) - affected by harsh environmental conditions (dirt, fog, dust, etc.)	✓ no absolute reliance on GNSS (functional in and near buildings, tunnels, etc.) - organizational measures (all objects have to be hardware-equipped) necessary	✓ reliable object classification - organizational measures (digital 3D-map) necessary - reliance on RTK-GNSS - collision with unknown obstacles	✓ reliable object classification - organizational measures (digital 3D-map) and infrastructure (site network, etc.) necessary - reliance on RTK-GNSS - collision with unknown obstacles

ACKNOWLEDGEMENTS

This Project is supported by the Federal Ministry for Economic Affairs and Energy (BMWi) on the basis of a decision by the German Bundestag.

Supported by:

Federal Ministry for Economic Affairs and Climate Action

on the basis of a decision by the German Bundestag

ZIM Zentrales Innovationsprogramm Mittelstand

REFERENCES

[1] Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Regeln zum Arbeitsschutz auf Baustellen (RAB), https://www.baua.de/DE/Angebote/Rechtstexte-und-Technische-Regeln/Regelwerk/RAB/RAB.html, visited on 12.01.2024.

[2] Risch, M., Arbeitsschutz und Arbeitssicherheit auf Baustellen, Springer Fachmedien Wiesbaden, 2016.

[3] Deutsche Gesetzliche Unfallversicherung, Meldepflichtige Arbeitsunfälle je 1.000 Vollarbeiter nach Bereich und Berufsgenossenschaft, https://www.dguv.de/de/zahlen-fakten/au-wu-geschehen/au-1000-vollarbeiter/index.jsp, visited on 12.01.2024.

[4] Mesafox, SICHERHEIT. DIGITALISIERUNG. Für Baustellen & Industrie mit dem mesafox PLINX System, https://plinx1.com/de/startseite/, visited on 12.01.2024.

[5] Ministry of Business, Innovation & Employment New Zealand, Case Study: KiwiRail - Transforming New Zealand's rail network with technology, https://www.constructionaccord.nz/good-practice/beacon-projects/case-study-kiwirail/ visited on 13.01.2024.

[6] Leica Geosystems, 3D Avoidance Zone system, https://leica-geosystems.com/de-de/products/machine-control-systems/awareness-solutions/3d-avoidance-zone-system, visited on 12.01.2024

[7] Liebherr, Assistance systems for wheel loaders, https://www.liebherr.com/shared/media/construction-machinery/earthmoving/brochures/wheel-loader/liebherr-assistance-systems-for-wheel-loaders.pdf, visited on 12.01.2024

[8] Beck, B. et al., Lösungen für die vernetzte und automatisierte Baustelle, ATZ heavyduty, 04/2022.

[9] ISO, ISO 15143-1 - Earth-moving machinery and mobile road construction machinery - Worksite data exchange Part 1: System architecture, 2010

[10] Köhler, S. et al., Netzwerkschnittstellen für mobile Arbeitsmaschinen im Kontext der digitalisierten Baustelle, 8. Fachtagung Baumaschinentechnik, Dresden, 2020.

[11] OPC Foundation, OPC Unified Architecture, https://opcfoundation.org/about/opc-technologies/opc-ua/, visited on 12.01.2024.

[12] ISO, ISO 15143-3 - Earth-moving machinery and mobile road construction machinery - Worksite data exchange Part 3: Telematics Data, 2010.

[13] Köhler, S., Weber, J., Integration of fault detection and diagnosis methods into OPC UA for mobile machinery, 13th International Fluid Power Conference, Aachen, 2022.

[14] Köhler, S. et al, Bauen 4.0: Vernetzter, automatisierter Tiefbau am Beispiel Mobilbagger, 10. Fachtagung Baumaschinentechnik, Dresden, 2022.

[15] OPC Foundation, OPC UA Companion Specification for Robotics (OPC Robotics) Part 1: Vertical integration, 2019.

[16] DIN EN IEC, 61784-3:Industrielle Kommunikationsnetze - Profile - Teil 3: Funktional sichere Übertragung bei Feldbussen - Allgemeine Regeln und Festlegungen für Profile, 2017

[17] DIN EN IEC, 61508-1: Funktionale Sicherheit sicherheitsbezogener elektrischer/elektronischer/programmierbarer elektronischer Systeme - Teil 1: Allgemeine Anforderungen, 2011

[18] OPC Foundation, OPC Unified Architecture Part 15: Safety, 2021.

[19] OPC Foundation, Safety, https://opcfoundation.org/markets-collaboration/safety/, visited on 12.01.2024.

[20] OPC Foundation, Relation to safety-, security- and OPC UA-standards, https://reference.opcfoundation.org/Safety/v104/docs/2.2, visited on 12.01.2024.

[21] Taesch, J, Entwicklung eines Konzeptes zur dynamischen Kollisionsvermeidung mobiler Arbeitsmaschinen auf digitalen Baustellen, TU Dresden, 2022

A COMPARISON OF STATE-OF-THE-ART NETWORK ARCHITECTURES FOR INSTANCE-SEGMENTATION IN FOREST ENVIRONMENTS.

Lukas Michiels[1]*, Manuel Westermann[1], Benjamin Kazenwadel[1], Chris Geiger[2], Marcus Geimer[1]

[1] *Karlsruhe Institute of Technology / Institute of Mobile Machines, Rintheimer Querallee 2, 76131 Karlsruhe, Germany*
2 Hohenloher Spezial-Maschinenbau GmbH (HSM), Im Greut 10, D-74635 Neu-Kupfer

* Corresponding author: E-mail address: lukas.michiels@kit.edu

ABSTRACT

Research and development have increasingly focused on automating mobile machines to reduce the negative influence of labor shortages and high labor costs. Object detection is a key requirement for the automation of mobile machines. The transfer of the developed methods to the environment of mobile machines, e.g. a forest, a building site, or in mining, is challenging. Objects of the same class can have significantly different phenotypes and the surroundings cannot be controlled, weather as well as lighting conditions can change. Neural networks are the state-of-the-art method for detecting and classifying objects for image sensors. The required datasets as well as network architectures mastering object detection across different forest areas have not yet been presented. We collected two datasets, *MobimaWoodlands* and *MobimaSkidRoads*, one with a handheld camera and one captured while driving on skid roads in different areas and in different seasons. Three network architectures for the instance segmentation with two different backbones were trained on the two datasets to segment stems, trees, and stumps. In a subsequent step, the trained networks were evaluated on two public datasets which have not been used in the training process. With an adapted training pipeline, we achieved a similar accuracy with a slight decrease in the AP of 0.1 on one of the unknown datasets with similar tree specimens. On the second unknown dataset, the AP decreased more significantly by 0.3. The findings highlight that generalization over various forest areas is possible, even in demanding outdoor settings. However, the portability to unknown domains cannot be guaranteed especially if different tree species are present, which continues to be an issue in many applications.

Keywords: Object Detection, Forest, Neural Networks, Instance Segmentation

1. INTRODUCTION

Object detection is one of the key requirements for the automation of mobile working machines [1]. In recent years, neuronal networks have become the state-of-the-art approach [2]. Object detection tasks are generally divided into three subdomains: object classification, semantic segmentation, and instance segmentation. Object classification generally provides knowledge of whether an object of a specific class is present in the picture; semantic segmentation provides the class for every object; and instance segmentation separates each instance of the same object class [3].

Various works have already evaluated partial aspects of instance segmentation with neural networks in forest environments, such as object detection or instance segmentation of individual object categories. Two main challenges remain regarding instance segmentation in forest environments. The distinction between individual instances and the background (stuff, [4]) is often ambiguous. In many

cases no distinct differentiation between individual trees (foreground) and trees in the background (stuff) is possible. Additionally, most research in the computer vision area is focused on the COCO Dataset ([5]) and only a handful of public datasets for forestry environments are available [6], [7]. Datasets are key aspects of object detection, especially in environments with many varying objects. An optimal algorithm can recognize all possible variations of the specific objects (perfect generalization). When the evaluation dataset deviates from the training dataset, e.g., if a tree species is not included in the training dataset, the algorithm needs to be able to extrapolate the features of the object class (e.g. trees) to recognize other possible variations. Encountering data outside the training dataset's scope (out-of-distribution data) is inevitable in forest environments. A dataset featuring four seasons, sunny and cloudy days, coniferous, deciduous, and mixed forests for three different age classes already has at least $4 \cdot 2 \cdot 3 \cdot 3 = 72$ different parameter combinations. In integrating different tree species and additional objects, the number of parameter combinations increases rapidly, and a huge number of images are already required to cover the parameter space, even without considering the different phenotypes of the same tree species. The same issue comes up regarding the bias of different object classes. Generating a forest dataset with an equal number of human instances and tree instances is unfeasible. In consequence, object recognition for autonomous mobile working machines requires robustness regarding biased and incomplete training datasets.

In recent years, many new architectures have been presented. Mostly the improved accuracy comes with increased complexity demanding higher computational power. In the environment of mobile machines, transferability to unknown areas is often more important than accuracy on one collected training dataset. Computational power is costly, and it is in dispute whether increasing the complexity is worthwhile for the application in mobile machines. In this paper, we compared three state-of-the-art instance segmentation architectures on two new datasets, *MobimaWoodlands*, and *MobimaSkidRoads*, as well as their performance when applied to a completely unknown dataset from a different area. The validation dataset consisted of images that had been separated from the dataset and had not been used in the training of the network. However, they share the same tree variations as already encountered in the training data set. In contrast to the validation data, we focused on the transferability of the trained networks to an unknown forest area with different tree variations and a different environment. For this purpose, we used the public *FinnWoodlands* and *CanaTree100* datasets [6], [7]. The implemented training pipeline focus on transferability of the trained models instead on maximizing the mean average precision (mAP) over a single dataset.

The paper is structured as follows: First, the state if the art is reviewed in Section 2. Section 3 presents the dataset and the object classes. In Section 4 the setup and the training pipeline are described. Section 5 evaluates the training results and transferability to other datasets before Section 6 concludes the paper and gives an outlook on further approaches.

2. STATE OF THE ART

2.1. Evaluation Metrics

Several evaluation metrics are known to evaluate the quality of the predictions generated by the instance segmentation algorithms [8]. We focus on the Intersection over Union (IoU) and the Average Precision (AP).

Intersection over Union

The intersection over union (IoU) is a metric of the accuracy of individual bounding boxes and masks. This metric, also known as the Jaccard index, is calculated from the predicted label B_{pr} and the ground truth B_{gt} according to the following equation [9]:

$$IoU = \frac{B_{pr} \cap B_{gt}}{B_{pr} \cup B_{gt}} \tag{1}$$

A predicted mask/bounding box was a correct detection, if its IoU with the ground truth was higher than a specified threshold. The IoU threshold is given as percent, e.g. AP_{50} for an $IoU > 0.5$

Average Precision and Recall

Based on the IoU with regard to the ground truth, the following cases were separated:

- True positive (TP): Correct detection of an object.
- False negative (FN): Not detected existing (labeled) object.
- False positive (FP): Incorrect detection of an object that does not exist or a misaligned detection of an existing object.

The true-negative case of a correctly undetected annotation is omitted in the context of object detection due to the large number of theoretically possible bounding boxes that should not be detected.

Based on these cases, the precision and the recall are defined as:

$$Precision = \frac{TP}{TP + FP}, \qquad Recall = \frac{TP}{TP + FN} \tag{2}$$

The precision is a metric for the performance of the network with regard to the detected objects. The recall is, on the other side, a metric for the performance with regard to all existing objects. Both the precision and the recall depend on the IoU threshold for true positives. If the IoU threshold is small, the recall is typically larger while the precision decreases. The precision-recall curve is constructed from all recall and precision values for an IoU threshold from zero to one. The interpolated area under the precision-recall curve for a given IoU threshold is denoted by Average Precision (AP) [8]. The $AP_{0.5:0.95}$ is the average of the AP for the IoU thresholds from 0.5 to 0.95, $IoU \in \{0.50, 0.55, \ldots, 0.95\}$ [9]. The mean Average-Precision (mAP) is the arithmetic mean of all AP values of each class.

2.2. Network Architectures

Three different network architectures, Mask R-CNN, Cascade Mask-RCNN, and Mask2Former were regarded in the following work and explained in detail. The Mask R-CNN architectures are two-stage architectures with a region proposal network and a region of interest head. The Mask2Former architecture on the other hand is a single-stage transformer-based architecture. All architectures require a backbone network for feature recognition. In the state of the art, further architectures were used in forest environments, e.g., Rotated Mask R-CNN, Yolact++ and EfficientPS which will not be discussed in detail.

Mask R-CNN

Mask R-CNN ([10]) is a two-stage detector architecture developed as an extension of the original Faster R-CNN architecture [11]. The first stage consists of the backbone (e.g. a fully convolutional network) and a region proposal network, which determines the regions of interest. In the second stage, the algorithm predicts the class, a bounding box, and a binary mask for each region of interest. In addition to the Faster R-CNN architecture, Mask R-CNN has additional convolutional layers for mask generation. Figure 1 depicts an overview of the Mask-RCNN architecture. In the second stage, all regions of interest can be evaluated in parallel.

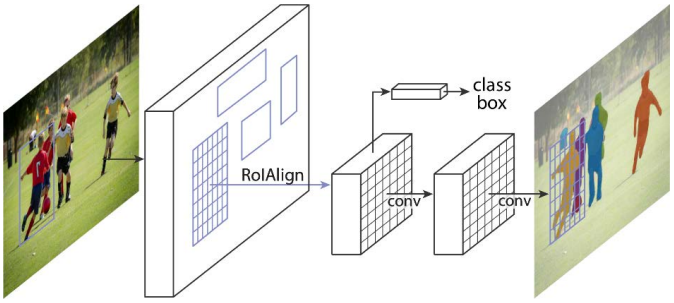


Figure 1: The Mask R-CNN framework for instance segmentation [10].

Cascade Mask-RCNN

Cascade Mask R-CNN is an extension to the original Mask R-CNN framework to improve high Intersection over Union (IoU) threshold detections [12]. Cascade Mask R-CNN is a two-stage architecture composed of a series of detectors in the second stage. The detectors are trained in parallel, each with the detector output of the previous. Each subsequent detector has a higher IoU threshold than the previous. This adapted architecture should improve the hypotheses' quality and guarantee a positive training set for each detector. Additionally, it aims to minimize overfitting. The first stage is identical to the Mask R-CNN architectures, using identical backbones.

Mask2Former

Mask2Former is a transformer-based universal algorithm for image segmentation [13]. In contrast to the other two architectures presented, it utilizes transformers for the segmentation tasks and can be theoretically used to perform panoptic, instance, and semantic segmentation without retraining. The algorithm consists of three modules: the backbone, the pixel decoder, and the transformer decoder. The backbone architecture (e.g. Swin or Resnet) extracts low-resolution feature maps from the input image. These are then upscaled to high-resolution feature masks by a pixel decoder. The transformer decoder uses the features of the pixel decoder to generate object queries. The output masks are a combination of the feature maps from the pixel decoder and the object queries from the transformer decoder.

2.3. Feature Backbones

All the presented instance segmentation architectures have in common that they use a backbone for feature extraction. There are a vast number of different backbones available, each with advantages and drawbacks. In this study, we focus on two backbones, the residual network ResNet presented by He et al., with a depth of 101 layers [14] (R101) and the Swin-T backbone [15] of Liu et al. The Swin-T backbone is a shifted window transform backbone. The Swin-T backbone addresses the challenge of high resolutions and different feature scales by a hierarchical transformer architecture. The image is partitioned into regular windows and the self-attention is computed within each window. In the next layers, the windows are shifted. This architecture aims to provide flexibility for different feature scales while maintaining linear computational complexity with respect to image size.

2.4. Applications in Forestry

The first application of instance segmentation for forestry equipment focused on log detection for autonomous grasping with a forwarder grapple. For this purpose, Fortin et al. recorded the TimberSeg 1.0 dataset, consisting of 220 images [16]. They used Mask R-CNN, Rotated Mask R-CNN, and Mask2Former for the segmentation. The lowest accuracy was achieved by Mask R-

CNN, followed by Rotated Mask R-CNN with a 12 percentage points higher mean Average Precision (mAP) of 31.38 %. By far the best results were achieved by Mask2Former with a mAP of 57.53%. The algorithms showed good robustness against changing external influences such as snow, glare, and darkness. Geiger et al. investigated the partial automation of the loading process [17]. The YOLACT++ architecture was used to detect and segment the stems before gripping. The trained network reached a mAP over all classes of 56.65.

Deep learning and image classification with neuronal networks were studied by Liu et. Al in [18]. Liu et al. classified tree specimens and stock volume to provide a more efficient and faster alternative to conventional ground surveys. Grondin et al. created the synthetic dataset *SYNTHTHREE43K* consisting of 43,000 images, and the *CanaTree100* dataset, consisting of 100 RGB and depth images of Canadian forests in different weather conditions [7], [19]. The annotated masks and bounding boxes include trees and other objects such as stumps and grass. In addition to the masks and bounding boxes, the labels also include the position of the cut, the diameter, and the inclination. They used Mask R-CNN and Cascade Mask R-CNN with different backbones for the instance segmentation. After initial training on synthetic images in [19], the model was adapted to the real images in [7]. The Cascade Mask R-CNN algorithm gave better results than Mask R-CNN on all backbones, however, both algorithms reached high mean average precision of AP-50 > 85 and AP-50:95 > 60 for both bounding boxes and segmentation. In the second step, the trained model was applied to an unknown Portuguese dataset. A significant degradation of the results was observed.

Lagos et al. aimed to create a dataset for data-driven methods in forest environments [6]. Their *FinnWoodlands* dataset contains 300 RGB stereo images, point clouds, and sparse depth maps. They provide manual annotations for semantic, instance, and panoptic segmentation. The instance categories include three types of trees and the obstacles: “Lake”, “Ground” and “Track”. In addition to providing the data, they evaluated the instance segmentation with Mask R-CNN and EfficientPS where they achieved an AP-50 of 28% and 50% respectively.

3. DATASETS

For this study, two annotated datasets of forest areas have been created. The *MobimaWoodlands* (doi:10.35097/1749) dataset consists of two subsets, *MobimaWoodlands/Winter* and *MobimaWoodlands/Summer*, with 126 images each. Both subsets are captured manually with handheld cameras. This dataset features typically middle-European mixed forests in summer and in winter. The summer subset is in 16:9 landscape format, while the winter subset is in 4:3 portrait orientation. The second dataset, *MobimaSkidroads* (doi:10.35097/1750), consists of 293 images captured while driving on a skid or forest road with an industrial camera mounted on a vehicle. The dataset is completely in 16:9 landscape format and includes mixed and coniferous forests in summer and winter. Example images with annotations are displayed in Figure 2.



(a) **(b)**

Figure 2: Example images with annotations (blue: stems, purple: trees, green: stumps) of *MobimaWoodlands* (a) and *MobimaSkidRoads* (b)

The datasets include mask and bounding box annotations for the five object classes described in Table 1. The number of objects per class is given in Table 2.

Table 1: Object classes

Class	Description
Stems	Trunks and cutted trees with a diameter of approx. $\geq 10\ cm$
Trees	Trees with a diameter of approx. $\geq 10\ cm$
Stumps	Tree stumps
Obstacles	Non-passable objects not being subject to any of the other groups (e.g. stones, raised hides, way signs)
Persons	Complete or partially visible humans

The datasets were randomly divided into a training set and a validation set by a ratio of 0.8/0.2. The number of objects per class shows a significant bias. The number of annotated tree instances exceeds largely the other classes, while obstacles and persons are almost not present. However, unbiased datasets are challenging as the number of trees exceeds the number of stems, stumps, and especially humans present in a forest.

Table 2: Number of objects per class in the training/validation sets

Class	*MobimaWoodlands/ Winter*	*MobimaWoodlands/ Summer*	*MobimaSkidRoads*
Stems	175/34	177/54	83/34
Trees	425/108	697/232	2762/672
Stumps	13/1	53/17	67/19
Obstacles	6/0	7/11	32/1
Persons	0/0	10/1	0/0
Images	*100/26*	*100/26*	*234/59*

4. TRAINING SETUP

Three different network architectures, Mask R-CNN, Cascade Mask-RCNN, and Mask2Former were implemented in the following comparison. The architecture implementations from the PyTorch-based library MMdetection were employed [20]. In this study, the Swin-T backbone is used for each architecture except for the Mask R-CNN architecture, where we compared the results of the R101

backbone with those of the more complex Swin-T backbone.

All backbones were initialized with pre-trained weights from the COCO dataset. The weights were obtained from the MMDetection library. Training detection backbones is time-consuming; pre-trained backbones allow fast transfer learning on new datasets and minimize the computational time for training adapted detectors. Additionally, the first stages of the backbone layers have been frozen, as studies indicate that retraining the complete backbone has no significant benefit [21]. For the Mask R-CNN and Cascade Mask R-CNN architectures, the first backbone stage was frozen. For the training of the Mask2Former architecture, all backbone stages were frozen to reduce the number of trainable parameters.

In addition to the two published datasets, we included the *Mobimalogs* dataset from [17] in our training. To prevent a possible bias due to the different dataset sizes, the data loader created a mini-batch of 5 samples with a sample ratio of [1,1,2,1] from *MobimaWoodlands/Winter*, *MobimaWoodlands/Summer*, *MobimaSkidRoads*, and *MobimaLogs* respectively.

The training results largely depend on the training pipeline. Typically, a random image section is extracted from the base image at each training epoch. This section can be a part of the image or, with a certain probability, be the complete image. In contrast to this approach, the training pipeline depicted in Figure 3 uses a random section of each image but never the complete image to prevent overfitting of the networks to the surrounding area. Each image was resized first to a maximal size randomly sampled from the interval $l_{px} \in [1280, 3072]\ px$ while conserving the aspect ratio. In a subsequent step, a random section with size: $width\ x\ height = 800x800px$ was taken from the resized image. The section was flipped horizontally with a probability of 50%. Hence, only image parts with a ratio of $0.625x$ to $0.26x$ of the original image were utilized in the training process. This training pipeline reduced the performance of the trained networks when trained on a single dataset but increased it when different datasets with various tree aspects were present.

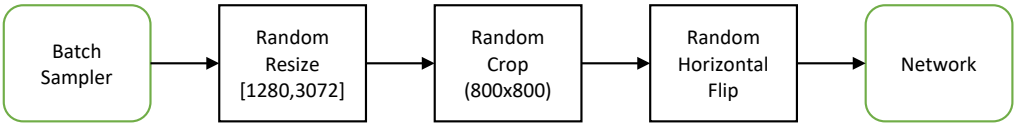


Figure 3: Training Pipeline

Due to the large bias in the datasets, evaluating the mAP would largely overrepresent the influence of the obstacle and person instances. In contrast to most studies evaluating the mAP, we evaluate the AP on a per-class basis.

5. RESULTS

The trained architectures are evaluated on the validation sets of the training datasets and the complete dataset of unknown forest areas represented by the *CanaTree100* and the *FinnWoodlands* datasets.

5.1. Training Datasets

At first, the trained network architecture was evaluated on the validation sets. Figure 4 depicts the AP on a per-class basis for the validation set of *MobimaWoodlands/Summer* for all four network architectures and the backbone (in parentheses). In the whole validation dataset, only one person is present, and all networks predicted the person with at least an IoU of 0.5. On the other hand, the random separation has put many of the obstacles in the validation set, and therefore all architectures struggle with the recognition of the obstacles, especially as this class includes various object types. This example illustrates, that the mean Average Precision is not a suitable metric when working with

biased datasets. All four architectures perform similarly on the tree class, with a small benefit for the Mask2Former architectures on stem and stump classes.

The results for all datasets are given in Table 3. The highest value for each class and dataset is displayed in bold. Both the MaskRCNN (Swin) and the Mask2Former perform better on some classes for the same datasets. In total, transformer backbones and the transformer architecture of Mask2Former did not lead to improvements regarding the average precision. In combination with the high AP levels of an IoU threshold of 0.5, this leads to the conclusion that the overall performance of the networks is not limited by the architecture but by the quality of the input datasets. Especially, the distinction between trees in the foreground, which are labeled, and trees in the background, which are not labeled, is ambiguous and therefore varies between images and datasets.

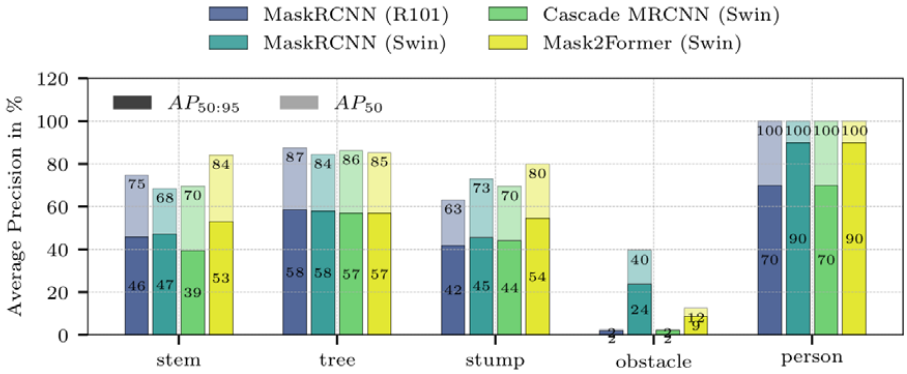


Figure 4: Average Precision (AP_{50} stacked on $AP_{50:95}$) of all architectures on the *MobimaWoodlands/Summer* validation set

5.2. Generalization

In the next step, the trained networks were evaluated on the dataset *CanaTree100* ([7]) and *FinnWoodlands* ([6]). As these datasets had not been used for the training, the networks were evaluated on the complete sets, including both, the training and the validation set. For the other datasets, that had been used for the training, the networks were evaluated on the validation set and not on the training set. The generalization of the networks was analyzed for the tree class, as the other classes are not consistent across all datasets. Figure 5 shows the AP of the tree class for all datasets. On the *CanaTree100* dataset, the AP remained in the same range of $AP_{50:95} \in [0.42,0.58]$ with no significant variation between the architectures. On the *FinnWoodlands* dataset, however, the AP of all architectures diminished significantly.

Table 3: Average Precision per class for each dataset (*bold: best AP per dataset*)

	Stem	*Tree*	*Stump*	*Obst.*	*Person*	*Stem*	*Tree*	*Stump*	*Obst.*	*Person*
MW/Winter		*MaskRCNN (R101)*					*MaskRCNN (Swin)*			
Box $AP_{50:95}$	0.62	0.49	0.70	-	-	0.63	0.47	0.80	-	-
Segm $AP_{50:95}$	0.39	0.48	**0.80**	-	-	0.42	0.45	**0.80**	-	-
		Cascade MRCNN (Swin)					*Mask2Former (Swin)*			
Box $AP_{50:95}$	0.65	0.48	**0.90**	-	-	**0.78**	**0.51**	0.80	-	-
Segm $AP_{50:95}$	0.44	0.44	**0.80**	-	-	**0.71**	**0.51**	0.70	-	-
MW/Summer		*MaskRCNN (R101)*					*MaskRCNN (Swin)*			
Box $AP_{50:95}$	0.54	**0.56**	0.39	0.01	0.70	**0.56**	**0.56**	0.41	**0.28**	**1.00**
Segm $AP_{50:95}$	0.46	**0.58**	0.42	0.02	0.70	0.47	**0.58**	0.45	**0.24**	**0.90**
		Cascade MRCNN (Swin)					*Mask2Former (Swin)*			
Box $AP_{50:95}$	0.47	0.57	0.42	0.02	0.80	0.55	0.50	**0.49**	0.08	**1.00**
Segm $AP_{50:95}$	0.39	0.57	0.44	0.02	0.70	**0.53**	0.57	**0.54**	0.09	**0.90**
MSkidRoads		*MaskRCNN (R101)*					*MaskRCNN (Swin)*			
Box $AP_{50:95}$	0.26	0.60	0.30	**0.90**	-	0.30	0.59	**0.37**	**0.90**	-
Segm $AP_{50:95}$	0.12	0.47	0.30	**0.80**	-	0.15	0.47	**0.35**	**0.80**	-
		Cascade MRCNN (Swin)					*Mask2Former (Swin)*			
Box $AP_{50:95}$	0.26	**0.62**	0.32	**0.90**	-	**0.34**	0.47	0.35	0.70	-
Segm $AP_{50:95}$	0.12	**0.48**	0.31	**0.80**	-	**0.30**	0.42	0.34	0.70	-

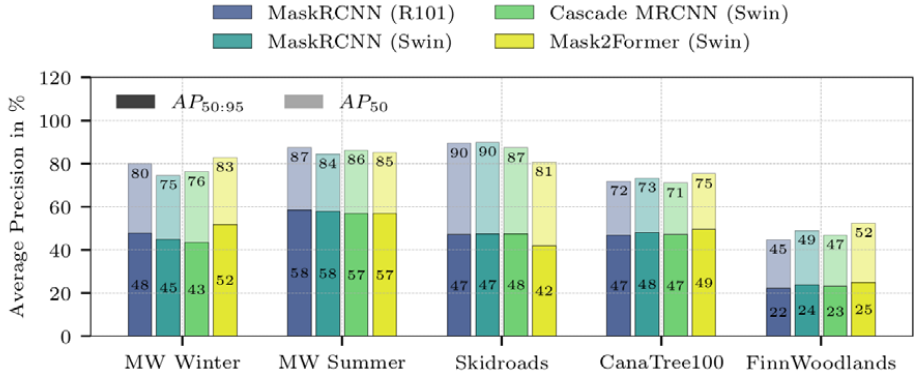
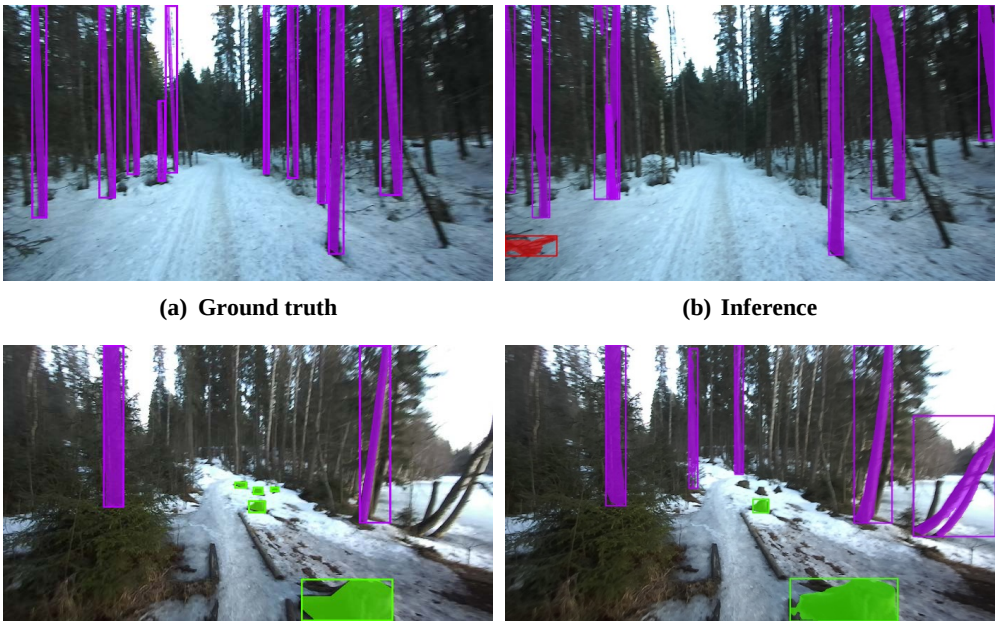


Figure 5: Average Precision (AP_{50} stacked on $AP_{50:95}$) of all architectures for the *Tree* object class

To explain the difference between the CanaTree100 and the FinnWoodlands results, three possible explanations are likely. The CanaTree100 dataset is probably more similar to our datasets and therefore requires less generalization of the network. Additionally, the FinnWoodlands dataset includes many birch trees which are not present in our datasets and therefore unknown to the networks. Figure 6 (a) and (b) illustrate False Negative detection of birch trees. Finally, missing labels reduce the AP, as seen in the example of Figure 6 (c) and (d). The network correctly identifies tree instances that were not labeled as trees. This ambiguity of the ground truth labels remains a major challenge when creating datasets and training image detectors for trees with different datasets.



(a) Ground truth **(b) Inference**

(c) Ground truth **(d) Inference**

Figure 6: Inference examples of the Mask2Former architecture on the FinnWoodlands dataset (blue: stems, purple: trees, green: stumps, red: obstacles).

The distribution of True Positives, False Negatives, and False Positives is displayed in Figure 7. The previous assumption that the remaining error on the training datasets is due to unlabeled trees, which are recognized as False Positives, is supported by the high percentage of False Positives, especially regarding the Mask2Former architecture. On the *FinnWoodlands* dataset, the number of False Positives detections is not larger than on the other datasets. On the contrary, many False Negatives were not detected, indicating that especially the undetected (birch) trees were responsible for the decreased AP.

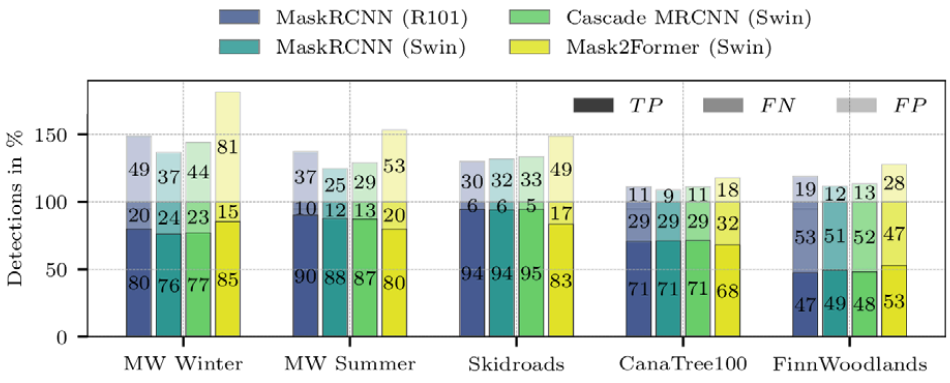


Figure 7: Error causes of tree classes (stacked: True Positive (IoU > 0.5), False Negative, False Positive, 100% = TP+FN)

6. CONCLUSION

In this paper, we trained four instance segmentation architectures with two different backbones on two newly generated datasets. All four architectures achieved a high Average Precision for the three primary object classes: stems, trees, and stumps. For IoU thresholds of 0.5 an Average Precision of above 0.80 is reached for the tree class with less than 10% False Negative tree detections. A qualitative evaluation of the detection results indicates that most of the remaining false detections are due to incorrect labels. Deciding whether a tree is in the foreground and should be labeled as a tree or whether it is part of the background is ambiguous and therefore prone to errors. Similar issues arise for stems and stumps, whether they are still distinguished objects or part of the forest floor. In a subsequent step, the trained networks were evaluated on two public datasets *CanaTree100* and *FinnWoodlands*. The generalization of the network was analyzed on the tree class as the other classes are inconsistent across the datasets. On the *CanaTree100* dataset, the AP_{50} decreased only about 0.1 for all architectures. On the *FinnWoodlands* dataset, however, the AP_{50} decreased by 0.3 points, and the detection result had around 50% false negatives detections. The high number of False Negatives could be explained by the fact that some of the tree specimens of the *FinnWoodlands* were not present in the training dataset. The findings indicate, that adapted training procedures and large datasets improve the transferability of the networks to unknown forest areas. However, the performance cannot be guaranteed, especially if the training dataset is from a different vegetation zone. More complex network architectures and new transformer architectures as well as transformer backbones, however, did not lead to better detection and generalization results than the Mask R-CNN architecture.

In the future, different approaches could be beneficial to further improve the generalization of object detectors in forest environments. The amount of available training data from different locations, representing different phenotypes and environmental conditions, needs to be increased and the quality of these datasets has to be ensured. New approaches, including stuff classes for background, could be beneficial to solve the ambiguity of background objects. Generative adversarial networks on the other hand could provide fictional images to train the detector on various object phenotypes and environmental conditions.

ACKNOWLEDGMENT

We like to thank ForstBW and Matthias Gauß for the possibility of collecting parts of the dataset on their skid roads. This publication is part of a project funded by the Federal Ministry of Education and Research (BMBF) under funding number 01IS23018 A-C.

REFERENCES

[1] T. Hellström, P. Lärkeryd, T. Nordfjell, and O. Ringdahl, *Autonomous Forest Machines - past, present and future*. 2008.

[2] R. Sharma, M. Saqib, C. T. Lin, and M. Blumenstein, "A Survey on Object Instance Segmentation," *SN COMPUT. SCI.*, vol. 3, no. 6, p. 499, Sep. 2022, doi: 10.1007/s42979-022-01407-3.

[3] O. Elharrouss, S. Al-Maadeed, N. Subramanian, N. Ottakath, N. Almaadeed, and Y. Himeur, "Panoptic Segmentation: A Review." arXiv, Nov. 19, 2021. doi: 10.48550/arXiv.2111.10250.

[4] G. Heitz and D. Koller, "Learning Spatial Context: Using Stuff to Find Things," in *Computer Vision – ECCV 2008*, D. Forsyth, P. Torr, and A. Zisserman, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2008, pp. 30–43. doi: 10.1007/978-3-540-88682-2_4.

[5] T.-Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context." arXiv, Feb. 20, 2015. doi: 10.48550/arXiv.1405.0312.

[6] J. Lagos, U. Lempiö, and E. Rahtu, "FinnWoodlands Dataset," in *Image Analysis*, R. Gade, M. Felsberg, and J.-K. Kämäräinen, Eds., in Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2023, pp. 95–110. doi: 10.1007/978-3-031-31435-3_7.

[7] V. Grondin, J.-M. Fortin, F. Pomerleau, and P. Giguère, "Tree detection and diameter estimation based on deep learning," *Forestry: An International Journal of Forest Research*, vol. 96, no. 2, pp. 264–276, Apr. 2023, doi: 10.1093/forestry/cpac043.

[8] R. Padilla, S. L. Netto, and E. A. Da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, 2020, pp. 237–242.

[9] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.

[10] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE international conference on computer vision (ICCV)*, Oct. 2017.

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., Curran Associates, Inc., 2015.

[12] Z. Cai and N. Vasconcelos, "Cascade R-CNN: High Quality Object Detection and Instance Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 5, pp. 1483–1498, May 2021, doi: 10.1109/TPAMI.2019.2956516.

[13] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 1290–1299.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[15] Z. Liu *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9992–10002, Oct. 2021, doi: 10.1109/ICCV48922.2021.00986.

[16] J.-M. Fortin, O. Gamache, V. Grondin, F. Pomerleau, and P. Giguère, "Instance Segmentation for Autonomous Log Grasping in Forestry Operations," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2022, pp. 6064–6071. doi: 10.1109/IROS47612.2022.9982286.

[17] C. Geiger, M. Weißenböck, and M. Geimer, "Assistance System for an Automatic Loading Process," in *Proceedings of The Joint Annual 43rd Annual Meeting of Council on Forest Engineering (COFE) & the 53rd International Symposium on Forestry Mechanization (FORMEC)*, Online, 2021, pp. 5–7.

[18] J. Liu, X. Wang, and T. Wang, "Classification of tree species and stock volume estimation in ground forest images using Deep Learning," *Computers and Electronics in Agriculture*, vol. 166, p. 105012, Nov. 2019, doi: 10.1016/j.compag.2019.105012.

[19] V. Grondin, F. Pomerleau, and P. Giguère, "Training Deep Learning Algorithms on Synthetic Forest Images for Tree Detection," May 2022.

[20] K. Chen *et al.*, "MMDetection: Open MMLab detection toolbox and benchmark," *arXiv preprint arXiv:1906.07155*, 2019.

[21] S. Hinterstoisser, V. Lepetit, P. Wohlhart, and K. Konolige, "On Pre-Trained Image Features and Synthetic Images for Deep Learning," presented at the Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0.

ASSISTED DRIVING MIDI-EXCAVATOR FOR AUGMENTED PERFORMANCES AND IMPROVED SAFETY

Andrea Cervi[1]*, Cesare Fantuzzi[2], Alberto Rocchi[1], Francesco De Martino[1], Fabio Tedeschi[1]

[1] *Walvoil SpA, via Adige 13/D, 42124 Reggio Emilia, Italy*
[2] *Università degli studi di Modena e Reggio Emilia, Italy*

* Corresponding author: Tel.: +39 0522 932411; E-mail address: cervi.an@walvoil.com

ABSTRACT

Numerous examples of assisted driving are present in the automotive sector to improve driver comfort and enhance safety. In the Off-Highway vehicles the examples of assisted driving are instead less widespread, although the stresses received by the driver are much higher, due to the more difficult environmental conditions and to the different operations that must be simultaneously controlled.

The purpose of this work is to develop a driving support system, mainly oriented to small-sized excavators, based on a software logic that can support the operator in the execution of predefined functions and paths.

Compared to previous works that dealt with the subject in a theoretical way, the present one starts from a robotic approach, to continue in an extensive simulation activity and conclude with an experimental application.

The case of a 5t Excavator is studied, equipped with 6 degrees of freedom inertial platforms, to dynamically monitor the geometry of the machine.

Keywords: assisted driving, safety, automation, excavator

1. BACKGROUND AND DRIVER NEEDS

There are many types of off-highway vehicles in the world and all of them can perform several different tasks. All these machine needs trained operators for their use, especially for more complex jobs that require constant attention and precision.

In addition to carrying out the specific work, the operator must also pay attention to his own safety and to that of the people around him; this can increase his level of stress.

The main goal of this job is to design, simulate and test logics that can assist the operator during the different working activities in order to both simplify the tasks and increase the safety.

The case study of this paper is a 5 ton excavator: the assisted functions are related to trajectory following tasks to simplify the creation of trenches and slopes with specific inclinations.

The second goal of this job is to perform the assisted operation only with a minimum of onboard sensors without using external equipment or sensors like GPS or Laser.

In this study the excavator is equipped only with 6 degrees of freedom inertial platforms to measure the inclination of the Boom, the Arm, the Bucket and the machine frame.

2. FROM ROBOTIC TO OFF-HIGWAY VEHICLES

An excavator is a typical hydraulic construction machine often used in dangerous and heavy duty working conditions, in adverse weather conditions and in dirty areas where it is not easy to drive the machine by a human operator.

System automation with electrical motors is well studied in the industrial robotics sector. Recently, there has been a growing interest in incorporating concepts from the robotics sector into the control of the hydraulic arms of excavators. In this way, the manual control of each joint would be automated, obtaining much greater performance and operational safety, thanks to the use of assisted systems for the generation of trajectories and joint control.

The excavator operations require coordinated movement of the swing, boom, arm and bucket to control the position of the End Effector (centre of the bucket tip, see figure 1) to obtain the desired trajectory. This can be achieved through the implementation of an automatic control system for the excavation task, which requires the understanding of the kinematics, the dynamics, and the control of the excavator.

Thanks to the large academic and industrial interest in this topic, several authors studied the kinematics of the excavator links [1], [2], the system mechanics and dynamics [5]. [6], and the control algorithms [3], [4], [7].

With respect to the relevant bibliography, in our work we aim to achieve a practical and effective approach that includes kinematic analysis and robust control algorithm development. The result of the theoretical research was tested through software (SIL) and hardware (HIL) in the loop simulations. Going into details, a model-based design approach was applied following these development steps:

- The kinematic and the dynamic models of the arm of the excavator have been deeply studied. This step also included the validation of the model using real data acquired from on-field experiments.
- The mathematics of the inverse kinematics have been derived in closed form to allow the implementation in cost effective ECU (electronic control unit), e.g. with limited computation power.
- A robust closed-loop control algorithm, derived from the implementation of trajectory follower in robotic applications, was finally developed and tested.

In the following sections the above development steps will be described in detail.

3. 5T EXCAVATOR

The hydraulic system of the Excavator analysed is a Load Sensing (LS) system [13] composed of a Variable Displacement Piston Pump with electronic flow control and a flow sharing Main Control Valve (MCV) provided of electrohydraulic commands.

All the HMI devices (joysticks, pedals and display) and the machine ECU are connected through a CANbus network.

The machine is equipped with 6 degrees of freedom inertial platforms (IMU), to dynamically monitor the geometry of the machine.

3.1. Vehicle sensors



Figure 1: 5T Excavator with Inertial Platforms

The sensors used [15] [17] are the following:

- Three single axis IMUs are used to detect the movement of the Boom, Arm and Bucket.
- One dual axes IMU is used to detect the frame movement (roll and pitch) respect to the ground.
- One Hall Effect sensor is used to detect the cabin swing.

The inertial measurement units used have the following characteristics:

Table 1: IMU characteristics

Characteristic	Value
Measurement Range	1 axis 360° / 2 axes ±90°
Absolute Accuracy	± 0.30°
Resolution	0.01°
Repeatability	≤ 0.05°
Dynamic Accuracy	± 0.50°
Hysteresis	≤ 0.05°
Signal	CANbus

The angles measured with the IMUs need to be elaborated in order to be used for kinematics purpose.

3.1.1. Boom Angle (Theta 3)

The Boom angle Theta 3 value is given by the IMU sensor plus an offset to consider the assembling orientation and a real time correction that consider the roll and pitch angles of the machine on the ground:

$$\vartheta_3 = \vartheta_{IMU_BOOM} + \vartheta_{IMU_BOOM_OFFSET} + \vartheta_{IMU_FRAME} \quad (1)$$

3.1.2. Arm Angle (Theta 4)

The Arm angle Theta 4 value is given by the IMU sensor plus an offset to consider the assembling orientation and a real time correction that consider the roll and pitch angles of the machine on the ground minus the Boom Angle:

$$\vartheta_4 = (\vartheta_{IMU_ARM} + \vartheta_{IMU_ARM_OFFSET} + \vartheta_{IMU_FRAME}) - \vartheta_3 \quad (2)$$

3.1.3. Bucket Angle (Theta 5)

Typically the Bucket IMU is not fitted directly in the bucket but on the linkage arm (see Figure 2 for details), and the calculation formula has to take into account this geometrical complexity.

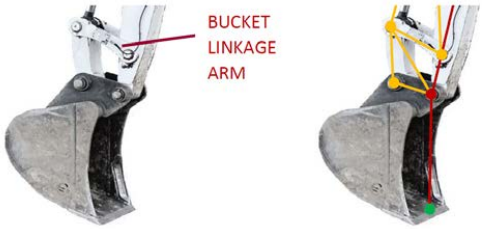


Figure 2: Bucket linkage arm

The angle value of the Bucket linkage Alpha 5 is given by the IMU sensor plus an offset to consider the assembling orientation and a real time correction that consider the roll and pitch angles of the machine on the ground minus the Boom and Arm Angles:

$$\alpha_5 = (\vartheta_{IMU_BUCKET} + \vartheta_{IMU_BUCKET_OFFSET} + \vartheta_{IMU_FRAME}) - \vartheta_4 - \vartheta_3 \quad (3)$$

The Bucket Theta 5 value is calculated using the linkage angle Alpha 5 with trigonometric approach.

3.1.4. IMU sensors VS Cylinder sensor

To validate the IMU solution, a comparison between the two sensor technologies has been made and in Figure 3 it is possible to notice the difference of the values of the two sensors for a typical working cycle.

The value of the difference of the two signals is normally less than +/- 0.5 degree.

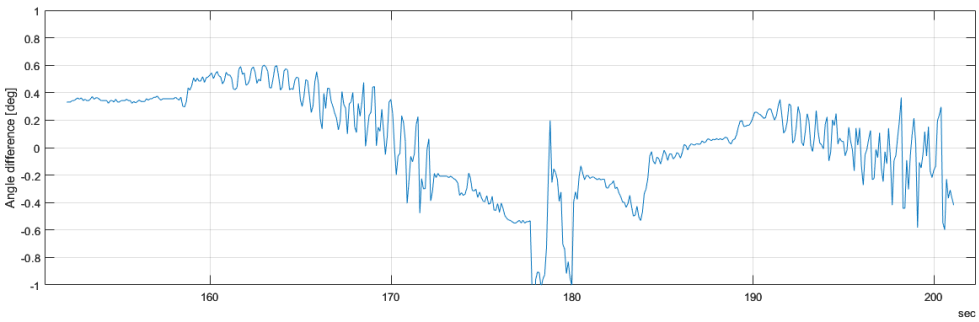


Figure 3: Comparison IMU VS Cylinder sensors

The use of the IMU sensors respect to other solutions like sensors inside the cylinder presents several benefits especially from the assembly and cost point of view.

3.2. Denavit–Hartenberg convention

The kinematic of the machine [10] has been obtained using the Denavit-Hartenberg convention.

The origin of the reference system (O0) is coincident with the centre of rotation of the cabin and positioned at the ground level (see figure 4).

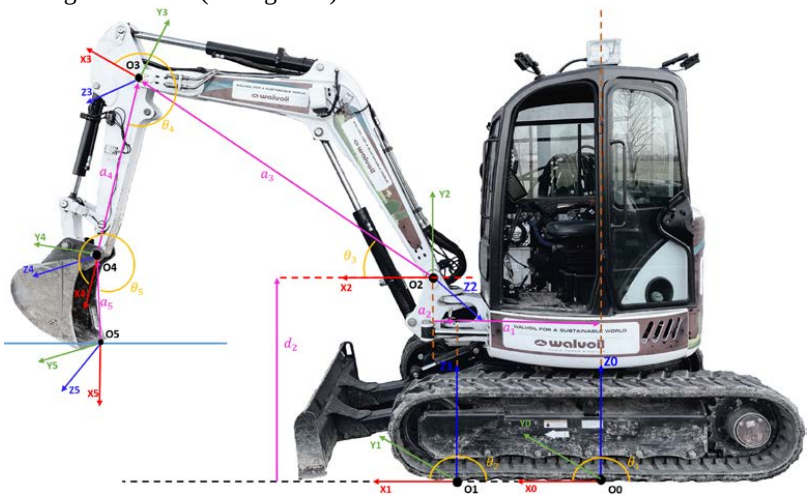


Figure 4: Denavit-Hartenberg reference systems

The Denavit-Hartenberg parameters of the specific excavator are detailed in the following table.

Table 2: Denavit–Hartenberg Parameters

Cylinder parameter	a_i [m]	d_i [m]	α_i [rad]	ϑ_i [rad]
Joint 1, Swing	0.799	0	0	ϑ_1
Joint 2, Boom Swing	0.155	1.155	Π / 2	ϑ_2
Joint 3, Boom	2.644	0	0	ϑ_3
Joint 4, Arm	1.370	0	0	ϑ_4
Joint 5, Bucket	0.725	0	0	ϑ_5

3.3. Trajectory definition

The operator defines the desired trajectory [14] [16] with the following parameters:

- Starting point; positioning the End Effector in the desired position.
- Inclination angle respect the horizontal and the vertical offset level; with machine HMI.

Algorithm calculate the trajectory up to the ending point where the joints reach them limits.

The geometry of this specific excavator allows a generation of a trajectory with a maximum length of 3000mm.

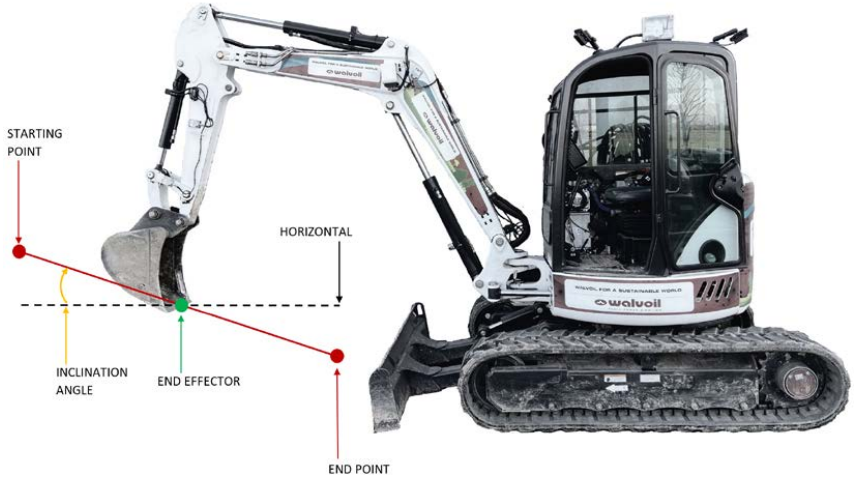


Figure 5: Trajectory generation example

In this paper two specific use cases have been considered:

- Trajectory with -1° of inclination; to simulate the final grading of a trench for pipe laying.
- Trajectory with 45° of inclination; to simulate the creation of an embankment.

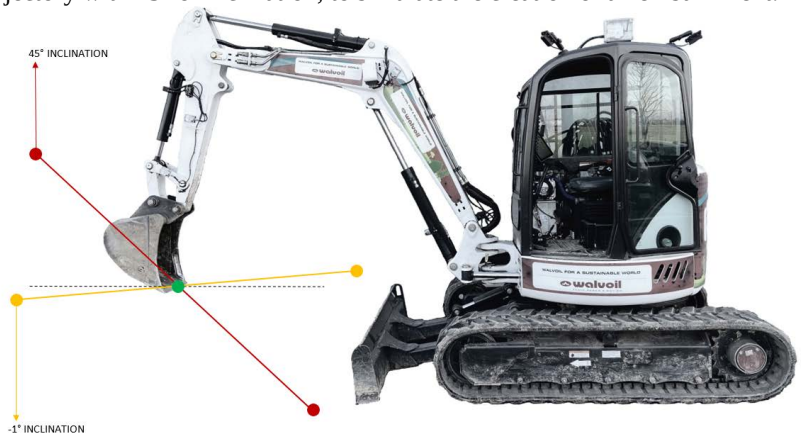


Figure 6: Trajectory used during the testing activities

3.4. Excavator Model

Validation of control algorithm is achieved using a multi-domain vehicle simulation [12].

In MATLAB® environment, a hybrid model can be built comprehensive of: CANbus command signals, hydraulic circuit and mechanical components.

The CANbus messages are converted via lookup-table into control current information, to be transmitted to the electro-hydraulic stage of the Main Control Valve.



Figure 7: Main Control Valve command

The Proportional Pressure Reducing Valves are modelled with all the characteristics of hysteresis and non-linearity to represent the real conversion from current command to pilot pressure.

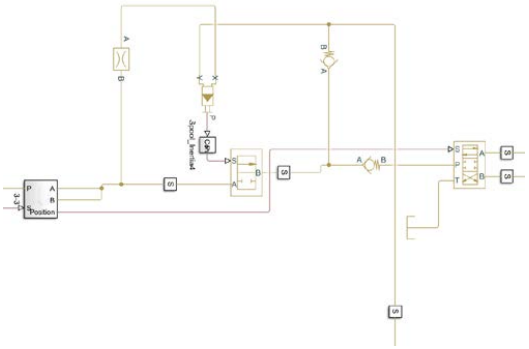


Figure 8: Proportional Pressure Reducing Valves and MCV

The pilot pressure is applied to drive the spools in the desired position. The Load Sensing MCV and the Pump are developed using the hydraulic components in the Simscape Fluids® library.

The purpose of modelling the MCV and its components in every detail (i.e. meter-in and meter-out area, local compensator, etc.) is to have a complete representation of the actuation speeds and any pressure phenomena on the vehicle. Thanks to these it's possible to check the control algorithm.

After verifying the correct functioning of the hydraulic power stage, the complete model of the vehicle, with its inertia, constraints and kinematics, a complete 3D CAD model is drawn (Solidworks®).

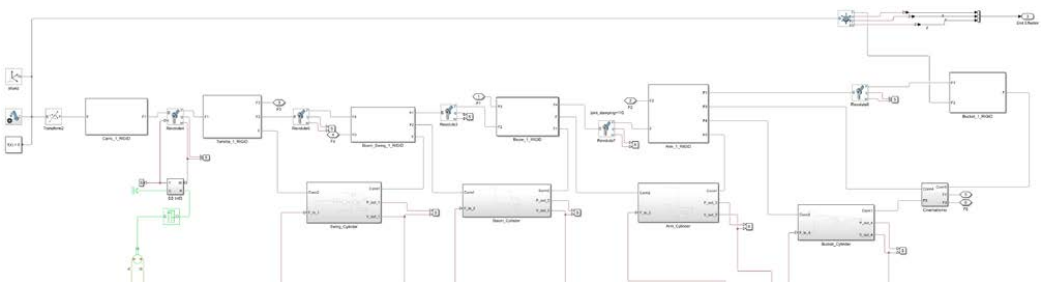


Figure 9: Mechanical System

The mechanical system is imported with a specific Simscape Multibody ® add on.

After export, the hydraulic system can be properly connected to the mechanical one to achieve the overall functionality of the vehicle.

As shown in Figure 9, the outputs of the model are the joints angle and the End Effector cartesian

coordinates.

Thanks to the Simscape Multibody ® tool it is also possible to see the 3D model of the machine during the simulation phase.

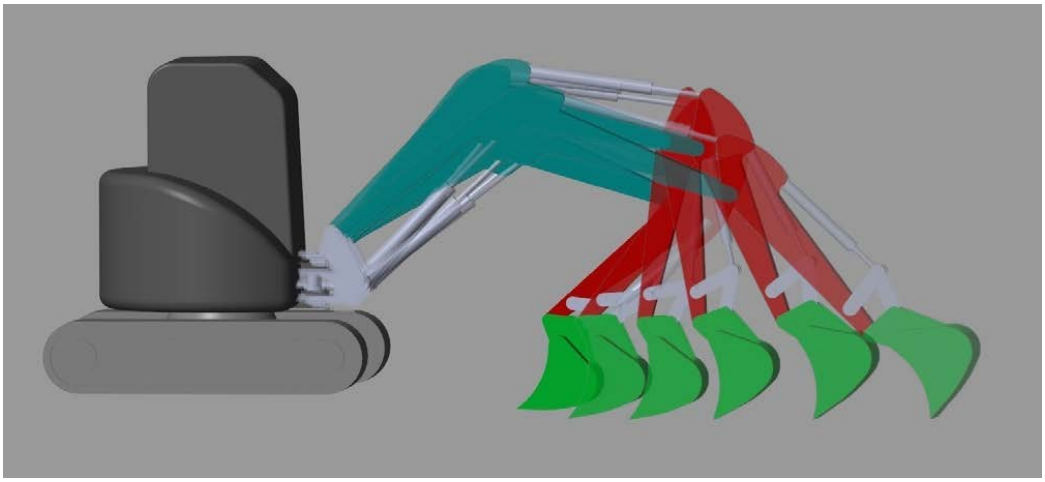


Figure 10: 3D model of the machine

3.5. Control algorithm

The model of the excavator has been inserted in a complete Simulink ® model with the specific control algorithm.

The controller used for tracking the reference position on the three joints is of the Proportional-Derivative (PD) type [8]. The main characteristic of the PD controller is to try to anticipate the position error, compensating the signal propagation delays, and improving the stability characteristics of the controlled system [11]. The equation describing the PD algorithm is

$$u(t) = K_p \cdot e(t) + K_d \cdot \frac{de(t)}{dt} \tag{4}$$

K_p is the proportional gain and K_d is the derivative term. In particular, the derivative term acts by compensating the variation of the joint position error, thus providing an anticipatory correction with respect to the estimate of the future joint position error.

4. SIMULATION RESULTS

First of all, the two defined use cases, trajectory with -1° of inclination and trajectory with 45° of inclination, have been tested in MATLAB Simulink ® environment with the defined control algorithm.

The benefit of simulation is that it is possible to test many control algorithms by varying the main configuration parameters until the optimal ones have been found.

In this way a huge amount of time dedicated to tests on the real machine on the field can be saved.

After the simulation loops, the optimal parameters values have been defined and the results are the following figures.

To evaluate the results of the simulation, the figures 11 and 12 show the trajectory error, defined as the difference between the computed and the simulated trajectory [9], and the joints angle error, defined as the difference between the computed and the simulated joint angle.

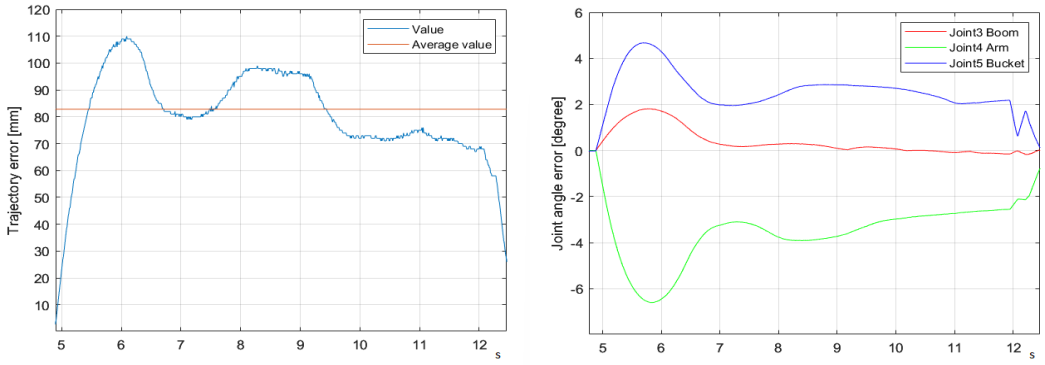


Figure 11: Trajectory and joints errors (-1° inclination)

The trajectory error average is 82.8 mm and the joints angle error is +4.5° / 6.5°.

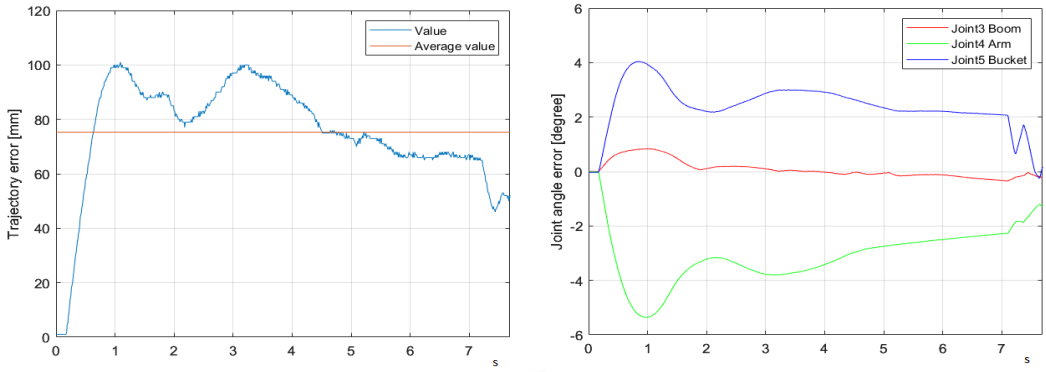


Figure 12: Trajectory and joints errors (45° inclination)

The trajectory error average is 75.3 mm and the joints angle error is +4° / -6°.

5. VEHICLE TEST RESULTS

At this point it is possible to start the real vehicle tests on the field initially with the optimal parameter values defined during the simulation.

After further tuning of the parameters on the machine, the best configuration has been obtained, and the results are in the following figures.

To evaluate the results of the experimental tests, the figures 13 and 14 show the trajectory error, defined as the difference between the computed and the measured trajectory [9], and the joints angle error, defined as the difference between the computed and the measured joint angle.

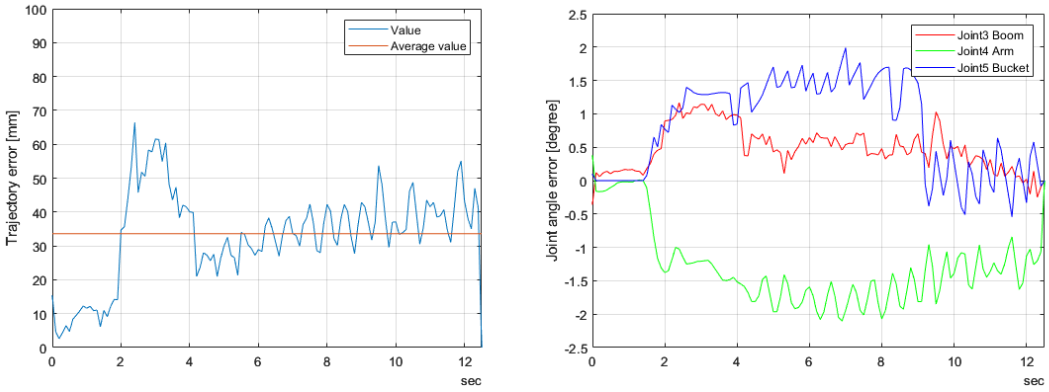


Figure 13: Trajectory and joints errors (-1° inclination)

The trajectory error average is 33.6 mm with a peak due to the initial acceleration phase and the joints angle error is less than ±2°.

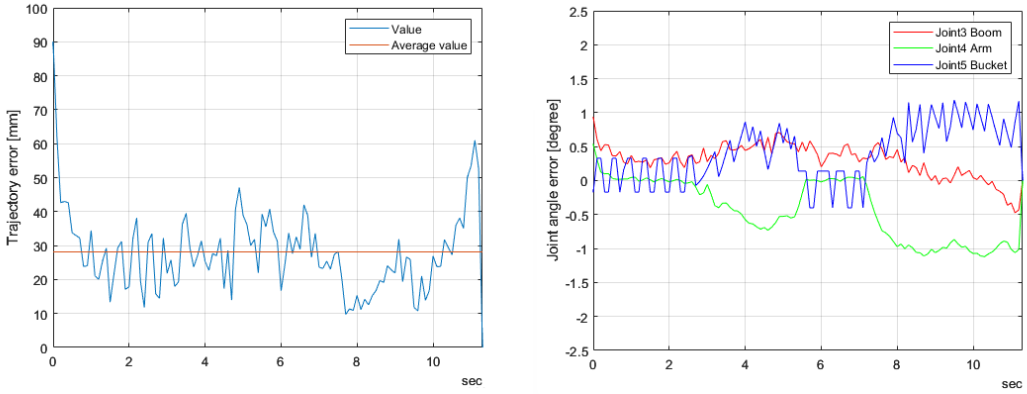


Figure 14: Trajectory and joints errors (45° inclination)

The trajectory error average is 28.1 mm and the joints angle error is less than ±1.5°.

6. CONCLUSION AND OUTLOOK

This paper confirms the benefit of working in simulation before starting any test on the vehicle, this gives the possibility to save resources and obtain an optimal result in a short time.

The main result of this work is the development of a linear trajectory driving system applied to a real excavator.

The experimental result show that the average difference between the theoretical trajectory and the real one is less than 50 mm, in line with the operator expectations.

This result that increases safety and accuracy with a reduced operator stress was achieved adding only a minimum number of sensors to the vehicle without using external equipment or sensors like GPS or Laser.

7. FURTHER DEVELOPMENTS

The next step of the work is to improve the assisted driving functions adding features that increase

the safety of the driver, the bystanders, and the vehicles in the working area.

The feature to improve the safety involves the adding of virtual wall that define the safety area where the machine is not allowed to work.

Further development is to define the safety area dynamically by means of a set of cameras that together with AI technology allows to detect and avoid workers or obstacles

NOMENCLATURE

LS	Load Sensing	
ECU	Electronic control unit	
PD	Proportional Derivative	
HMI	Human Machine Interface	
MCV	Main Control Valve	
K_p	Proportional gain	
K_d	Derivative gain	
α_5	Bucket linkage	rad
ϑ_3	Boom Angle	rad
ϑ_4	Arm Angle	rad
ϑ_5	Bucket Angle	rad
ϑ_{IMU_i}	Imu Angle	rad
a_i	Distance between two joints	m
d_i	Distance from ground	m
$e(t)$	Joint position error	rad
IMU	Inertial Measurement Unit	

REFERENCES

[1] Patel, B.P. and Prajapati, J.M. (2013) 'Kinematics of mini hydraulic backhoe excavator – part II', Int. J. Mechanisms and Robotic Systems, Vol. 1, No. 4, pp.261–282.

[2] Patel, Bhavesh & Prajapati, Jagdish. (2011). A REVIEW ON KINEMATICS OF HYDRAULIC EXCAVATOR'S BACKHOE ATTACHMENT. International Journal of Engineering Science and Technology. 3.

[3] D. Lee, I. Jang, J. Byun, H. Seo and H. J. Kim, "Real-Time Motion Planning of a Hydraulic Excavator using Trajectory Optimization and Model Predictive Control," *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, 2021, pp. 2135-2142, doi: 10.1109/IROS51168.2021.9635965.

[4] Christian Plaza, Olga Ramos, Dario Amaya, ,''Algorithm for the Control of an Autonomous Excavator", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 15 (2018) pp. 12229-12236

[5] Yoshida, T., Koizumi, T., Tsujiuchi, N., Jiang, Z. et al., "Dynamic Analysis of an Excavator During Digging Operation," *SAE Int. J. Commer. Veh.* 6(2):419-428, 2013, https://doi.org/10.4271/2013-01-2410.

[6] Wang, Xiaojun & Sun, Haoran & Feng, Minghao & Ren, Zhigui & Liu, Jurong. (2021). Dynamic Analysis of Working Device of Excavator under Limit Digging Force. Journal of The Institution of Engineers (India): Series C. 102. 10.1007/s40032-021-00725-4.

[7] Jeon, Pilsoo, and Giseo Park. "Automatic Excavator Control Using Model-Based Control Algorithm."

2023 26th International Conference on Mechatronics Technology (ICMT). IEEE, 2023.

[8] Wen Yu and Xiaoou Li, Hong-Nian Yu (2001). PD control of robot with velocity estimation and uncertainties compensation: Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228), Orlando, FL, USA, 2001, pp. 1162-1167 vol.2, doi: 10.1109/CDC.2001.981042

[9] Yang Liu; Mohammad Shahidul Hasan; Hong-Nian Yu (2010). Modelling and Remote Control of an Excavator. International Journal of Automation and Computing: DOI: 10.1007/s11633-010-0514-8

[10] A.J. Koivo; M.C. Ramos Jr; M. Thoma (1994). Dynamic Model For Excavators (and backhoes). Preprints of the Fourth *IFAC Symposium on Robot Control*.

[11] Le Quang Hoan, Chan Se Jeong, Hack Sun Kim, He Lim Yang and Soon Yong Yangd (2011). Study on modeling and control of excavator. *Preprints of the Fourth IFAC Symposium on Robot Control*. ISBN 978-89-954572-4-5, ISSN 2413-5844

[12] Zhang Jingming, Zhang Zhen and Luo Nianning (2020). Kinematics Analysis and Trajectory Planning of the Working Device for Hydraulic Excavators. *Journal of Physics Conference Series.* DOI: 10.1088/1742-6596/1601/6/062024

[13] Haoju Song, Guiqin Li, Zhen Li and Xin Xiong (2023). Trajectory Control Strategy and System Modeling of Load-Sensitive Hydraulic Excavator. *Machines* 2023, *11*, 10

[14] Fattah Hanafi Sheikhha, Ali Afzalaghaeinaeini and Jaho Seo (2021). Collaborative Tracking Control Strategy for Autonomous Excavation of a Hydraulic Excavator. *The 8th International Electronic Conference on Sensors and Applications.* https://doi.org/10.3390/ecsa-8-11333.

[15] B. Son, C. Kim, C. Kim and D. Lee, "Expert-Emulating Excavation Trajectory Planning for Autonomous Robotic Industrial Excavator," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2020, pp. 2656-2662, doi: 10.1109/IROS45743.2020.9341036.

[16] Guosheng Xu, Guangming Lv, Nianli Lu. Adaptive Sliding Mode Control for the Excavator Manipulator. Journal of Harbin Institute of Technology (New Series), 2020, 27(2): 27-35. DOI: 10.11916/j.issn.1005-9113.18117

[17] Dominic Jud, Simon Kerscher, Martin Wermelinger, Edo Jelavic, Pascal Egli, Philipp Leemann, Gabriel Hottiger, Marco Hutter (2021). HEAP - The autonomous walking excavator, *Automation in Construction, Volume 129,* 103783, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2021.103783.

Chapter 22

Actuators and Sensors-I

EXPERIMENTAL ANALYSIS OF ENERGY CONSUMPTION OF PIEZO ACTUATORS USED IN HYDRAULIC SWITCHING VALVE

Marko Simic[1]*, Niko Herakovic[1]

[1] *University of Ljubljana, Faculty of Mechanical Engineering, Askerceva 6, 1000 Ljubljana, Slovenia*

* Corresponding author: Tel.: +386 1 4771 797; E-mail address: marko.simic@fs.uni-lj.si

ABSTRACT

This paper presents a four-way digital piezo valve system (*4WDPVS*) composed of four high-response switching piezo valves. The main part of the new switching valves introduces the piezo actuator system (PAS) used instead of conventional solenoids. One of the important aspects in digital hydraulic systems is the use of energy-efficient switching hydraulic valves that assure low switching energy and low holding power consumption in the steady-state open valve state. First, the theoretical background of piezo technology is presented. Second, a detailed description of the new piezoelectric actuator system used in switching valves and four-way digital piezo valve system. The main part presents experimental analysis of the electrical energy consumption of the new piezoelectric actuator system. To validate and compare the electrical switching energy and the holding power consumption, several existing switching valves were used as reference. The results show a huge reduction of electrical energy consumption when using the new PAS for hydraulic switching valve actuation. Especially the holding power consumption in the stationary active state is about 6 times lower compared to other existing hydraulic switching valves controlled by solenoids.

Keywords: hydraulic switching valves, piezo actuators, energy consumption, digital piezo valve system

1. INTRODUCTION

The development of smart and energy-efficient hydraulic and pneumatic systems is demanded to design sustainable production systems for the factories of the future. Besides the hydraulic energy efficiency, the improvement of the electrical energy efficiency of the control components, i.e. the valves, has a major potential on the overall energy consumption reduction of the hydraulic and pneumatic system.

There were several developments in the past two decades that show the potential how to improve energy efficiency of hydraulic and pneumatic systems. Optimizing the shape of hydraulic valve housing and spools to reduce both, the static and dynamic flow forces, presents one of the solutions [1]. Lower flow forces lead to a reduced actuation force requirement for the valve, enabling the use of low-power solenoids.

A more promising solution for energy-efficient hydraulic and pneumatic systems is digital fluid power technology. The primary challenge in digital hydraulics and switching technology remains the development of fast, reliable, and energy-efficient on/off valves [2]. To significantly reduce electrical energy consumption, a single on/off valve within a digital fluid control unit (*DFCU*) can benefit from appropriate boosting control methods [2]. This involves using a high initial switching control voltage and then maintaining a lower voltage to keep the valve open in a stationary state, leading to reduced holding power and energy consumption in steady-state conditions. Alternatively, high-response, low-

power valve actuators are being explored. Concerning switching energy, miniature valves with flow rates up to 6 l/min have reasonably low switching energy ranging from 0.7 to 0.011 J [3]. Developing high-response and energy-efficient on/off valves with a larger nominal size remains a challenge. A promising solution involves the use of piezoelectric actuators. FESTO has demonstrated their integration into pneumatic high-response valves for pneumatic linear drives, with piezo on/off valves boasting a remarkable 20-times reduction in switching energy compared to conventional solenoid counterparts [4]. There are some studies focused on developing low-power, high-response piezo valves suitable for electrohydraulic direct-drive servo valves and pilot-operated servo valves [5]. In terms of energy efficiency, amplified piezoelectric actuators, combining flexure joints with piezo stacks, have been employed in brake systems. These piezo actuators consume a mere 21 mJ of energy, as opposed to the 950 J consumed by solenoid actuators, showcasing substantial energy savings [6]. Other studies in piezo technology concentrate on methods for conserving and reusing energy through piezo effects and efficient control systems. Energy-harvesting units, featuring piezoelectric elements placed in oscillating systems like bridges, walkways, motorways, and railways, generate and store electrical power. These units can produce up to 11.7 mW of output power using piezo stacks [7]. Additionally, units embedded in walkways offer 3-120 mW of output power [8]. Energy harvesting can extend to rotational motion through rotational shafts with integrated bending piezo actuators [9]. Results suggest suitability for low-speed rotation, with an average power output of up to 0.564 mW. Research in the field also explores how bridge vibrations impact energy harvesting, with optimal electronic resistance and location enhancing the potential energy savings, reaching approximately 579 µJ [10].

Existing literature lacks a detailed analysis of the electrical energy consumption of direct-operated on/off piezo valves. This paper addresses this gap by providing a comprehensive examination of the energy consumption of a novel piezo actuator system (*PAS*), which serve as an actuator in direct-operated hydraulic on/off valves. Understanding the behaviour and electrical characteristics of piezo stacks during hydraulic valve operation empowers us to anticipate the electrical energy consumption of *DFCU* during development.

2. PIEZOELECTRIC ACTUATORS

Considering the piezoelectric stacks as main elements of the hydraulic on/off valve actuators, the energy consumption depends on several parameters such as the maximum operating voltage, the capacitance of the piezo stacks and the switching frequency. In this investigation the focus is in the determination of the switching energy per single piezo stack, that is the energy, needed to fully charge the piezo stack and maintain active state.

2.1. Piezoelectric actuator as RC electrical circuit

The piezoelectric actuator can be described with an *RC* electrical circuit (**Figure 1**), where *R* is the electrical resistance and *C* the capacitor representing the piezo stack.

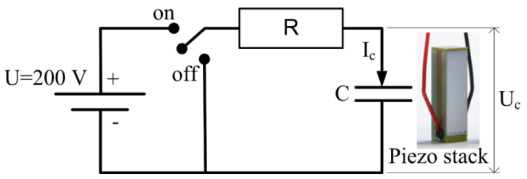


Figure 1: Piezo stack as *RC* electrical circuit

Charging and discharging the piezo stack is achieved by applying the electrical voltage to the

capacitor through the switching connector (on-off). The electrical current I_c appears in the RC circuit in the charging and discharging phase, which is determined by equation (1), where q is the electrical charge, C is the capacitance of piezo stack, and U_c is the measured electrical voltage.

$$q = C \cdot U_c \tag{1}$$

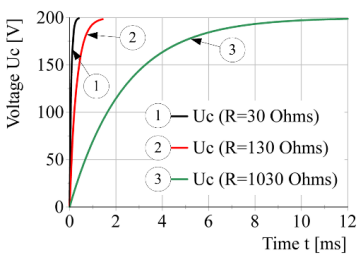
The step response of the first-order system is characterized by time constant τ presented by equation (2). By considering the RC electrical circuit, the final equation describing the electrical voltage $U_c(t)$ and the electrical current $I_c(t)$ can be formed with equations (3) and (4), respectively.

$$\tau = R \cdot C \tag{2}$$

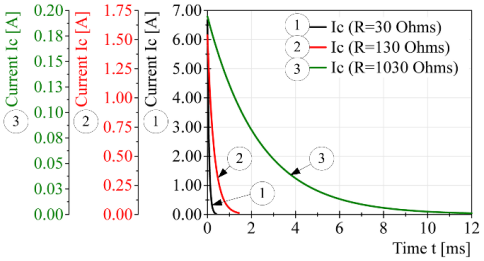
$$U_c(t) = U_0 \cdot \left(1 - e^{\frac{-t}{R \cdot C}}\right) \tag{3}$$

$$I_c(t) = I_0 \cdot e^{\frac{-t}{R \cdot C}} \tag{4}$$

Figure 2 and **Figure 3** represent the theoretical curves of electrical voltage and the electrical current. For the given examples including different electrical resistances R and consequently different time constant τ, the parameters of the mathematical models were set to the real conditions of piezo control electronics and piezo stack. In this investigation the maximal operating voltage U_0 =200 V, the electrical resistance of the electronics R=30 Ω, 130 Ω, 1030 Ω and the capacitance of the piezo stack C = 2.285 µF were considered. The minimum electrical resistance R=30 Ω (curve No. 1) is chosen according to the initial measurements of the step response of the piezoelectric actuators and the on/off valves. Other possible electrical resistances, such as R=130 Ω (curve No. 2) and R=1030 Ω (curve No. 3) can be achieved by adding the external resistance unit to the primary piezo control electronic circuit.



a) Charging the piezo



b) Discharging the piezo

Figure 2: Electrical voltage U_c and current I_c as a function of electrical resistance R

The theoretical results of the electrical voltage and electrical current presented in **Figure 2a** and **Figure 2b** show major differences in terms of the voltage increase and current decrease. The low electrical resistance R=30 Ω leads to the superior step response of the electrical voltage (U_c) increase, as well as the electrical current (I_c) decrease. The step response is approximately 0.2 ms. On the other hand, higher electrical resistance (R=1030 Ω) leads to the poor step response of the voltage U_c, and the electrical current decrease I_c.

2.2. The electrical power and energy consumption determination

The power $P(t)$ can be calculated by using the equation (5). A graphical presentation of the power curves for the different electrical resistances R is shown in **Figure 3**. Very high peak power, up to 330 W, is achieved due to high electrical current when using low electrical resistance R=30 Ω. On

the other hand, high power is needed for a very short time period (approximately 0.2 ms to charge the piezo stack up to 99%). Curves No. 2 and No. 3 represent the power defining the electrical resistances of 130 Ω and 1030 Ω, respectively. According to the theory, there is no need for any power to maintain the piezo stack's elongation after the piezo stack is charged. One of the important tasks of our investigation is to validate this theoretical statement considering the real working conditions.

$$P(t) = U_c(t) \cdot I_c(t) \tag{5}$$

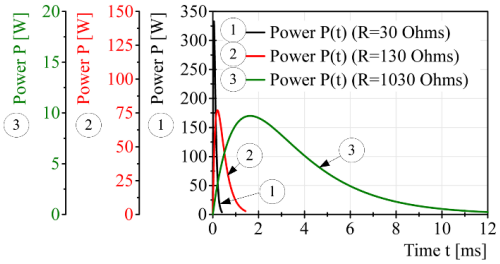
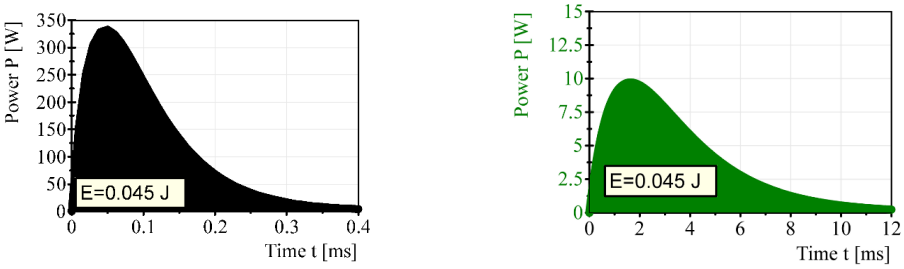


Figure 3: Power *P* as a function of electrical resistance *R* - charging the piezo stack

The energy consumption, i.e., the switching energy (E_{switch}) of the piezo stack for all three cases, presented in **Figure 3**, is defined as the area under the power curve. For a better understanding the power curves and the areas representing the switching energy for the electrical resistances of R=30Ω and R=1030Ω are shown in **Figure 4a Napaka! Vira sklicevanja ni bilo mogoče najti.**and **Figure 4b**. Equation (6) is used to calculate the switching energy. By comparing both switching energies, we can conclude that the energy needed to charge the piezo stack remains the same and it is not dependent on the electrical resistance *R* (E_{switch}=0.045 J).

$$E = E_{switch} = \int_{t1}^{t2} P(t)dt \tag{6}$$



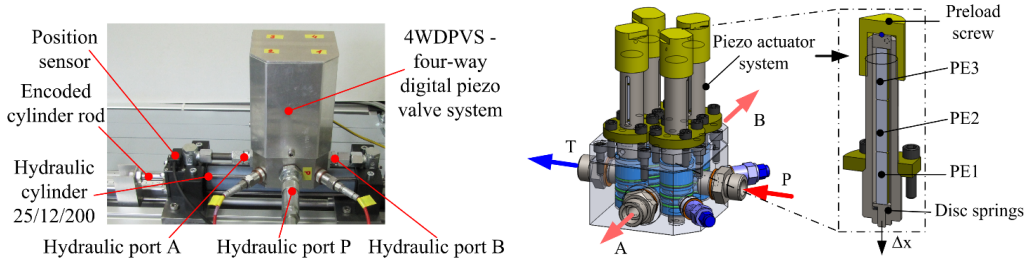
a) R=30 Ω b) R=1030 Ω

Figure 4: Switching energy as a function of electrical resistance - charging the piezo stack

3. DIGITAL PIEZO VALVE SYSTEM

The main component of the hydraulic linear drive (**Fugure 5a**) is a 4-way digital piezo valve system (*4WDPVS*). It consists of 4 on/off seat valves and 4 piezoelectric actuator systems (*PAS*), one per each valve. Peizoelectric actuator system is composed of 3 piezo stacks (*PE*) placed in serial (**Figure 5b**). The *PAS* with three piezo stacks enables several positions of the main spool of on/off valve and thus several discrete flow rates. The proper initial pretension of the piezo stacks is achieved with a preload screw and three disc-type springs placed in serial, thus preventing decoupling of piezo stack ceramic layers and achieving stable control. Activation of the piezo stacks corresponds to the

piezoelectric actuator system stroke Δx. The different activation of piezo stacks corresponds to discrete *PAS* stroke Δx and discrete flow rates of a single on/off valve. As we deal with on/off control technology, activation of one piezo stack results in 28 microns of *PAS* stroke, the activation of two or three piezo stacks results in 52 and 72 microns of *PAS* stroke, respectively. Based on the static performance testing the stability of the *PAS* stroke as well as the valve spool displacement (valve opening) is high, which means the stroke error below 1%. *PID* control in combination with Pulse Number Modulation strategy was used to control all 12 piezoelectric actuators (4 on/off valves). For the experimental investigation, gear pump with a constant flow rate Q_p=3,7 l/min, the pressure filter with 3µm and return line filter with 10 µm filtration rate were used. The pressure relief valve with the nominal size of Q_{RV} = 10 l/min, was set to p_s= 60 bar. Two stage, single rod hydraulic cylinder PARKER HMI ISO (25/12/200, piston/rod/stroke) was used for all tests. The nominal size of the valve is up to 20 l/min at pressure drop Δp=5 bar. The position sensor was integrated with the position resolution of 1µm in order to perform the closed-loop position control.



a) Hydraulic linear drive under investigation b) *4WDPVS*

Figure 5: Hydraulic linear drive and new piezo valve.

4. MEASUREMENT METHODOLOGY AND EXPERIMENTAL SETUP

The experimental tests are performed in order to analyse and verify the theoretical prediction of the energy consumption of individual piezo stack. The scheme of hydraulic linear drive is shown in **Figure 6.**

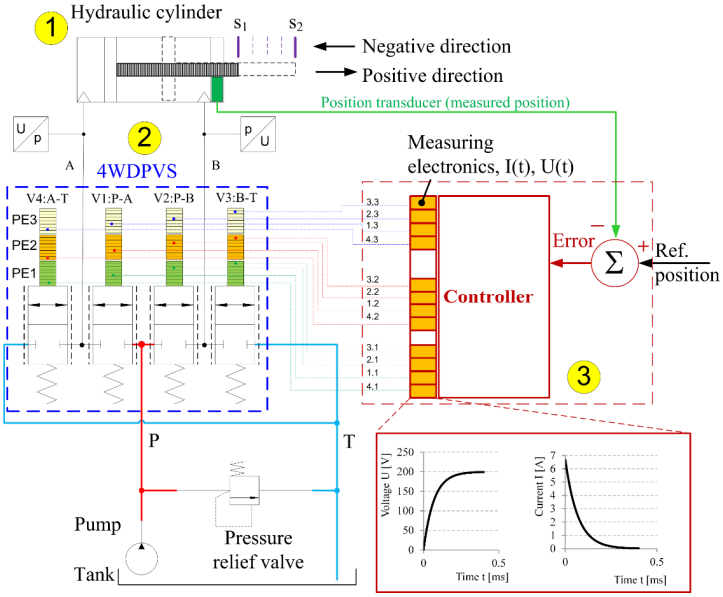


Figure 6: Measurement of electrical voltage and electrical current

The electrical energy consumption of the entire *4WDPVS* can be defined by considering all the active piezo stacks during the control cycle. It is also important to verify the theoretical prediction regarding the electrical energy consumption while the piezo stacks have already been charged and remaining in the stationary active state. Based on the fact that the *PAS* is exposed to external variable forces such as variable spring force of *PAS* and fluid forces during the operation we assume to measure small values of electrical current *I(t)*.

The detailed analyses of correlation between the active state of the piezo stacks and the electrical energy consumption of the piezo valves demands the electrical voltage (control signal) and electrical current measurement of the individual piezo stack during the test cycles. In order to measure the electrical current, the external measuring electronics with electrical current sensors ACS714 (measuring range up to ±30A and total output error of ±1.5%) was integrated in the controller unit [11]. The electrical voltage (U_c) is measured directly at the output signal of the control electronics. By measuring U_c and I_c, the power curve (P_c) and the energy consumption (E_c) are calculated. As we deal with on/off piezo valves and *PNM* control method, the electrical energy consumption represents the switching energy, which is needed to fully charge the piezo stacks, i.e., to open the single on/off valve.

The test cycle consists of the cylinder extension (positive direction) and cylinder retraction (negative direction) to the desired reference position. Four different reference positions s_2=25 50, 75 and 100 mm for the positive stroke of the cylinder and one reference position s_1=0 mm for the negative stroke of the cylinder were chosen for the test cycles (**Figure 7**).

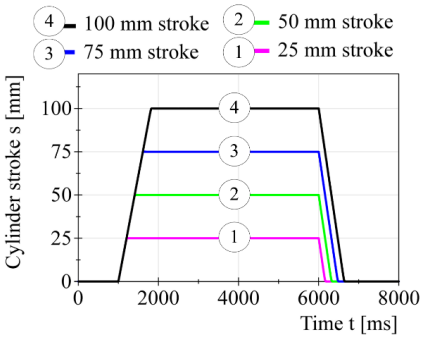
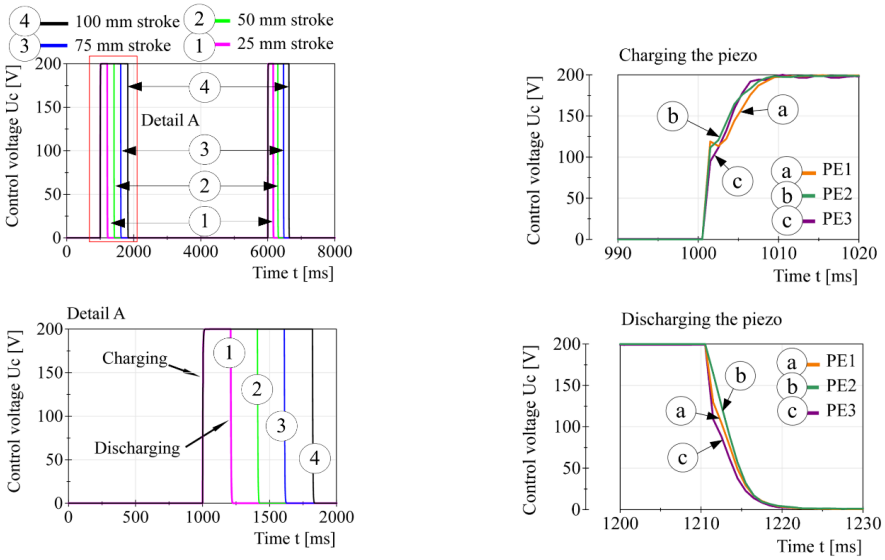


Figure 7: Cylinder stroke for test cycles

For each reference stroke of hydraulic cylinder five measurements were performed to measure *U(t)* and *I(t)*. The standard deviation of the measured parameters and the mean value were calculated using the statistical function in the NI DIAdem. The average results, calculated on the basis of five tests, of the electrical voltage and electrical current were considered in the interpretation of the results and in the final calculations of electrical power *P* and electrical energy consumption *E*. Since new *4WDPVS* operates with 12 control signals, one for each piezo stack *PE* of the on/off valve, the highest energy consumption is expected when all three piezo stacks of the on/off valve are activated at the same time during the test cycle (*P-A* and *B-T* valves for positive stroke and *P-B* and *A-T* for negative stroke of the cylinder). For our experimental tests the electrical resistance of the piezo control electronics R=1030 Ω was used that defines a measured step response of approximately 10 ms. We chose such electrical resistance for practical reasons. First, the *4WDPVS* has a sufficiently fast step response for closed-loop position control. Second, using a lower resistance would risk damaging the piezo stacks or necessitate increased pretension, reducing stack elongation. Third, theoretical analysis confirms that varying the R has no effect on energy consumption measurement results.

5. RESULTS AND DISCUSION

Figure 8a displays measured voltage curves for various test cycles, each based on five measurements per piezo stack. The standard deviation of measured voltage is less than 1 V. Notably, these curves correspond to individual reference position cycles. In **Figure 8a**, we observe two periods of maximum voltage amplitude for charging the piezo stacks during positive and negative cylinder strokes. During the positive stroke, each piezo stack requires a 200 V control signal (Detail A, Charging). After reaching the desired position, all piezo stacks in active on/off valves are switched off (Detail A, Discharging). This activation/deactivation strategy repeats for the negative cylinder stroke. For a 25-mm stroke cycle in valve *V1:P-A*, **Figure 8b** presents details of the average voltage curves for all three piezo stacks (*PE1*, *PE2*, and *PE3*) during the charging and discharging phases.



a) Measured U_c at stroke cycles b) Detailed U_c (charging, discharging)

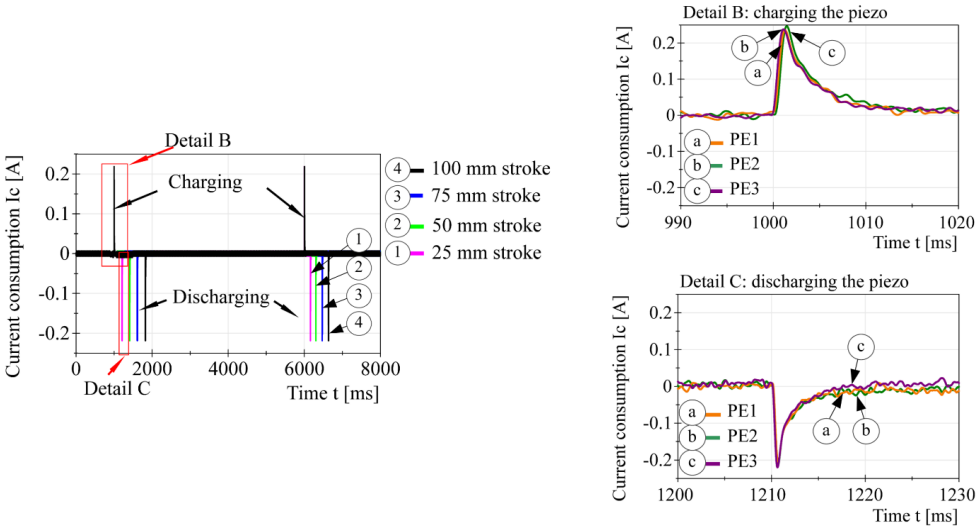
Figure 8: Measured electrical voltage during the test cycles

The results presented in **Figure 9a** and **Figure 9b** show the average electrical current consumption measured at piezo stacks in valve *V1:P-A* when performing test cycles. The average current-consumption curves are calculated based on five measurements for each piezo stack. The standard deviation of the measured current consumption does not exceed 0.01 A.

In line with theory, power is calculated using the average voltage and electrical current consumption for each piezo stack. **Figure 10a** illustrates the average voltage and current for piezo stack *PE1* in *V1:P-A*, specifically for the 25 mm test cycle. In **Figure 10b**, we calculate power (P_c) using equation (5). The crucial aspect of the P_c curve, regarding energy consumption, is the high-power demand during the charging phase. The peak power reaches 9.96 W, reflecting the piezo stack's maximum switching power. While this peak power is notably high, it's required for a short period, lasting around 3-4 ms for values exceeding electrical current 5 A. The calculated average power over the 210 ms duration of active piezo stack operation is a modest 0.46 W. When considering the activation of all three piezo stacks in the actuator system, the average power increases to 1.38 W.

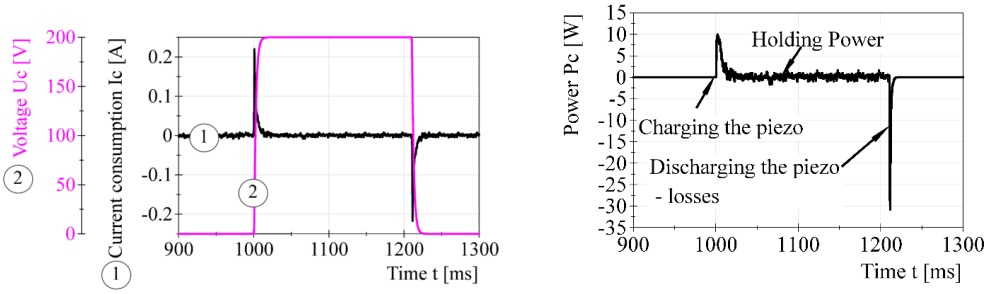
Analyzing signal U_c (**Figure 10a**) alongside P_c (**Figure 10b**), we determine the energy consumption of a single piezo stack, *PE1* (**Figure 11**, curve No. 2). While charging the piezo stack, the energy consumption, denoted as $E_{c,PE}$, amounts to 0.07 J, exceeding the theoretical prediction of E_c=0.045 J. This increased energy consumption arises from the necessary low constant holding power during the

stationary active state, approximately 11.9 mW. Over a 210 ms period, the energy consumption rises from 0.07 J to 0.095 J, impacting the overall energy usage of a single piezo stack in an on/off piezo valve. In essence, to maintain constant piezo stack elongation and keep the valve open, energy consumption increases by 0.0119 J per second.



a) Measured I_c during the test cycles, *PE1* b) Detailed I_c (charging, discharging)

Figure 9: Measured electrical current during the test cycles



a) Average voltage and current consumption b) Power curve of a single piezo stack

Figure 10: Average voltage and current consumption for 25 mm test cycle, *PE1*

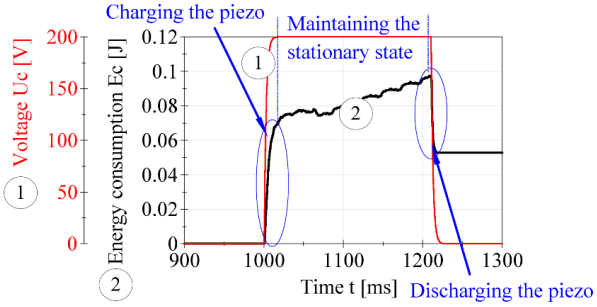


Figure 11: Energy consumption of piezo stack *PE1*, *P-A*

The energy-consumption curve shows an opportunity to save electrical energy when discharging the piezo stack. As we can see from **Figure 11**, the energy consumption drops from 0.095 J to approximately 0.055 J in the discharging phase. So potentially, 0.040 J of energy can be saved (up to 55% by considering the charging phase and the necessary switching energy). The energy savings can be made with proper new control electronics (energy harvesting circuit) and a new control algorithm, which is one of the main future goals.

The overall energy consumption of *4WDPVS* during the test cycles is presented in **Table 1**. The energy consumption considers the activation of all three piezo stacks per valve at the same time (activation of *PE1+PE2+PE3* of a single on/off valve). The energy consumption can be calculated by using empirical equation (7), where n is the number of activated piezo stacks, E_{switch} is the switching energy for the single piezo stack (0.07 J), T_{period} is the time period of active state of the piezo stacks and W_{hold} is the holding power (0.012 W) needed to maintain stationary state.

$$E\,[J] = n \cdot \left(E_{switch}[J] + \left(T_{period}[s] \cdot W_{hold}[J/s]\right)\right) \tag{7}$$

We also have to consider that the switching energy depends on piezo stack characteristics (capacitance), the type of the valve and operating pressure, i.e., the size of disturbance forces acting on the actuator during the valve operation. We can also conclude that the energy consumption is proportional to the time period and the number of the active piezo stacks.

Table 1: Energy consumption as a function of the active state of piezo stacks in *4WDPVS*

Cylinder stroke cycle [mm]	Active state *4WDPVS*	Period T_{period} [ms]	E_{4WDPVS} [J]
0 - 25	*P-A (3PE)* *B-T (3PE)*	210	**0.435**
0 - 50		410	**0.449**
0 - 75		610	**0.464**
0 - 100		820	**0.479**
25 - 0	*P-B (3PE)* *A-T (3PE)*	160	**0.432**
50 - 0		300	**0.442**
75 - 0		470	**0.454**
100 - 0		630	**0.465**

One of the goals of this study is to compare the energy consumption of the new on/off piezo valve with other on/off valves on the market or prototypes. **Table 2** presents the characteristics of our *4WDPVS* unit compared with some other on/off valves. The new *4WDPVS* has three possible discrete opening and thus three piezo actuators states (1*PE*, 2*PE* or 3*PE*). The remarks in the table mean: (p) - peak power, (h) - holding power and / - no data available.

Traditional 2/2-way direct operated seat valves are often powered by 22 W solenoids, leading to high holding power and energy consumption. Some commercial valves have been modified to use low-power solenoids, while others employ a boosting control method. This approach uses a high voltage signal initially and then reduces the holding voltage to the minimum required to keep the valve open, thereby decreasing holding power. For example, valves like SUN DLV and SI-1000 have reduced peak power from 22 to 15 W and operate with 2 W and 3 W of holding power. The optimized on/off valve (*WHV*), coupled with new control electronics and an intelligent boosting control method, achieves significantly lower switching energy and holding power consumption compared to traditional valves. Although the switching energy is comparable to our new on/off piezo valve, the prototype *WHV* still exhibits over six times higher average holding power consumption for valves of the same size. Furthermore, a prototype small water hydraulic valve for pilot-stage control achieves low switching energy of 0.011 J through innovative pilot-stage design and advanced *PWM* control.

Table 2: Characteristics of *4WDPVS* in terms of energy consumption

Characteristics Manufacturer	Valve type	Switching energy [J]	Power [mW]
UL FME LASIM *4WDPVS*	**Directional seat, piezo actuator**	**0.07 (1*PE*) 0.14 (2*PE*) 0.21 (3*PE*)**	**$9.96*10^3$ (p) 12 (h) 14 (h) 36 (h)**
Rexroth KSDER0 N/P [12]	2/2 seat valve	/	$22*10^3$(h)
Hydac WS08W [13]	2/2 seat valve	/	$20*10^3$(h)
Parker GS02-73 [14]	2/2 seat valve	/	$22*10^3$(h)
SUN DLV [15]	2/2 seat valve	/	$15*10^3$ (p)
Bucher WS22GD [16]	2/2 seat valve	/	$15*10^3$(h)
Tampere University of Technology WHV [17]	2/2 seat valve	≈0.05 (WHV1 AC) ≈0.07 (WHV2 AC) ≈0.15 (WHV2)	120(h) 80(h) 75(h)
Tokio Institute of Technology [18]	Pilot-operated seat valve	0.011	/

6. CONCLUSIONS

This paper presents an extensive experimental analysis of a novel piezo actuator system (*PAS*) comprising three piezo stacks integrated into hydraulic on/off valves, forming a four-way digital piezo valve system (*4WDPVS*). These digital valve systems serve as the primary control components for high-performance, energy-efficient hydraulic drives, replacing traditional hydraulic solenoid on/off valves in the *DFCU*.

The experimental findings demonstrate a substantial reduction in electrical energy consumption for hydraulic on/off piezo valves compared to conventional direct operated solenoid on/off valves. Switching energy is comparable to the best low-power on/off valve prototypes, ranging from 0.1 to 0.3 J. The key distinction lies in the stationary state, characterized by holding power consumption. Conventional on/off valves for digital hydraulics require high holding power, typically around 20 W, to keep the valve open. This power demand can be minimized by employing low-power solenoids and an intelligent boosting control method for activation, resulting in a minimum holding power of approximately 80 mW.

With the new *PAS* and on/off piezo valves integrated into the *4WDPVS*, a single piezo valve consumes 0.07 J when activating one piezo stack, 0.14 J with two, and 0.21 J with three. The holding power needed to sustain piezo stack elongation is low at 12 mW. Consequently, these new on/off piezo valves (comparable in size) require over six times less holding power than the best existing on/off valves suitable for digital hydraulics.

One significant advantage of utilizing piezo on/off valves is their suitability for various *DFCU* configurations that demand numerous on/off valves to achieve high-performance hydraulic drives. This, in turn, translates into substantial energy savings.

NOMENCLATURE

C	Capacitor, capacitance	[F]
DFCU	Digital Flow Control Unit	/
E, E_c	Energy consumption	[J]

E_{switch}	Switching energy	[J]
$E_{c,PE}$	Energy consumption of a single piezo stack	[J]
I_c	Electric current consumption at piezo stack	[A]
$I_c(t)$	Electric current consumption at piezo stack as a function of time	[A]
p_s	Hydraulic system pressure	[MPa]
$P(t)$	Electric power as a function of time	[W]
PAS	Piezo actuator system	/
PE	Piezo element, piezo stack	/
q	Electric charge	[C]
Q	Flow rate	[l/min]
Q_1	Flow rate of a single on/off piezo valve	[l/min]
R	Electrical resistance	[Ω]
RC	Resistance-capacitance electrical circuit	/
s_1	Reference position of hydraulic cylinder (negative stroke)	[m]
s_2	Reference position of hydraulic cylinder (positive stroke)	[m]
t	Time	[s]
T_{period}	Period in time domain	[s]
U	Voltage	[V]
U_c	Voltage at piezo stack	[V]
$U_c(t)$	Voltage at piezo stack as a function of time	[V]
U_o	Operating voltage	[V]
$U_{o,max}$	Maximum operating voltage	[V]
V	Valve, hydraulic on/off valve	/
W_{hold}	Holding power	[W]
Δx	Stroke of piezoelectric actuator system	[m]
τ	Time constant of RC electrical circuit	[s]
$4WDPVS$	four-way digital piezo valve system	

REFERENCES

[1] Li, R.; Sun, Y.; Wu, X.; Zhang, P.; Li, D.; Lin, J.; Xia, Y.; Sun, Q. Review of the Research on and Optimization of the Flow Force of Hydraulic Spool Valves. Processes 2023, 11, 2183. https://doi.org/10.3390/pr11072183

[2] Huova M, Linjama M, Huhtala K. (2015) Energy Efficient Digital Hydraulic Valve Control. (Tampere University of Technology. Publication; Vol. 1298). Tampere: Tampere University of Technology.

[3] Johnson, B., Massey, S. & Sturman, O. 2001. Sturman Digital Latching Valve. Proceedings Seventh Scandinavian International Conference on Fluid Power, May 30 – June 1, 2001, Linköping, Sweden, pp. 299–314 (Vol. 3).

[4] Wirtl H, Sixt U. White paper - Piezo technology in pneumatic valves. Festo AG & Co. KG, Germany, June 2017.

[5] Reichert M. Development of High-Response Piezo-Servovalves for Improved Performance of Electrohydraulic Cylinder Drives. Von der Fakultat fur Maschinenwesen der Rheinisch-Westfalischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades eines Doktors der Ingenieurwissenschaften genehmigte Dissertation, 04.02.2010.

[6] Tsukahara S., et al. “Design Method for Buckling Amplified Piezoelectric Actuator Using Flexure Joint and Its Application to an Energy Efficient Brake System.” Volume 2: Legged Locomotion;

Mechatronic Systems; Mechatronics; Mechatronics for Aquatic Environments; MEMS Control; Model Predictive Control; Modeling and Model-Based Control of Advanced IC Engines, 17-19 October, 2012, Fort Lauderdale, Florida, ASME, 2012, p. 95. © 2012 by ASME.

[7] Wang Ch, Wang Sh, Li J Q, Wang Xi, Gao Z, Zhang L. Fabrication and performance of a power generation device based on stacked piezoelectric energy-harvesting units for pavements. Energy Convers Manage 2018; 196:207-163. https://doi.org/10.1016/j.enconman.2018.02.045

[8] Ju Song G., Cho J. Y., Kim K-B., Ahn J. H., Song Y., Hwang W., Hong S. D., Sung T. H. Development of a pavement block piezoelectric energy harvester for selfpowered walkway applications. Applied Energy 256 (2019) 113916. https://doi.org/10.1016/j.apenergy.2019.113916

[9] Zou H-X, Zhang W-m, Li W-b, Wie K-X, Gao Q-H, Peng Z-K, Meng G. Design and experimental investigation of magnetically coupled vibration energy harvester using two inverted piezoelectric cantilever beams for rotational motion. Energy Convers Manage 2017; 1391:1398-148.

[10] Zhang Z, Xiang H, Zhan J. Experimental investigation on piezoelectric energy harvesting from vehicle-bridge coupling vibration. Energy Conversion and Management 2018; 169:179-163.

[11] Alegro MicroSystems, LLC. ACS714, Automotive Grade, Fully Integrated, Hall Effect-Based Linear Current Sensor IC with 2.1 kVRMS Voltage Isolation and a Low-Resistance Current Conductor. Alegro MicroSystems, LLC 2012; Rev.9 November 16.

[12] Rexroth Hydraulics. Hydraulic valves, on/off directional seat valves, direct operated, https://www.boschrexroth.com/en/xc/products/product-groups/industrial-hydraulics/on-off-valves/directional-valves/directional-seat-valves/direct-operated [accessed 22 April 2023].

[13] HYDAC International. Directional direct operated hydraulic valves, https://www.hydac.com/no-en/products/valves/directional-valves.html [accessed 22 April 2023].

[14] Parker Hydraulics. Bi-Directional Poppet Type, 2-Way Valve, Series GS02 72/73, https://www.parker.com/literature/Hydraulic%20Cartridge%20Systems/Product%20Series%20PDFs/SV/2%20Way%20Poppet%20Type%20Bi-Directional%20Solenoid%20Valve.pdf [accessed 22 April 2023].

[15] Scully L. New Line of Solenoid Valves Outperforms Previous Models, Hydraulics & Pneumatics, AUG 18, 2016, https://www.hydraulicspneumatics.com/technologies/hydraulic-valves/article/21885246/new-line-of-solenoid-valves-outperforms-previous-models [accessed 22 April 2023].

[16] Bucher Hydraulics. 2/2 Cartridge Seat Valve, Size 5, Series WS22GD…/ WS22OD…. Reference: 400-P-121110-EN-00, Issue: 09.2015.

[17] Paloniitty M. Novel Water Hydraulic On/Off Valves and Tracking Control Method for Equal Coded Valve System. Doctoral thesis, Tampere University of Technology. Publication 1577, 2nd of November 2018. http://urn.fi/URN:ISBN:978-952-15-4239-8.

[18] Park, S.-H., Kitagawa, A., Kawashima, M., Lee, J.-K. & Wu, P. 2002. A Development of Water Hydraulic High Speed Solenoid Valve. January 2005 TRANSACTIONS OF THE JAPAN FLUID POWER SYSTEM SOCIETY 36(1):8-14, DOI: 10.5739/jfps.36.8.

ENERGY HARVESTING FROM HYDRAULIC PRESSURE FLUCTUATIONS USING AN OSCILLATING PISTON

peer reviewed

Hauke Lerche[1]*, Alexander Kolberg[2], Jürgen Weber[1]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*
[2] *AMR-Hydraulik Chemnitz GmbH, Nordstraße 23, 09247 Chemnitz*

* Corresponding author: Tel.: +49 351 463-33619; E-mail address: hauke.lerche@tu-dresden.de

ABSTRACT

This paper presents an analytical model of a novel energy harvester for use in hydraulic systems. The pressure pulsation caused by the pump is used as the energy source. The harvester can be connected to a hydraulic system's line and is composed of an oil volume that serves as an oil spring, a vibrating piston excited by the pressure pulsation in the hydraulic system, and an electromagnetic linear generator. The harvester is designed to be compatible with a broad static pressure range due to static pressure compensation between the oil volume and the system. The analytical description of the piston motion is derived and validated by measurements. Further, the mechanical harvester parameters are optimized to achieve maximum piston speed. Lastly, the possible performance of the harvester is estimated.

Keywords: Pressure Ripple, Autarkic Sensor

1. INTRODUCTION

The functionality of hydraulic systems, such as in sheet metal forming technology or rolling mills, is heavily dependent on sensor technology. The automated control units are based on the measured values supplied by sensors in order to manage the autonomous functional sequence of the machine. Furthermore, measurements are used to monitor and control the systems. Today's measurement technology in the field of hydraulic systems usually works with wired sensors. However, the limited flexibility in the application leads to a limited planning freedom in the positioning of the sensor, because cable guides cannot be placed everywhere. In extreme cases, especially in the industrial sector, a cable break can destroy the complete functional use of the sensor and the operation of a system may not be continued. When implementing a machine design, one must consider not only the technical risks but also the expenses for planning, materials, and cable routing. Moreover, retrofitting wired sensors can incur high costs. As an alternative, there has been a growing trend in the use of battery-operated sensors which require less effort in cabling planning and provide a higher level of safety, given the absence of cables. However, maintenance accessibility, such as for battery replacement, needs to be guaranteed, which curtails the versatility of these sensors.

In the field of research, there are already approaches for self-sufficient sensor technology through energy harvesting. Numerous publications and meta-studies [1]–[5] on the subject of energy harvesting have already been published. While many studies concern energy harvesting from vibration, some delve into energy harvesting from pressure pulsation in fluid power systems. Skow

presents a harvesting concept designed for use in hydraulic systems [6]. Here, the pressure pulsation in the fluid is transmitted to a piezostack by using a metallic diaphragm. With a pressure pulsation of 4 bar, a power output of 1.2 mW can be achieved. However, the static pressure range is restricted as the load on the piezo elements becomes excessive beyond 100 bar resulting in decreased power. To achieve higher static pressure levels the system can be preloaded, but it results in no further energy generation at low static pressure ranges. In Ren et al.'s study [7], a cylindrical magnet that is supported magnetically is excited to oscillate by pressure pulsations found in pneumatic systems. Coils near the cylinder magnet generate an induction bias. The system is stimulated with a pressure amplitude of 0.8 bar and 30 Hz, and it attains a maximum power of 2 mW.

Examining the acoustic power density of pressure pulsation (refer to **Figure 1**) [8], it is clear that the current percentage yield from pulsation is very low. This indicates the potential for optimization to increase yield.

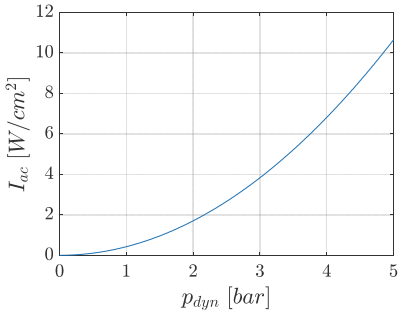


Figure 1: Acoustic Intensity of the pressure ripple in a hydraulic system

$$I_{ac} = \frac{p_0^2}{2\rho c} \tag{1}$$

2. STRUCTURE OF THE ENERGY HARVESTER

The schematic structure of the developed harvester is shown in **Figure 2**. The electromagnetic linear generator consisting of a coil and magnet is to be driven by a piston. The piston is excited to oscillation by the pressure pulsation in the system, which induces a voltage in the coil. An oil spring is to serve as the counterforce to the dynamic pulsation force of the hydraulic system. In simplified terms, the harvester can be seen as a single-mass oscillator. In order to compensate for the load caused by the static system pressure, the piston is guided in an annular gap seal through which a small volume flow can pass and thus acts as a kind of mechanical low-pass filter, whereby static pressure equalization can take place between the hydraulic system and the cavity of the harvester. The oscillation of the piston can thus be considered almost independently of the static pressure and is excited via the dynamic pressure. The influence of the static pressure on the compression modulus and the viscosity of the oil must be taken into account.

The linear generator is to provide sufficient energy for the operation of a pressure sensor as well as a radio unit. To ensure supply continuity even during standstill, a supercap will be utilized, capable of being charged by the generator as required. Nonetheless, the present study concentrates on the analytical modelling and validation of the piston oscillation.

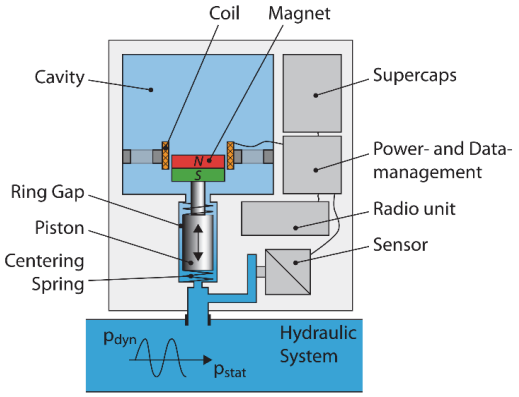


Figure 2: Schematic representation of the energy harvester

3. ANALYTICAL MODELING

The motion of the piston in the energy harvester can be described analytically on the basis of the acting forces (see **Figure 3**) using Newton's equation of motion:

$$\dot{v}(t)\, m_P + F_g + F_S - F_F + p_{Cav}(t) A_P = p_{sys}(t) A_P \tag{2}$$

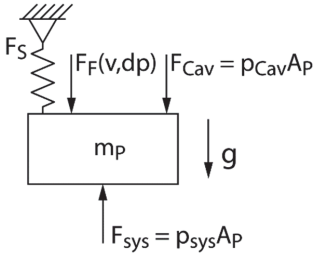


Figure 3: Forces acting on the piston of the energy harvester

Velocity and mass of the piston and the area of the piston's face are represented by v, m_P and A_P, respectively. F_F is the resultant frictional force on the piston attributed to the velocity gradient of the fluid within the annular gap. The weight force $F_g = m_P \cdot g$ and the spring force F_S of the centering springs are neglected. The pressure in the cavity p_{Cav} is denoted by the following differential equation, where the equivalent compression modulus is marked by K' and V_{Cav} represents the oil volume in the cavity:

$$\dot{p}_{Cav} = \frac{K}{V_{Cav}} (v\, A_P + Q_{Gap} + Q_{GapDrag}(v)) \tag{3}$$

The pressure gradient is determined by the speed of the piston and the volume flow through the annular gap. The drag flow $Q_{GapDrag} = 0.5 \cdot v\, A_{AnnularGap}$ caused by the piston movement can be neglected, since $A_P \ll A_{AnnularGap}$ applies. The pressure difference between system port and cavity also generates a volume flow Q_{Gap}, which can be calculated according to the following equation:

$$Q_{Gap} = \underbrace{\frac{\pi\, d_P\, s_{Gap}^3}{12\,\eta\, L_{Gap}} \left(1 + \frac{3}{2} \cdot \left(\frac{s_{exz}}{s_{Gap}}\right)^2\right)}_{=C_{Gap}} \cdot \left(p_{sys} - p_{Cav}\right) \tag{4}$$

Here, the diameter of the piston is referred to as d_P, with s_{Gap} denoting the height and L_{Gap} indicating the length of the annular gap. The dynamic viscosity of the oil is represented by η. Since the piston is not concentrically fixed in the sleeve, eccentric displacements can occur, which are taken into account by the eccentricity s_{exz}. The dynamic viscosity η is calculated as a function of the static pressure p_{stat} (in bar) [9]. For a static pressure $p_{stat} = 0\ bar$ a viscosity of $\eta_0 = 0.04\ Pa \cdot s$ is assumed.

$$\eta = \eta_0 e^{1.7 \cdot 10^{-3}\ \frac{1}{bar} p_{stat}} \tag{5}$$

To determine the frictional force, it is necessary to analyze the velocity profile v_f of the flow in the annular gap, as illustrated by **Figure 4**. It is assumed that this can be calculated from the summation of the simplified linear profile $v_{f,drag}$ due to the drag flow caused by the piston motion (see equation (6)) and the quadratic profile $v_{f,dp}$ due to the laminar flow caused by the pressure difference between the system and the cavity.

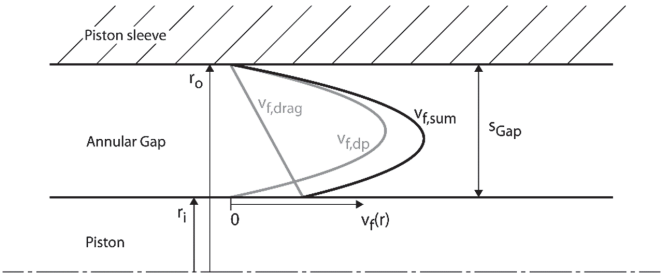


Figure 4: Velocity profiles of the fluid in the annular gap

$$v_{f,drag}(r) = -\frac{v}{r_o - r_i} r + \frac{r_a v}{r_o - r_i} \tag{6}$$

$$v_{f,dp}(r) = -\frac{6Q_{Gap}(r - r_o)(r - r_i)}{\pi(r_o - r_i)^3(r_o + r_i)} \tag{7}$$

$$v_{f,sum}(r) = v_{f,drag}(r) + v_{f,dp}(r) \tag{8}$$

Here, r_i and r_o are the inner and outer radius of the annular gap, respectively and v is the speed of the Piston. The frictional force acting on the piston can now be calculated from the velocity gradient at the point $r = r_i$ as follows, where A_{PS} is the surface area of the piston:

$$F_F = \left.\frac{dv_{f,sum}}{dr}\right|_{r=r_i} \cdot \eta A_{PS} = \frac{6Q_{Gap}}{\pi s_{Gap}^2 (r_a + r_i)} - \frac{v}{s_{Gap}} \tag{9}$$

Substituting equations (4) and (9) into (2) and setting $r_o + r_i = r_o$, we obtain the following differential equation.

$$\left(\frac{V_{Cav}\,\ddot{p}_{Cav}}{K'} - C_{Gap}(\dot{p}_{sys} - \dot{p}_{Cav})\right)\cdot m_P + \left(\left(\frac{1}{A_P} - \frac{6}{\pi r_o s_{Gap}}\right) C_{Gap}(p_{sys} - p_{Cav}) - \frac{V_{Cav}\,\dot{p}_{Cav}}{K'}\right)\frac{\eta A_{PS}}{s_{Gap}} + p_{Cav}A_P^2 = p_{sys}A_P^2 \quad (10)$$

Laplace transformation can be used to solve the equation in the image domain and form the transfer functions $G_p(s)$.

$$B_1 = \pi r_o A_P^2 s_{Gap}^2 + 6A_{PS}C_{Gap}A_P\eta - \pi A_{PS}C_{Gap}r_o s_{Gap}\eta \quad (11)$$

$$G_p(s) = \frac{\hat{p}_{Kav}(s)}{\hat{p}_{sys}(s)} = \frac{\frac{\pi C_{Gap}m_P r_o s_{Gap}^2}{B_1}s + 1}{\frac{\pi V_{Cav}m_P r_o s_{Gap}^2}{K'B_1}s^2 + \frac{\pi r_o s_{Gap}(C_{Gap}m_P K' s_{Gap} + A_{PS}V_{Cav}\eta)}{K'B_1}s + 1} \quad (12)$$

In addition, the transfer function $G_v(s)$ is required. From equations (3) and (4), the following equation can be generated after reshaping and transformation into the image domain.

$$s\hat{p}_{Cav}(s) = \frac{K'}{V_{Cav}}(\hat{v}\,(s)A_P + C_{Gap}(\hat{p}_{sys}(s) - \hat{p}_{Cav}(s))) \quad (13)$$

If this is converted to p_{Cav} and inserted into the Laplace Transform equation (2), the transfer function $G_v(s)$ can be calculated:

$$G_v(s) = \frac{\hat{v}(s)}{\hat{p}_{sys}(s)} = \frac{\overbrace{\frac{V_{Cav}(\pi A_P r_o s_{Gap}^2 + 6A_{PS}C_{Gap}\eta)}{K'B_1}}^{=T_D}s}{\underbrace{\frac{\pi V_{Cav}m_P r_o s_{Gap}^2}{K'B_1}}_{=T_1^2}s^2 + \underbrace{\frac{\pi r_o s_{Gap}(C_{Gap}m_P K' s_{Gap} + A_{PS}V_{Cav}\eta)}{K'B_1}}_{=2DT_1}s + 1} \quad (14)$$

It can be seen that the transfer function $G_v(s)$ corresponds to a DT2 term. Consequently, the time constants T_1 and T_D can be represented as in equation (14) and the damping D can be calculated as follows:

$$D = \frac{\pi r_o s_{Gap}(C_{Gap}m_P K' s_{Gap} + A_{PS}V_{Cav}\eta)}{2\sqrt{\frac{B_1\pi V_{Cav}m_P r_o s_{Gap}^2}{K'}}} \quad (15)$$

In contrast to a conventional single-mass oscillator, another factor that affects the damping occurs here in addition to the friction in the annular gap, namely the volume flow through the annular gap.

4. EXPERIMENTAL STUDIES

The validation and parameter identification of the system equations is carried out by means of high-frequency pressure measurements on a demonstrator (see **Figure 6**). Since only the behavior of the piston oscillation is to be investigated here, no linear generator is used in this setup. However, placeholder masses can be attached to the piston rod. In addition, the cavity is designed in such a way that the volume can be variably adjusted. The demonstrator is installed in a simple hydraulic circuit as shown in **Figure 6** below, and the pulsation of the system and cavity pressure can be measured. To avoid falsification of the measured values due to reflections and standing waves in the system, the system pressure is measured as close as possible to the energy harvester. If the system pressure is

recorded too far away, the error between the measured value and the true dynamic pressure p_{sys} at the input of the harvester - which is used in the analytical calculation - is too large. For the validation, different static pressure ranges from 50 bar to 200 bar were investigated.

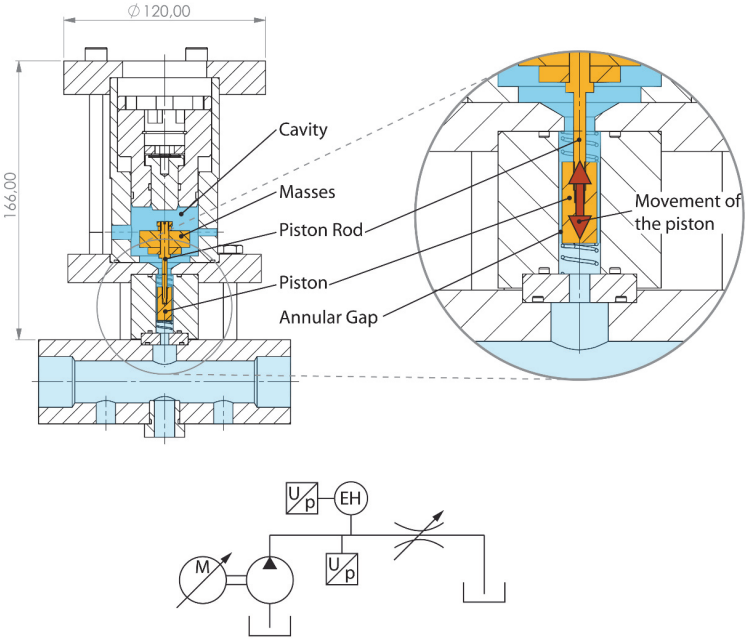


Figure 6: Mechanical structure of the demonstrator (top) and simplified hydraulic circuit diagram of the measurement setup (bottom) (EH...Energy Harvester)

Figure 5 illustrates the time dependent curve and the frequency spectrum of the system pressure and the cavity pressure at a speed of 1500 rpm. As expected, the fundamental frequency of 225 Hz corresponds to the speed of the axial piston pump multiplied by the number of pistons. In addition, several harmonic vibrations occur. Their characteristics depend, among other things, on the structure of the hydraulic system, since reflections and thus standing waves occur. These can be suppressed by the use of a RALAs [10]. Since the pulsation of the system pressure is recorded directly at the energy

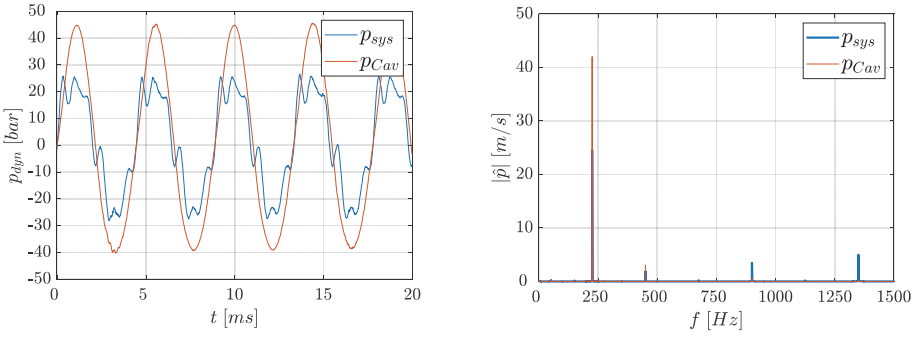


Figure 5: Time dependent curve (left) and frequency spectrum (right) of the dynamic pressure in the system and in the cavity of the Harvester at a pump speed of 1500 rpm and a static pressure of 150 bar. Parameters: $V_{Cav} = 40.9\ cm^3$, $m_P = 0.046\ kg$, $s_{Gap} = 0.08\ mm$, $d_P = 10\ mm$

harvester and the relationship between system and cavity pressure is considered for the validation, the reflection of the pressure waves has no influence on the basic measurements for investigating the behavior of the harvester. Therefore, the use of a RALA is not necessary.

The measurement data can be used to create a transfer function by analyzing the frequency and phase spectra of both the system and cavity pressure. For this purpose, the ratio of the amplitudes and phases of the system pressure at the five harmonics to those of the cavity pressure at identical frequency is determined. In this way, a Bode chart is created based on measurements at different pump speeds. **Figure 7** shows this for the case of a static pressure at 150 bar, a cavity volume of 40.9 cm³, a piston mass of 46 g and an annular gap height of 0.08 mm. Also shown in the diagram is the curve of the analytical transfer function. This was fitted to the measured data on the basis of the phase curve in order to determine the equivalent compression modulus K' and the eccentricity of the piston s_{exz}. It can be seen that the behavior of the harvester can be reproduced in wide ranges by the linear transfer function. Above the natural frequency, however, deviations occur due to nonlinear effects.

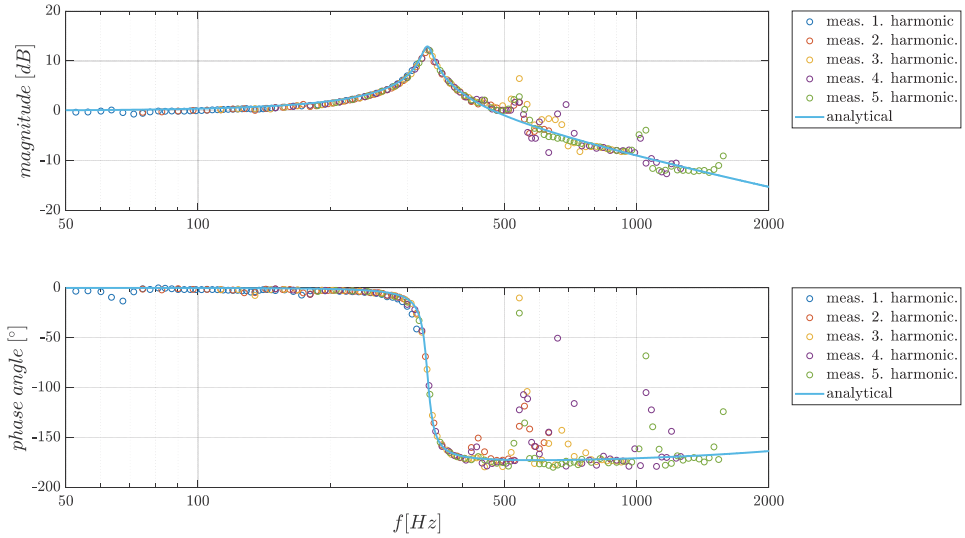


Figure 7: Experimentally and analytically determined Bode chart of $G_v(s)$ at a static pressure of 150 bar. Parameters: $V_{Cav} = 40.9\ cm^3$, $m_P = 0.046\ kg$, $s_{Gap} = 0.08\ mm$, $d_P = 10\ mm$

In the range of the natural frequency of the demonstrator, there is an approximately 16-fold increase in the system pulsation, or an increase of 12 dB. This means that with an assumed pulsation amplitude in the system of 10 bar, the pressure in the cavity pulsates at 160 bar. However, only higher harmonics of the pressure pulsation were employed in experimental measurements to stimulate the resonant frequency, which led to the cavity pressure not achieving such high absolute pulsation. At high-pressure pulsations and correspondingly high piston velocities, there may be damping effects that are not considered analytically, for instance, caused by cavitation in the cavity.

To ensure that the analytical model is also valid at other operating points, additional measurements were carried out with varying parameters. **Table 1** shows the results for these measurements. Using the parameters K' and s_{exz} determined from the fitting function, the time constants T_1 and T_D as well as the damping D of the transfer function $G_p(s)$ can be calculated. The experimentally determined parameters K' and s_{exz} as well as the resulting damping with respect to the static pressure are shown in **Figure 8**.

Table 1: Parameters of the demonstrator an the transfer function at different measurements

meas	mP [g]	sGap [mm]	VCav [cm³]	T1 [ms] (pstat=150bar)	TD [ms] (pstat=150bar)	D (pstat=150bar)
1	46	0,08	40,9	0,477	0,022	0,0256
2	46	0,08	37,6	0,453	0,023	0,0268
3	16	0,08	46,3	0,307	0,008	0,0166
4	16	0,08	41,3	0,300	0,008	0,0169
5	16	0,13	46,3	0,309	0,019	0,0328
6	16	0,13	41,3	0,300	0,019	0,0348

The left side of **Figure 8** displays the changes observed in the equivalent compression modulus regarding static pressure for various harvester configurations. The outcomes show substantial fluctuations between measurements, however, a noticeable increase in the equivalent compression modulus with an increasing static system pressure is observable. This causes a rise in the natural frequency of the harvester. The low values of the equivalent compression modulus compared to the compression moduli reported in the literature (for example, $K_{oil}(200\ bar) \approx 18500\ bar$ [9]) are due to the compliance of the bolts and seals, as well as a the presence of free air. For measurements with an annular gap height of 0.08 mm, the eccentricity is relatively constant and is just below or at the maximum value of 0.08 mm. For measurements with an annular gap height of 0.13 mm, the eccentricity fluctuates more, but does not reach the maximum value. The damping decreases slightly across the static pressure. This decrease can be attributed to the increasing viscosity and subsequent reduction in volume flow through the annular gap (compare equation (4)). The damping increases with increasing gap size, which is in line with expectations.

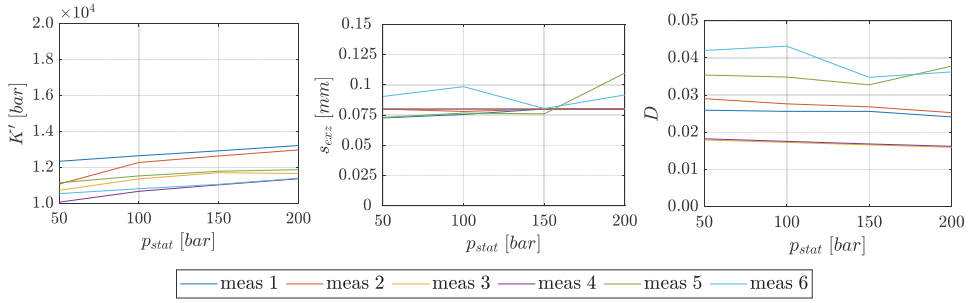


Figure 8: Parameters $K'$and s_{exz} determined using the fitting function and the resulting attenuation for different measurement setups (see **Table 1**)

Overall, it can be seen that the principle behavior of the harvester can be predicted to a large extent via the linear, analytical transfer function. This can also be shown in **Figure 9**, where the comparison between the measured values of the cavity pressure and the cavity pressures calculated via the inverse transfer function $G_p(s)$ is presented.

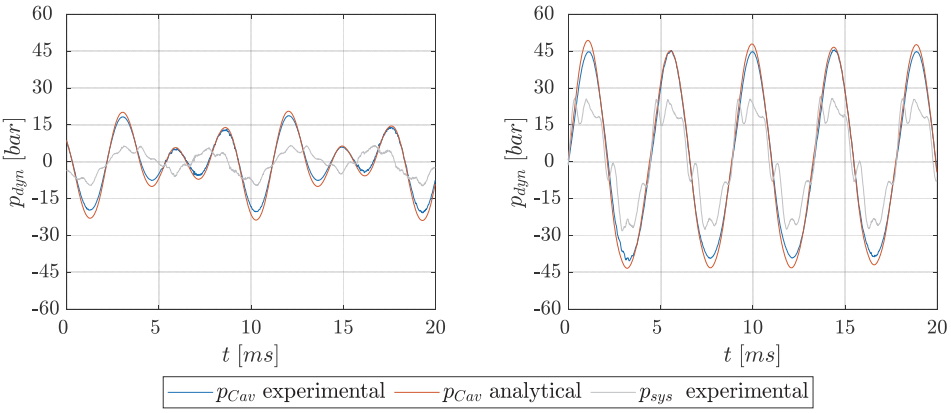


Figure 9: Comparison of the measured data and the analytical calculation of the pressure pulsation in the cavity at a pump speed of 750 rpm (left) and 1500 rpm (right)

5. OPTIMIZATION OF THE PISTON MOVEMENT

The objective of the harvester is to maximize energy extraction from pressure pulsations. An electromagnetic linear generator is implemented to induce voltage through piston movement. In a fundamental coil-solenoid combination, as shown in **Figure 2**, the induced voltage can be determined via the following equation:

$$U_{ind} = -N\frac{d\phi}{dt} = -N\frac{d\phi}{dx}\cdot\frac{dx}{dt} \tag{16}$$

Here N represents the number of turns of the coil, $d\phi/dx$ indicates the magnetic flux gradient in the direction of movement of the piston and dx/dt the piston speed. Using the induced voltage and the ohmic resistance of the coil R_C as well as the load resistance R_L, the power at the load resistance of the generator can be calculated. The inductance of the generator coil is considered negligible due to the low frequency of the piston movement (< 1 kHz).

$$P_{Last} = \frac{R_L}{(R_L + R_C)^2}\, U_{ind}^2 \tag{17}$$

The flux gradient $d\phi/dx$ is reliant on the linear generator's design, and several factors affect it, including the magnet geometry, type, number of windings, and wire diameter. However, these parameters will not be explored in greater depth in the following optimization of the mechanical components. Along with the flux gradient, the generator's performance is also influenced quadratically by the piston speed and will be optimized in the following.

In order to achieve the highest possible energy yield from the harvester, the design is to be optimized with regard to maximum piston speed. The first step is to examine the influence of the cavity volume. Inserting the resonance case, $s = j/T_1$ - which provides the highest velocity - into $G_v(s)$ produces the following equation:

$$G_v\left(s = j\frac{1}{T_1}\right) = \frac{V_{Cav}(\pi A_P r_o s_{Gap}^2 + 6A_{PS}C_{Gap}\eta)}{\pi r_o s_{Gap}(C_{Gap}m_P K' s_{Gap} + A_{PS}V_{Cav}\eta)} \tag{18}$$

As the cavity volume increases, the transfer function and consequently, the piston speed also increase.

This can be explained by the fact that the pressure gradient in the cavity, due to the annular gap volume flow, decreases as the volume increases and so does the resulting damping (refer to equation (15)). This effect becomes more pronounced with increasing annular gap. In practice, however, due to installation space limitations, the cavity volume can only be increased to a limited extent. Therefore, for the following calculations, it is assumed constant with $V_{Cav} = 40\ cm^3$.

Figure 10 shows the maximum speed of the piston as a function of the relative eccentricity of the piston and at various annular gap heights. It is observed that with low ring gaps the piston speed is almost independent of the eccentricity. However, as the ring gaps increase, the piston speed falls more precipitously in relative terms.

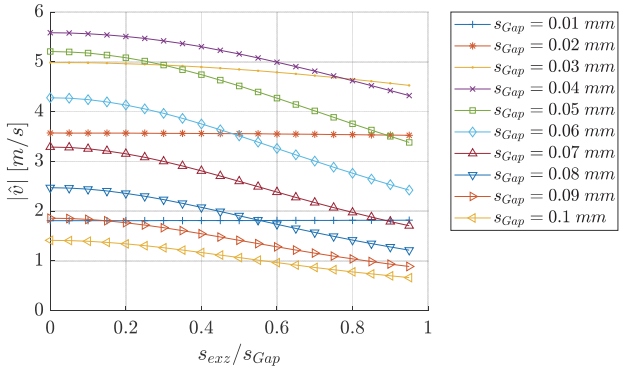


Figure 10: Maximum piston speed as a function of relative eccentricity at different ring gap heights. $m_P = 0{,}05\ kg$, $d_P = 11{,}5\ mm$, $p_{dyn} = 1\ bar$

Figure 11 depicts the behaviour of the maximum piston speed as a function of piston mass and ring gap height at maximum (left) and minimum (right) eccentricity. To ensure comparability, the natural frequency according to equation (19) is assumed to be $f_0 = 225\ Hz$, which aligns with the excitation frequency of a 9-piston pump operating at a speed of 1500 rpm. Additionally, the cavity volume ($V_{Cav} = 40\ cm^3$) remains constant. As a result, with varying piston mass, the piston diameter can be expressed as $d_P = d_P(m_P)$.

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{K' A_P^2}{V_{Cav} m_P}} \approx \frac{1}{2\pi T_1} \tag{19}$$

$$d_P = \sqrt{8 f_0 \sqrt{\frac{V_{Cav} m_P}{K'}}} \tag{20}$$

Figure 11 displays that the optimal mass for the piston rises as the annular gap diminishes. Despite an increase in mass, the absolute value of the respective optimum falls only insignificantly and is just below 6 m/s at an amplitude of the system pressure of 1 bar. The small impact of the mass can be explained by the increase of the piston diameter d_P according to equation (20) alongside its mass, leading to the increase in the excitation force $F_{psys} = p_{sys}\, A_P$. The increase in excitation force can also account for the shift of the optima towards larger masses with decreasing annular gap and the consequent increase in friction. However, it is worth noting that the calculations were carried out with

certain simplifications, and the fact that a larger mass may generate a higher flow resistance under certain circumstances was not taken into consideration.

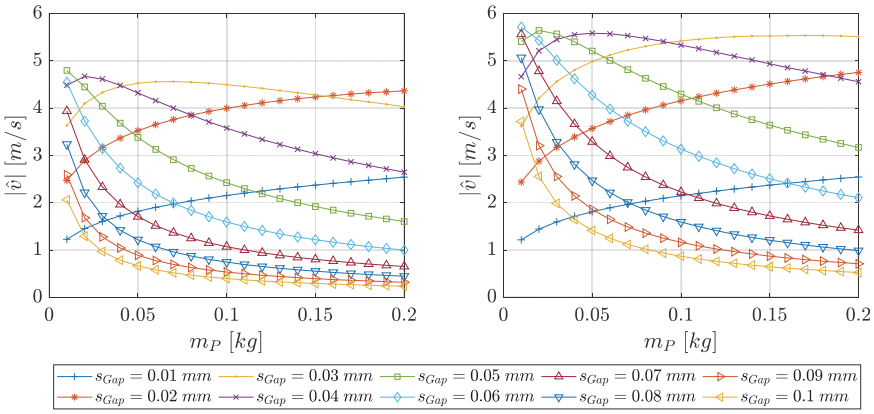


Figure 11: Maximum velocity of the piston as a function of the piston mass and the height of the annular gap at maximum (left) and minimum (right) eccentricity of the piston. Dynamic pressure amplitude of the system $p_{dyn} = 1\ bar$.

6. LINEAR GENERATOR

A linear generator composed of a cylindrical magnet and a coil is utilized to generate energy from the movement of the piston. The simple design ensures robust behaviour and low acquisition costs. The calculations in the previous section show that the maximum speed of the piston is almost independent of the mass under the given assumptions. Therefore, the cylinder magnet can be attached to the piston and the coil can be permanently installed as a stator in the harvester. This significantly simplifies the design of the linear generator and reduces its susceptibility to errors that would otherwise increase due to a moving coil.

An example to demonstrate the achievable performance of the harvester will be briefly explained here. First, the electromagnetic damping of the generator is calculated using the following equation

$$D_{em} = \frac{1}{R_L + R_c + j\omega L_c}\left(N\frac{d\phi}{dx}\right)^2 \tag{21}$$

Here, R_C refers to the ohmic coil resistance, R_L refers to the load resistance, and N signifies the number of windings. The inductance of the generator coil L_c is considered negligible due to the low frequency of the piston movement (< 1 kHz). The magnetic flux gradient $d\phi/dx$ was determined through an FEM analysis of a cylindrical magnet with a diameter of 20 mm and a height of 13 mm, using a coil with 250 turns and a wire diameter of 0.35 mm. The magnet has a mass of 0.03 kg, resulting in the piston having a total mass that corresponds with the experimental measurement of $m_P = 0.046\,kg$. The chosen example exhibits an achieved electromagnetic damping of $D_{em} = 0.21\,Ns/m$ for a load resistance of $R_L = R_C = 6.3\,\Omega$, enabling optimal utilization if D_{em}<< parasitic damping [2]. To determine the transfer function $G_{vDem}(s)$, this must be incorporated into the analytical equation of motion of the piston or linear generator.

$$B_2 = \pi r_o A_P^2 s_{Gap}^2 + 6A_P A_{PS} C_{Gap}\eta \tag{22}$$

$$B_3 = \pi D_{em} V_{Cav} r_o s_{Gap}^2 + \pi A_{PS} V_{Cav} r_o s_{Gap} \eta + \pi m_P r_o C_{Gap} K' s_{Gap}^2 \tag{23}$$

$$G_{vDem}(s) = \frac{\hat{v}(s)}{\hat{p}_{sys}(s)} = \frac{s \cdot \frac{V_{Cav}}{A_P K'} \left(1 - \frac{C_{Gap}}{V_{Cav}} \cdot \frac{B_3}{B_2}\right)}{\frac{\pi V_{Cav} m_P r_o s_{Gap}^2}{K' B_2} s^2 + \frac{B_3}{K' B_2} s + 1} \tag{24}$$

The Power $\hat{P}_{el}$ of the generator is:

$$\hat{P}_{el}(s) = D_{em} \cdot \hat{v}(s)^2 \tag{25}$$

Figure 12 displays the amplitude curve of the energy harvester's power under a pressure pulsation of 1 bar. A power of slightly over 2 W is attained at the natural frequency of 225 Hz. However, the power decreases quickly outside the resonant frequency and is only 15 mW at 200 Hz.

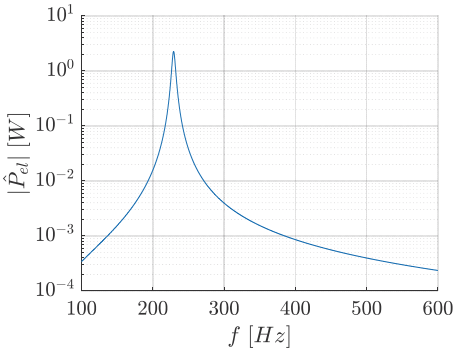


Figure 12: Magnitude plot of the generator power with the following parameters: $p_{dyn} = 1\ bar$, $V_{Cav} = 40\ cm^3$, $m_P = 0.05\ kg$, $s_{Gap} = 0.03\ mm$, $N = 250$, $d_w = 0.35mm$

7. CONCLUSION AND SUMMARY

This paper presents an analytical model for mapping the piston oscillation of a newly developed energy harvester with an electromagnetic generator. The model was validated using several measurements done on a demonstrator. Using the analytical model, it was possible to optimize the mechanical parameters of the energy harvester for maximum piston speed and the resulting high output of the generator. It was shown that the maximum piston speed depends only slightly on the mass of the oscillating piston, thus enabling a design of the linear generator with the coil as stator. Furthermore, the speed of the piston depends significantly on the height of the annular gap. The specific size of the gap depends, among other factors, on the mass of the piston and is situated at $s_{Gap} = 0.03$ mm for $m_P = 0.05\ kg$. In order to calculate the power of the generator, the analytical model was extended accordingly. At a pressure amplitude of 1 bar a power of more than 2 W could be calculated. However, since this is a resonance energy harvester, this only applies to a limited frequency range. Therefore, the application is mainly useful in systems which are driven by a pump with constant speed. Among other factors, adjusting the cavity volume can change the natural frequency of the harvester to be adapted to the specific system.
Due to the relatively low piston speed during validation, as the resonant frequency of the demonstrator was only excited by higher harmonics of the pressure pulsation, further

investigations should be made here. In this context, it must be checked whether damping effects not previously considered occur at higher piston speed. Furthermore, it is necessary to validate the entire analytical model including the linear generator on the basis of measurement data.

This Project is supported by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag.

NOMENCLATUR

A_P	Area of the piston's face	m²
A_{PS}	Surface area of Piston	m²
$B_{1\dots3}$	Auxiliary parameter	
C_{Gap}	Auxiliary parameter annular Gap	m³/(sPa)
D	Dampening	
D_{em}	Electromagnetic Dampening	Ns/m
d_P	Piston diameter	m
d_w	Wire diameter	m
f_0	Natural frequency	Hz
G_p	Transfer function	
G_v	Transfer function	m/(sbar)
K'	Equivalent Bulk modulus	Pa
L_c	Coil Inductance	H
m_P	Piston Mass	kg
N	Number of turns	
p_{dyn}	Dynamic Pressure	Pa
p_{stat}	Static Pressure	Pa
p_{sys}	System Pressure	Pa
p_{Cav}	Cavity Pressure	Pa
R_c, R_L	Coil, Load resistor resistance	Ohm
r_o, r_i	Outer, inner Radius annular Gap	m
s_{Gap}	Height annular Gap	m
T_1, T_D	Time constant	s
v	Piston velocity	m/s
v_f	Flow velocity	m/s
V_{Cav}	Cavity Volume	m³
η	Dynamic Viscosity	Pas
ϕ	Magnetic flux	
ν	Kinematic Viscosity	m²/s
ρ	Density	kg/m³

REFERENCES

[1] M. Ahmad und F. Khan, „Review of vibration based electromagnetic-piezoelectric hybrid energy harvesters“, *Int. J. Energy Res.*, Bd. 45, Dez. 2020, doi: 10.1002/er.6253.

[2] S. Priya und D. J. Inman, *Energy Harvesting Technologies*. New York, NY, UNITED STATES: Springer, 2008. Zugegriffen: 4. Oktober 2021. [Online]. Verfügbar unter: http://ebookcentral.proquest.com/lib/slub/detail.action?docID=417906

[3] M. Safaei, H. A. Sodano, und S. R. Anton, „A review of energy harvesting using piezoelectric materials: state-of-the-art a decade later (2008–2018)“, *Smart Mater. Struct.*, Bd. 28, Nr. 11, S. 113001, Okt. 2019, doi: 10.1088/1361-665X/ab36e4.

[4] C. Cepnik, R. Lausecker, und U. Wallrabe, „Review on Electrodynamic Energy Harvesters—A Classification Approach“, *Micromachines*, Bd. 4, S. 168–196, Juni 2013, doi: 10.3390/mi4020168.

[5] J. J. Aranda, S. Bader, und B. Oelmann, „Self-Powered Wireless Sensor Using a Pressure Fluctuation Energy Harvester“, *Sensors*, Bd. 21, Nr. 4, Art. Nr. 4, Jan. 2021, doi: 10.3390/s21041546.

[6] E. Skow, „Harvesting energy from acoustic pressure fluctuations within hydraulic systems via excitation of piezoelectric stacks“, Mai 2017. Zugegriffen: 2. August 2023. [Online]. Verfügbar unter: https://www.semanticscholar.org/paper/Harvesting-energy-from-acoustic-pressure-within-via-Skow/25508b5889d98cb4478d8a119baab22a7d7683ba

[7] H. Ren und T. Wang, „Development and modeling of an electromagnetic energy harvester from pressure fluctuations“, *Mechatronics*, Bd. 49, S. 36–45, Feb. 2018, doi: 10.1016/j.mechatronics.2017.11.008.

[8] E. Hering, R. Martin, und M. Stohrer, *Physik für Ingenieure*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2021. doi: 10.1007/978-3-662-63177-5.

[9] D. Will und N. Gebhardt, Hrsg., *Hydraulik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. doi: 10.1007/978-3-662-44402-3.

[10] H. Theissen, „Die Berücksichtigung instationärer Rohrströmungen bei der Simlulation hydraulischer Anlagen Thiessen“, RWTH Aachen, 1983.

LOAD HOLDING VALVES WITH INTEGRATED FLOW SENSORS

peer reviewed

Dr. Bernd Zähe[1]*, M.Sc. Florian Müller[2], Lyle T. Hayes-Macaluso[3]

[1, 2] *Sunhydraulik GmbH D-41812 Erkelenz, Brüsseler Allee 2,* [3] *Bachelor of Science in Mechanical Engineering, University of Florida*

** Corresponding author: E-mail address: Bernd.Zaehe@sunhydraulics.com*

ABSTRACT

Counterbalance valves are widely used in mobile hydraulics, they preload the return line of motors and cylinders to ensure the load doesn't move when the proportional control valve is in open center position. They also preload the return line when the cylinder moves to lower a load. In that condition the circuit can be unstable and restrictive counterbalance valves are needed. They stabilize the circuit but require higher inlet pressures. The paper describes a way to stabilize the circuit without adding pressure losses. A flow sensor can be built into the counterbalance. The signal can be used for a closed loop control. A simulation shows that the controller can improve the stability.

An alternative to the flow sensor in the counterbalance valve is a velocity sensor on the cylinder. Again, a controller can use the signal: it closes the loop on the velocity of the cylinder. The paper compares the closed loop controls.

Keywords: Counterbalance, Flowmeter, Stability, Load lowering

1. FUNCTION OF COUNTERBALANCE VALVES

Counterbalance valves are hydromechanical poppet type valves. A common design is a screw-in cartridge with 3 ports. See fig. 1.

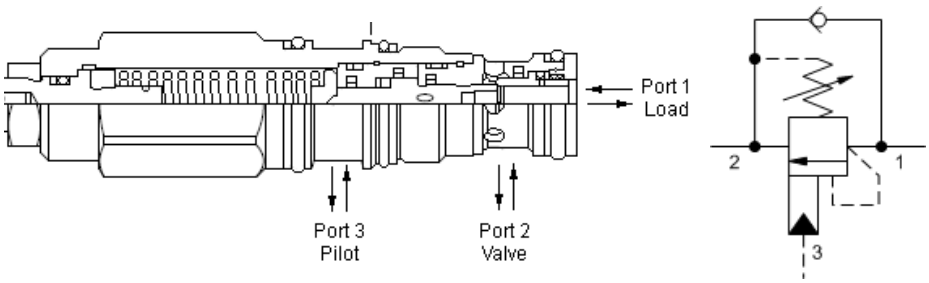


Figure 1: Counterbalance valve as screw-in cartridge

The counterbalance valve acts like a relief valve between port 1 and 2 but has an additional third port. Pilot pressure on port 3 helps to open the valve. The valve is usually built into the hydraulic line downstream of a cylinder that supports an external force if that force pushes or pulls the in the direction of the intended movement (negative load). Fig.2 on the left (A) shows a typical circuit: The cylinder doesn't move if the proportional valve is in the centre position and the counterbalance is set

high enough. A recommended setting is 30% above the highest expected load induced pressure F/A

$$\text{set} = 1.3 * \text{F/A} \tag{1}$$

The setting is the pressure on port 1 that starts to open the valve against an adjustable spring. Since a pilot pressure on port 3 also helps to open the valve, it is important to set the valve with no pilot pressure on port 3. When the proportional valve in the circuit opens, the pressure p_A increases. It works on port 3 of the counterbalance valve. It also works on the cylinder and increases the pressure p_B between cylinder and counterbalance valve. The pressures p_A and p_B that open the counterbalance valve can be calculated from two equations: The force balance across the cylinder:

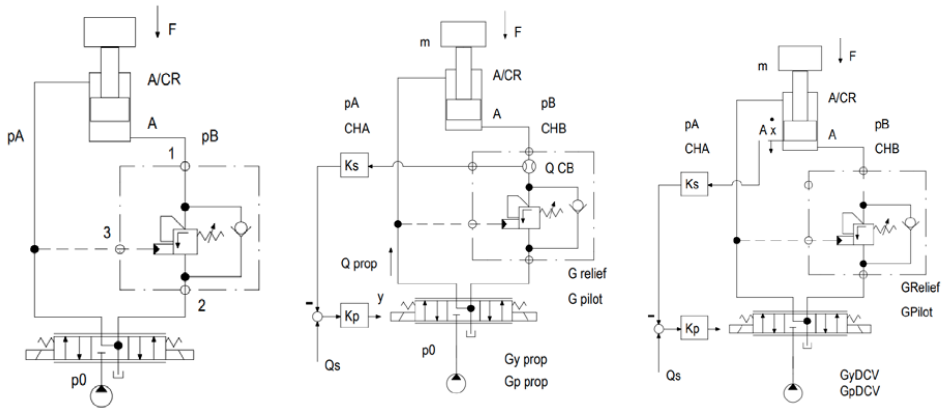
$$p_B = \frac{F}{A} + p_A/\text{CR} \tag{2}$$

and the force balance for the counterbalance valve when it starts to open:

$$\text{set} = p_B + p_A * PR \tag{3}$$

From the equations (2) and (3) the pressure p_A at that the movement starts can be calculated:

$$p_A = \frac{set - F/A}{\frac{1}{CR} + PR} \tag{4}$$



A: Standard circuit B: Circuit with flow feedback C: Circuit with cylinder speed feedback

Figure 2: Circuits with load reactive counterbalance valves

The pressure $p_A$in equation (4) must be positive to avoid cavitation. A low pressure is desirable because the inlet pressure p_A determines the required pressure to move the cylinder.
The difficulty in selecting a counterbalance valve is in finding the one that is stable (low pilot ratio, restrictive) but also doesn't generate excessive pressure losses.

This paper shows a counterbalance valve with a built-in flowmeter. The goal is using the flow signal in a simple closed loop (fig. 2B) to improve the performance and achieve stability with a low inlet pressure. The paper compares the stability of that circuit to a circuit with a feedback of the cylinder speed (fig. 2 C).

2. COUNTERBALANCE VALVE WITH INTEGRATED TURBINE AS FLOW SENSOR

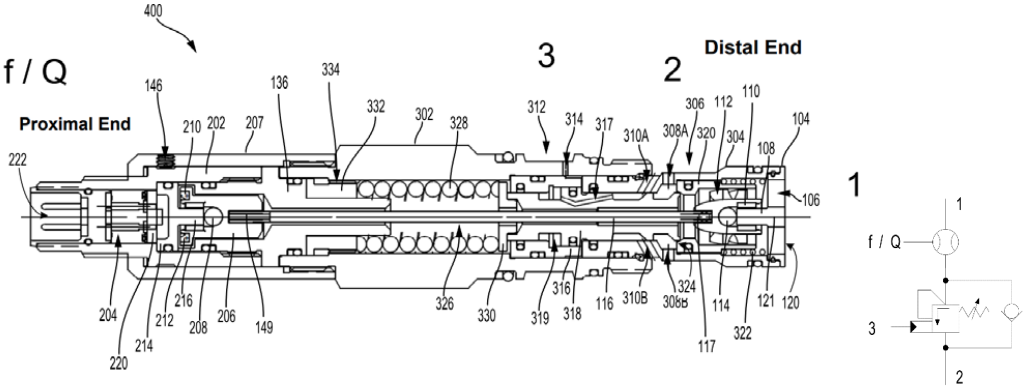


Figure 3: Cross section of an adjustable counterbalance valve with integrated turbine as flowmeter (Pat. Pending)

Fig. 3 shows a cross section of a counterbalance with the usual porting and function. The load is on port 1. Pilot pressure on port 3 moves the poppet (317) against a spring to open a flow path from 1 to 2. A turbine (112) near port 1 starts to spin when the valve opens. A long shaft (326) connects the turbine (320) to a wheel (206) with permanent magnets (219). A hall sensor on a circuit board (220) detects the speed of the wheel that is proportional to the flow. A digital signal is available through an M12 connector (222) at the proximal end of the counterbalance valve.
The valve has not been built as shown. But a very similar turbine in a different valve showed the correlation between the measured rpm of the turbine and the flow measured through a fixed displacement gear flow meter. The turbine reacts within about 10 milliseconds to changing flows.
Customers are interested in flowmeters if they can replace expensive sensors on long telescopic cylinders, but they are skeptical about the benefits of the integrated flow sensor with respect to stability. A common practice is using encoders to measure the position of work platforms. The signal is required for the correct positioning of the platform, but the circuit is still known for low damping. So, the question came up, whether measuring the flow instead of the movement of the cylinder could be more useful to improve the performance.

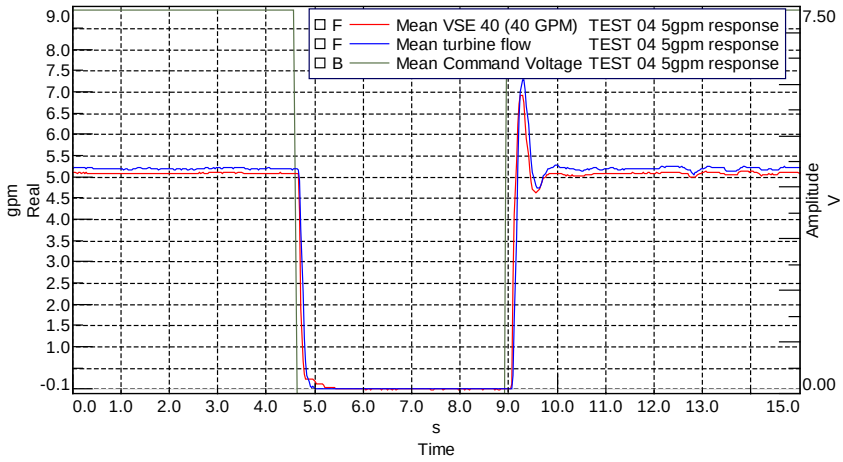


Figure 4: Flow measured with a gear flow meter (red) and with a turbine (blue), built into a cartridge.

3. LINEARISED MODEL OF A LOAD LOWERING CIRCUIT WITH COUNTERBALANCE VALVES AND FLOW SIGNAL FEEDBACK

In a previous paper [1] a mathematical, linear model was presented for a circuit with counterbalance valves. The paper also showed counterbalance valves with adaptive pilot ratios that have a low pilot ratio only in operating points where stability is critical. The overall efficiency of circuits with adaptive pilot ratios is higher. The adaptive counterbalance valves have been introduced into the market.

This paper describes another approach to improve load lowering circuits: The goal is finding out whether the signal from a flowmeter near the or built into the counterbalance could be used to improve the stability of the circuit. The goal is again to reduce the inlet pressure p_A when the load is being lowered and maintain stability. As in the previous paper, a state-space representation is chosen to calculate the stability of the circuit.

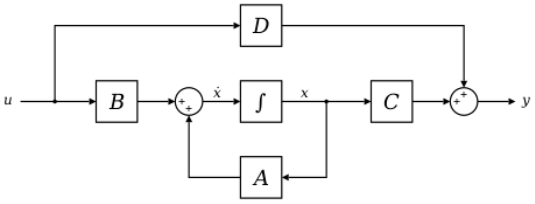


Figure 5: State-space representation

Figure 5 shows the general format: 'A' is the system matrix, 'B' the input matrix and 'C' the output matrix. We want to model the hydraulic system shown in fig. 2 B. The feedthrough matrix D is zero. Figure 2B describes the same circuit as in the previous paper [1] except for a sensor for the flow across the counterbalance valve. So, we can reuse the model for all components except the additional sensor. Please refer to that paper for the complete modelling of the circuit. The block diagram in

figure 6 shows the circuit including the new sensor and the controller. The linearized flow Q_{CB} through the counterbalance valve is a function of pilot pressure p_A and load pressure p_B :

$$Q_{CB} = G_{pilot} * p_A + G_{relief} * p_B \tag{5}$$

The assumption is that a sensor measures that flow without delay. The block diagram shows the sensor with a factor K_S . Changing that value from 0 to 1 activates the sensor. The parameter K_P describes the controller for the proportional valve. The simple proportional controller opens the proportional valve without delay. The variable y is the stroke the proportional valve:

$$\mathrm{y} = K_p * (Q_s - Q) \tag{6}$$

The linearized flow across the proportional valve is a function of stroke y and the pressure differential $p_O - p_A$:

$$Q_{Prop} = Gy_{prop} * y + Gp_{prop}(p_0 - p_A) \tag{7}$$

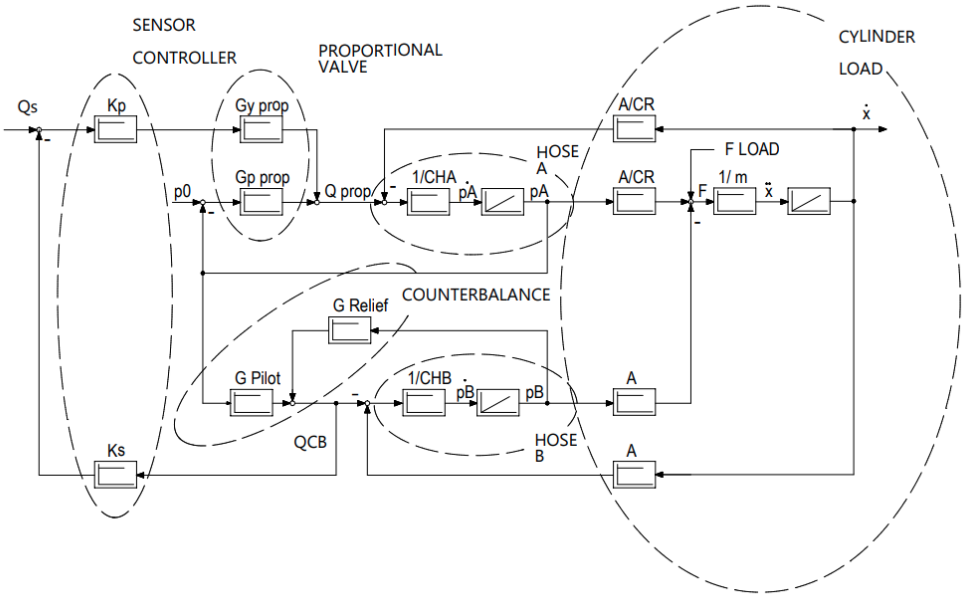


Figure 6: Block-diagram for the circuit in figure 2B.

Energy is stored in three components of the circuit: potential energy is stored in the capacitances CH_A, CH_B of the hoses and the attached volumes of the cylinder with their pressures p_A and p_B , kinetic energy is stored in the masses of the cylinder and the attached load with their velocity $\dot{x}$. So, the state variables of the vector x in figure 5 are p_A , p_B ,and $\dot{x}$. As input variable u in figure 5 we choose the desired flow Q_s. Other input variables could be p_0, and F_{Load} with their input matrices B.

$$\begin{pmatrix} \dot{p}_A \\ \dot{p}_B \\ \ddot{x} \end{pmatrix} = A * \begin{pmatrix} p_A \\ p_B \\ \dot{x} \end{pmatrix} + \begin{pmatrix} \frac{Gy_{prop} * Kp}{C_{HA}} \\ 0 \\ 0 \end{pmatrix} * Q_s + \begin{pmatrix} \frac{Gp_{prop}}{C_{HA}} \\ 0 \\ 0 \end{pmatrix} * p_0 + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{m} \end{pmatrix} * F_{Load} \quad (8)$$

In equation (8) A is the system matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

$$\begin{pmatrix} -\frac{Gp_{prop} + Kp * K_s * G_{pilot} * Gy_{prop}}{C_{HA}} & \frac{Kp * K_s * G_{Relief} * Gy_{prop}}{C_{HA}} & -\frac{A}{CR * C_{HA}} \\ -\frac{G_{Pilot}}{C_{HB}} & -\frac{G_{Relief}}{C_{HB}} & \frac{A}{C_{HB}} \\ \frac{A}{CR * m} & -\frac{A}{m} & 0 \end{pmatrix} \quad (9)$$

In this paper we assume that the variables Q_s , p_0, and F_{Load} are constant and independent of each other. That is why they are called input variables and not state variables. So, the paper does not describe instabilities of systems with load sensing systems for instance where the supply pressure p_0 is a function of the load pressure. Also, external forces on the cylinder F_{Load} that change with velocity -cylinder friction for instance- or stroke have an influence on stability but are not considered in this model.

4. STABILTY CRITERION OF THE LOAD LOWERING CIRCUIT WITH FLOW SIGNAL FEEDBACK

Whether the hydraulic system is stable or not depends solely on the system matrix A. The goal is to find out how the feedback of the flow signal and the controller influence the stability. The values of the elements a_{11}, a_{22} , a_{33} on the main diagonal have a positive influence on stability. So, K_p and K_s in the the a_{11} element can contribute to stability in combination with positive values for G_{pilot} and Gy_{prop}. G_{pilot} describes the change in flow across the counterbalance valve per change in pilot pressure, Gy_{prop} describes the change in flow across the proportional valve per stroke. Both gains are usually positive. So, the parameters K_p and K_s in the a_{11} element can have a similar influence as the parameter Gp_{prop}. That gain is the change in flow per change of the pressure differential across the valve.

$$Gp_{prop} = {dQ_{prop}}/{d(p_0 - p_{A)}} \quad (10)$$

An infinitely high Gp_{prop}describes a reducing valve that keeps the pressure constant. That is obviously good for stability since the counterbalance valve would see a constant pilot pressure. The parameters in the elements of the main diagonal have a positive influence on stability if they appear only on the main diagonal. But the controller and sensor are also seen in the a_{12} element of the system

matrix. Therefore, the Routh criterion has been used to calculate whether the system is stable. It uses the coefficients of the characteristic equation.

$$a_4 * s^3 + a_3 * s^2 + a_2 * s + a_0 = 0 \tag{11}$$

That equation can be calculated from the system matrix:

$$\det(A - s * I) = 0 \tag{12}$$

The Routh criterion leads to three inequations. The system is stable if:

$$\begin{aligned} &\frac{A^2 * \left(C_{HA} * C_{HB} * G_{relief} * Gy_{prop} * CR + C_{HB}^2 * G_{pilot} * Gy_{prop}\right) * (K_P * K_S)^2 * G_{Pilot} * Gy_{prop}}{NOM} \\ &+ \frac{A^2 * C_{HA}^2 * G_{relief} * CR^2 * G_{pilot} * Gy_{prop} * K_P * K_S}{NOM} \\ &+ \frac{C_{HA} * C_{HB} * G_{Pilot} * CR}{NOM} \\ &+ \frac{C_{HB}^2 * Gp_{prop} * G_{pilot} * Gy_{pilot} * K_P * K_S}{NOM} + \frac{G_{relief} * Gp_{prop}}{C_{HA} * C_{HB}} > 0 \end{aligned} \tag{13a}$$

With

$$\begin{aligned} NOM = {} & C_{HA} * C_{HB} * CR^2 * m \\ & * \left(C_{HB} * G_{Pilot} * Gy_{prop} * K_P * K_S + C_{HB} * Gp_{prop} + C_{HA} * G_{relief}\right) \end{aligned} \tag{13b}$$

And

$$\frac{G_{Pilot} * Gy_{prop} * K_s * K_P}{C_{HA}} + \frac{Gp_{prop}}{C_{HA}} + \frac{G_{relief}}{C_{HB}} > 0 \tag{14}$$

And

$$\left(G_{relief} - G_{pilot} * CR\right) * \left(1 - K_P * K_S * Gy_{prop} * CR\right) + Gp_{prop} * CR^2 > 0 \tag{15}$$

The equations 13-15 are identical with equations 6-10 in the previous paper [1] for K_S = 0. Equations 13 and 14 show that the additional controller has no negative effect on stability. The factor $K_P * K_S$ is additive to parameters that need to be positive to achieve stability. But all three inequations 13-15 need to be fulfilled for the system to be stable. Equation 15 shows:

- The term $\left(G_{relief} - G_{pilot} * CR\right)$ needs to be positive. It means there is an upper limit for G_{pilot}, which describes the change in flow across the counterbalance valve per change in pilot pressure. Counterbalance valves with a with a high pilot ratio tend to be instable in load lowering circuits. Valves with lower pilot ratio and more restrictive valves are more stable but more power is wasted in the circuit to lower a load.

- The term $\left(1 - K_P * K_S * Gy_{prop} * CR\right)$ needs to be positive. It means there is an upper limit for the factor K_P of the controller. We can't conclude from equations 12-13 that the flow feedback improves stability under all circumstances.

- The term $Gp_{prop} * CR^2$ needs to be positive. Gp_{prop} is the change in flow across the proportional valve per change in pressure differential. The parameter can describe a reducing valve with a high change in flow and a pressure compensated flow control with a low change in flow. The disadvantage of using a reducing valve in the load lowering circuit is that the velocity of the cylinder is difficult to predict. Sensing the flow and closing the loop allows the application of reducing valves or proportional valves with a low flow gain Gp_{prop}. The flow gain of proportional valves is low at a low pressure-differential across the proportional valve -see [1]. So, less restrictive counterbalance valves can be applied if stability is achieved with a lower flow gain of the proportional valve: that lowers the inlet pressure and saves energy.

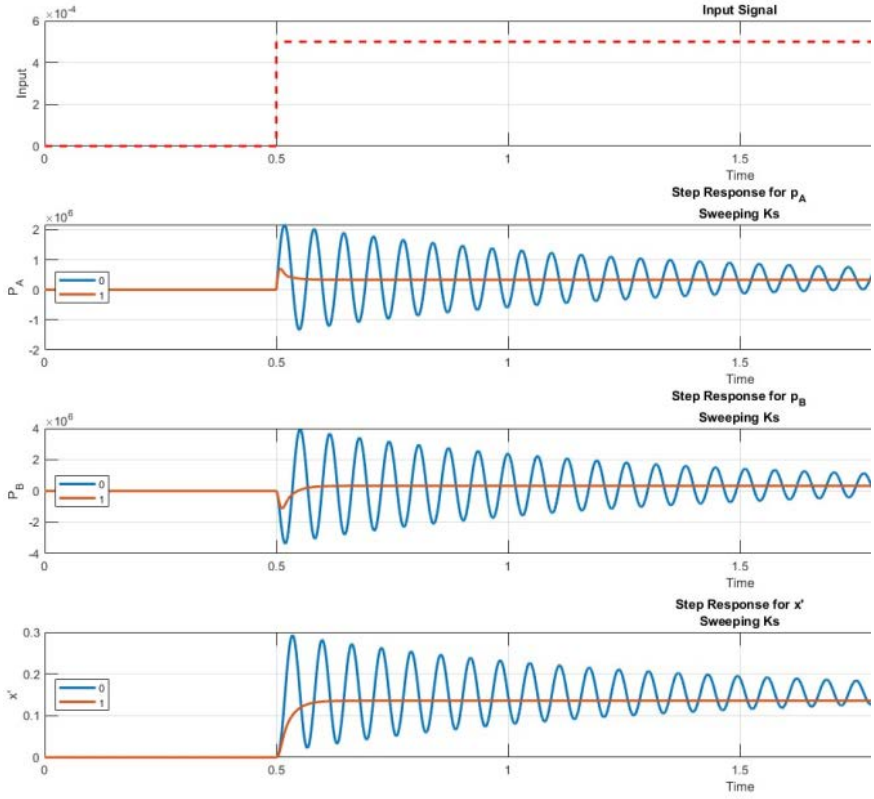


Figure 7: Step response for circuit - see fig. 3 -with ($K_S = 1$) and without ($K_S = 0$) flow feedback. $A = .001256\, m^2, \boldsymbol{C_{HA} = 1e^{-12}\, m^3/Pa}, C_{HB} = 7.59e^{-13}\, m^3/s * Pa$, $CR = 1$, $Gp_{prop} = 1.33\, e^{-11}\, m^3/(s * Pa)$, $G_{pilot} = 3.5\, e^{-10}\, m^3/(s * Pa)$, $G_{relief} = 1.65\, e^{-10}\, m^3/(s * Pa)$, $Gy_{prop} = .133\ m^3/(s * m)$, $K_p = 3\ m * s/(m^3), m = 500\, kg$

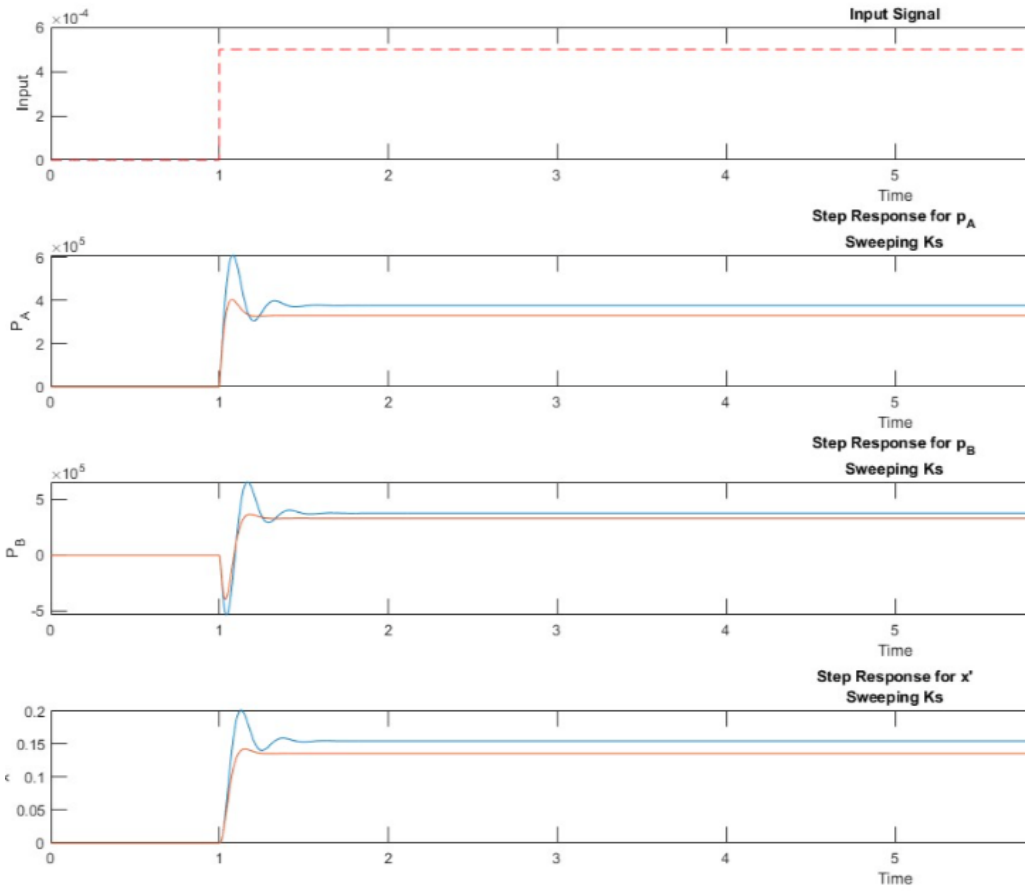


Figure 8: Step response for circuit - see fig. 3 -with ($K_S = 1$) and without ($K_S = 0$) flow feedback. $A = .001256\ m^2, \boldsymbol{C_{HA} = 1.5e^{-11}\ m^3/Pa}, C_{HB} = 7.59e^{-13}\ m^3/(s * Pa)$, $CR = 1$, $Gp_{prop} = 1.33\ e^{-11}\ m^3/(s * Pa)$, $G_{pilot} = 3.5\ e^{-10}\ m^3/(s * Pa)$, $G_{relief} = 1.65\ e^{-10}\ m^3/(s * Pa)$, $Gy_{prop} = .133\ m^3/(s * m)$, $K_p = 3\ m * s/m^3, m = 500\ kg$

Figure 7 shows how the step response to a changing flow command Q_s changes when the feedback signal is introduced. For the blue curves K_S is 0, so, the sensor is not active. For the orange curve K_S is 1. From top to bottom we see the input signal Q_s , pressures p_A, p_B, and the velocity $\dot{x}$ of the cylinder. The parameters are taken form a simulation in a program that has been presented in [1]. The parameters in SI units describe a hydraulic circuit with realistic numbers (cylinder diameter $28\ mm$, pilot ratio of the counterbalance valve $G_{pilot}/G_{relief} = 2\ldots$) The intention is not to simulate a real step response with best possible accuracy but to show the influence of the flow feedback under otherwise unchanged conditions.

In the example the damping of the system improves a lot with the additional flow feedback. That was not to be expected: closing a loop usually improves the accuracy, but often it comes at a price: stability is reduced or can only be maintained with a good control algorithm. In this case the performance of the load holding circuit improves with a simple proportional controller.

Figure 8 shows the same comparison of the circuit, but the capacitance C_{HA} is much larger. As a result, the positive influence of the flow feed back is weaker. In systems with long hoses between the proportional valve and the cylinder a controller on that valve can't influence the stability of the circuit much.

6. STABILTY OF THE LOAD LOWERING CIRCUIT WITH CYLINDER SPEED SIGNAL FEEDBACK

It seems natural to measure the speed of the cylinder instead of the flow across the counterbalance valve. The stability has been calculated for that alternative circuit, shown in fig. 1C, figure 8 shows the block diagram.

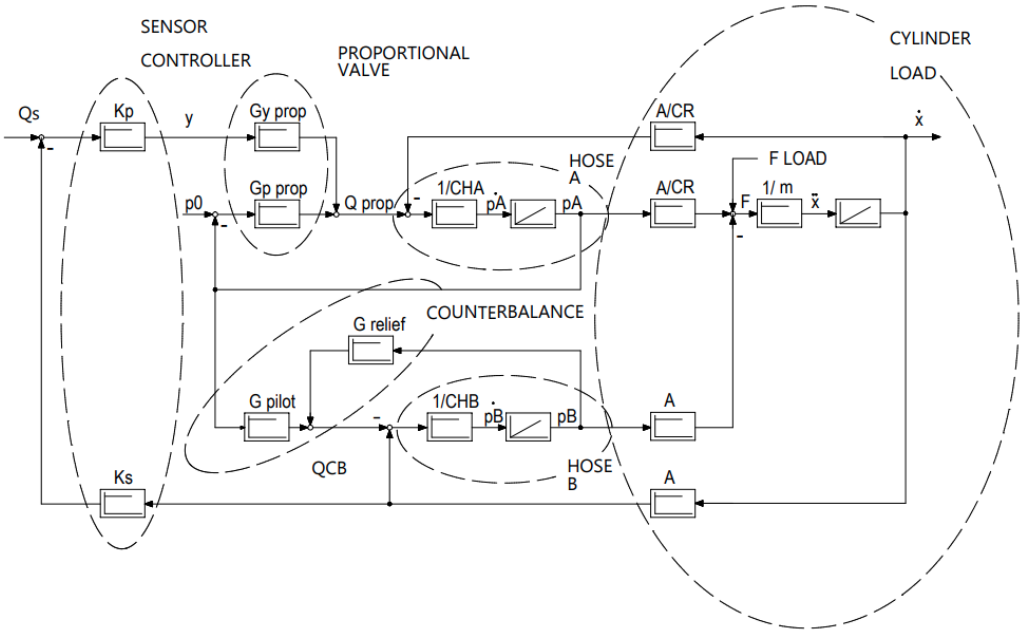


Figure 9: Block diagram of the load holding circuit with feedback of the cylinder velocity $\dot{x}$

The state variables are the same as in the system with flow feedback but the system matrix A is different, the controller and sensor appear in the a_{31} element, not in the a_{11} and a_{21}element:

$$\begin{pmatrix} \dot{p}_A \\ \dot{p}_B \\ \ddot{x} \end{pmatrix} = A * \begin{pmatrix} p_A \\ p_B \\ \dot{x} \end{pmatrix} + \begin{pmatrix} \frac{Gy_{prop} * K_S * K_S}{C_{HA}} \\ 0 \\ 0 \end{pmatrix} * Q_s + \begin{pmatrix} \frac{Gp_{prop}}{C_{HA}} \\ 0 \\ 0 \end{pmatrix} * p_0 + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{m} \end{pmatrix} * F_{Load} \qquad (15)$$

With:

$$A = \begin{pmatrix} -\frac{Gp_{prop}}{C_{HA}} & 0 & -\frac{A*\left(\frac{1}{CR}+K_S*K_P*Gy_{prop}\right)}{C_{HA}} \\ -\frac{G_{pilot}}{C_{HB}} & -\frac{G_{relief}}{C_{HB}} & \frac{A}{C_{HB}} \\ \frac{A}{CR*m} & -\frac{A}{m} & 0 \end{pmatrix} \tag{16}$$

As a result, the stability and the step responses are different. The feedback of the flow across the counterbalance gave better results than the feedback of the cylinder speed. Figure 10 shows a step response in that the system is at least stable with velocity feedback. The capacitance C_{HA} had to be chosen much smaller to avoid instability.

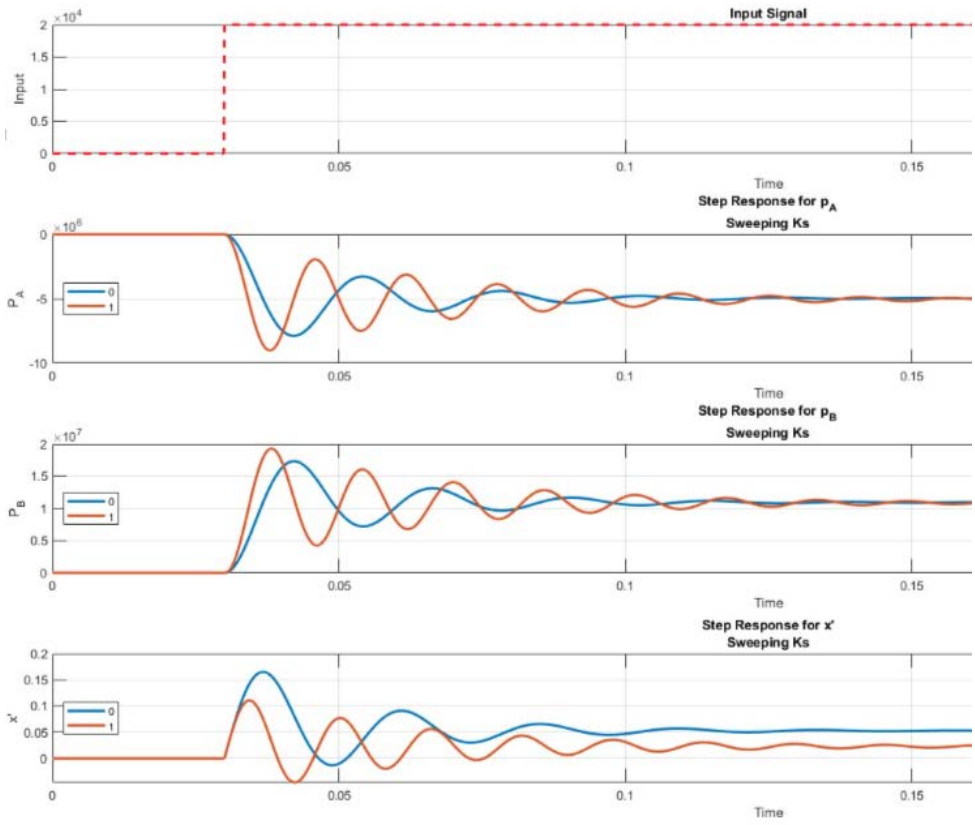


Figure 10: Step response for circuit - see fig. 8 -with ($K_S = 1$) and without ($K_S = 0$) cylinder speed feedback.

$A = .001256\ m^2, \boldsymbol{C_{HA} = 1.5e^{-13}\ m^3/Pa}, C_{HB} = 7.59e^{-14}\ m^3/(s*Pa)$, $CR = 1$, $Gp_{prop} = 1.33\ e^{-11}\ m^3/(s*Pa)$, $G_{pilot} = 3.5\ e^{-10}\ m^3/(s*Pa)$, $G_{relief} = 1.65\ e^{-10}\ m^3/(s*Pa)$, $Gy_{prop} = .133\ m^3/(s*m)$, $K_p = 10\ m*s/m^3$, $m = 500\ kg$

The system matrix in equation 16 (cylinder speed feedback) compared to the matrix in equation 9 is an indicator for the poor results of the velocity feedback. With cylinder velocity feedback there is no additional term on the main diagonal. The velocity changes after pressures p_A and p_B changed. The high inertia of the attached load delays the acceleration of the cylinder. So, when the counterbalance valve opens too far the cylinder velocity reacts with a delay. A turbine built into or near the counterbalance valve gives an earlier signal when the counterbalance valve opens too far. Therefore, a controller that uses that signal can stabilize the system more effectively.

SUMMARY

A counterbalance valve can be built with an integrated turbine as a flowmeter. A controller for a proportional valve can use that signal in a load lowering circuit.
The circuit is more stable and more accurate with the flow signal used for a simple proportional controller but there is an upper limit for the gain of the controller.
A reducing valve is more stable in load lowering circuit but is rarely used since the speed of the cylinder is difficult to predict. That problem can be solved with a flow feedback signal. The controller opens the reducing valve until the desired flow passes through the counterbalance valve.
Measuring the flow and using that signal for a closed loop gives better results with respect to stability than measuring the velocity of the cylinder and using that for a closed loop control since the cylinder with the attached mass is a slower flow meter than the turbine in or near the counterbalance valve.

NOMENCLATURE

A	Area	m^2
C_{HA}, C_{HB}	Capacitance	m^3/Pa
CR	Cylinder ratio	1
G_{relief}	Counterbalance: delta flow per delta load pressure	$m^3/(s*Pa)$
G_{pilot}	Counterbalance: delta flow per delta pilot pressure	$m^3/(s*Pa)$
Gy_{prop}	Proportional valve: delta flow per delta stroke	$m^3/(s*m)$
Gp_{prop}	Proportional valve: delta flow per delta pressure differential	$m^3/(s*Pa)$
K_S	Flow sensor factor	1
K_P	Prop controller factor	$m*s/m^3$
p_A, p_B	Pressure in Volume A, B	Pa/m^2
PR	Counterbalance pilot ratio (effective area for pilot pressure per effective area for load pressure)	1

REFERENCES

[1] Zähe B., Anders P, Ströbel S., (2016) New, Energy Saving Counterbalance Valve with Adaptive Pilot Ratio, IFK Dresden.

[2] Riccardo Bianchi, Steven T. Weber (2016) Hydraulic system with a counterbalance valve as a meter-out element and controlled by an independent pilot signal, Patent No.: US 10,920,799

Chapter 23

Hydrogen

HYDROGEN POWERED HYDRAULIC POWERPACK

peer reviewed

Dipl.-Ing. Lukas Trommler[1]*, Dipl.-Ing. Frank Hänel[1], Prof. Dr.-Ing. Frank Will[1]

[1]*Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden*

* Corresponding author: Tel.: +49 351 463-39478; E-mail address: lukas.trommler@tu-dresden.de

ABSTRACT

A general challenge in the construction machinery sector is that 90% of the current fleet of machines uses diesel as fuel. In contrast to the positive developments in the field of particle mass and nitrogen oxide emissions, CO_2 emissions are stagnating at a constantly high level. The use of hydrogen fuel cell technology results in CO_2 neutral operation and is in line with current CO_2 reduction policies. A hydrogen driven powerpack is used to generate hydraulic power, which is then used to drive a stationary concrete pump. This paper deals with the transformation of a conventional drive system to a fuel cell drive system. Furthermore, two different fuel cell operating strategies are discussed. The challenge in using fuel cells is that due to the efficiency and the balance of plant components about 50 % of the electrical power is converted into heat. Furthermore, the volumetric energy density of hydrogen currently limits the unrestricted use of the CO_2 neutral fuel. As a result, a sufficiently large cooling system must be kept in place. If the same performance and the same operating time shall be achieved as with a conventional diesel combustion machine, it results an increase of the installation space by a factor up to 8 for a fuel cell drive. For the use of such machines, it is important to ensure that future construction sites have an appropriately prepared H_2 infrastructure.

Keywords: Construction Machinery, Fuel Cell, Hydrogen, Powerpack

1. INTRODRUCTION AND MOTIVATION

Hydrogen as an energy carrier is playing an increasingly important role in the implementation of future CO_2 neutral drive technologies. The construction industry is currently facing structural changes and needs to deal with the new challenges that come along with it. The new challenges are significantly shaped by legislative regulations, by society and by the increasing awareness of environmental protection. The Federal Republic of Germany has committed itself to reduce greenhouse gas emissions (GHG) by 65 % by 2035 compared to the year 1990. Furthermore, the target is to achieve 100 % greenhouse gas neutrality by 2045 [1].

To achieve this ambitious goal, a change is needed in all social, economic and technical sectors of a society. The current state of the art shows that in construction machinery industry, 90 % of all existing machines still use diesel as a fossil fuel [2]. In order to fulfil all requirements, new ways, systems and machine technologies must be defined, implemented and launched on the market to accomplish the set goal. Driven be the regulation of the European Union (EU) for exhaust aftertreatment, a very large part of the reduction in local emissions has already been successfully achieved in recent years. The following **Figure 1** shows the development of emissions over time between the years 1980 to 2050. The figure is based on the Transport-Emission-Model for Mobile Machinery (TREMOD-MM), Version 5.1, and includes all non-road mobile machinery in Germany. The emission model takes into account the following sectors: Agriculture-, Forestry-, Construction-, Residential- and other non-road mobile machinery (NRMM) industry branches. In this diagram, the temporal course of the particulate mass (PM), the nitrogen oxides (NO_x) and carbon dioxide (CO_2) are shown graphically. During the

past few years, the machines have always been adapted to the next exhaust emission stage. That is the reason why there is a significant decrease in PM- and NO_x- emissions. The amount of PM- emissions reduced between 1980 and 2018 by 79 % and the amount of NO_x- emissions by 48 %. Against this positive trend, the CO_2 balance stagnates at a constantly high level over time. As a major driver of climate change, CO_2 will be one of the main challenges for construction machinery [3].

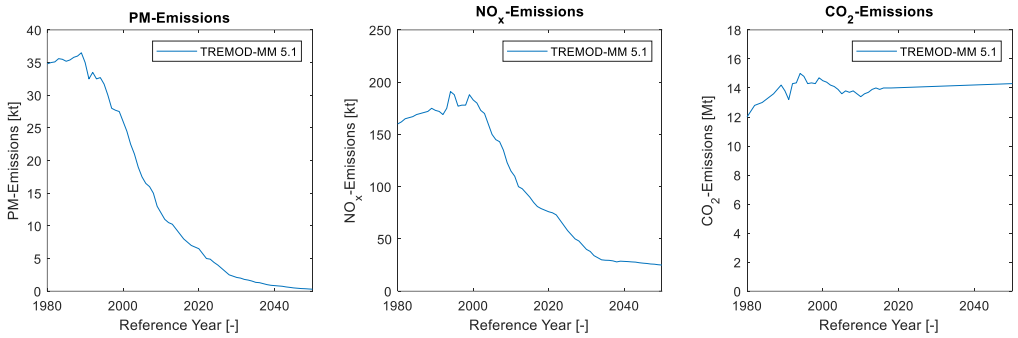


Figure 1: PM-, NO_x- and CO_2- emissions between 1980 - 2050 [3]

This issue must be countered by alternative drive technologies, like battery or hydrogen driven systems. In the low power range up to 80 kW, battery driven machines are available on the market today. For higher power, cable based solutions are the state of the art. The limits here are the available electrical load connection and the cable routes interfering with the construction workflow. The wide range of applications for construction machinery requires outputs up to 1000 kW. Especially the hydraulic power density needs to be maintained for the vast majority of construction machinery. This will not fundamentally change even with the proactive electrification of NRMM.

In the research project “Hydrogen 2 Hydraulics” at TU Dresden / Chair of Construction Machinery, the use of hydrogen together with a fuel cell as a primary power source for hydraulic powered construction machinery is being investigated. The goal is to develop a hydraulic powerpack that provides required hydraulic power with the help of a H_2 fuel cell. A semi-stationary concrete pump is used as an application scenario. Using this machine, the various challenges in the field of H_2 technology are presented in this paper.

2. STATE OF THE ART

A fuel cell is an electrochemical energy converter that converts the internal energy of a fuel into electrical energy and heat. Various types of fuel cells have established their presence on the market. Depending on the fuel cell type, different fuels are used to run the machines. According to this classification, there are fuel cells that use hydrogen (H_2), methanol (CH_3OH) or methane (CH_4) as fuel. Furthermore, the fuel cell types differ in their reaction temperatures. **Table 1** summarizes most common fuel cell types, their fuel and operating temperatures. The functional principle of a fuel cell is explained by using a polymer electrolyte membrane fuel cell (PEMFC) as an example. The reason for that is, because this type of fuel cell has become established in most mobile machine applications and is therefore also used in the research project mentioned above. [4]

Table 1: Fuel cell classification [4]

Fuel cell type	Acronym	Fuel	Operating temperature [°C]
Proton exchange membrane fuel cell	PEMFC	H_2	60 - 160*
*Low temperature PEMFC	LT - PEMFC	H_2	60 - 80
*High temperature PEMFC	HT - PEMFC	H_2	120 - 160
Direct methanol fuel cell	DMFC	CH_3OH	60 - 120
Alkaline fuel cell	AFC	H_2	60 - 80
Molten carbonate fuel cell	MCFC	CH_4	650
Solid oxide fuel cell	SOFC	CH_4	800 - 1000

The core of each fuel cell is formed by two electrodes, the anode (+) and the cathode (-), which are separated from each other by a semi-permeable polymer electrolyte membrane. The hydrogen fuel (H_2) and oxygen (O_2) are continuously supplied to the reaction chambers of each single cell from the outside via the flow channels of the bipolar plates arranged on both sides. Each electrode is coated with an electrolyte consisting of a noble metal layer. This noble layer is usually platinum or a platinum alloy. The electrolyte has a high porosity and consequently increases the active area which is available for the chemical reaction. [4][5]

With the help of **Figure 2**, the functioning of a PEMFC is explained in detail. On the anode side, hydrogen is supplied and diffuses through the Gas Diffusion Layer (GDL) to the catalyst. At the catalyst layer, the hydrogen molecule is absorbed and oxidizes to a positively charged hydrogen proton (H^+) while it releases a negative charged electron (e^-). This reaction is called oxidation and can be written as formula (1). The semi-permeable membrane allows only the protons to diffuse towards the cathode side. The electrons need to take an external by-pass way. As a result of the external electrical flow, electrical work is performed, which can be used. [5]

On the cathode side, the oxygen atom recombines with two electrons and two protons. This reaction is called reduction and can be written as (2). The uptake of the electrons and protons consequently leads to the formation of a water molecule (H_2O), which is discharged to the outside of a cell. The total reaction, shown in (3), is called redox reaction. [5]

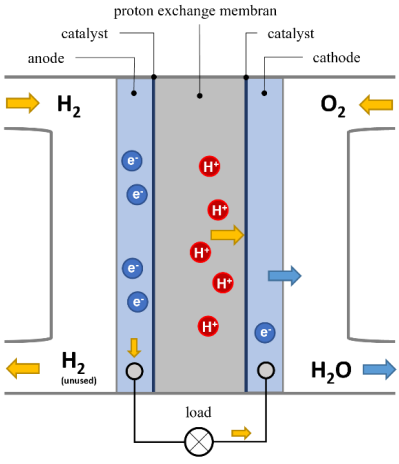


Figure 2: Working principle of a PEMFC [6]

Table 2: Fuel cell reactions

Location	Reaction	Formula	
Anode	Oxidation	$2H_2 \rightarrow 4H^+ + 4e^-$	(1)
Cathode	Reduction	$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$	(2)
Total	Redox reaction	$2H_2 + O_2 = 2H_2O$	(3)

Each fuel cell has a reversible cell voltage E^0 of 1,229 V under thermodynamic steady-state conditions and can be calculated according to equation (4). The basis of the calculation is the change in the free reaction enthalpy $\Delta_R G_M^0$, the charge number z and the Faraday constant F. [5]

$$E^0 = \frac{\Delta_R G_M^0}{z * F} = 1{,}229\ V \tag{4}$$

The open cell voltage is below the standard potential by an amount of delta E and counts 0.9 - 1,0 V. Under load, the real cell voltage decreases due to further losses with increasing current. The losses can be divided into three main categories. First, activation losses become effective, second, the ohmic resistance becomes stronger with increasing current density and third, gas transport losses occur at maximum power. **Figure 3** shows the typical current-voltage-behaviour of a fuel cell. [7]

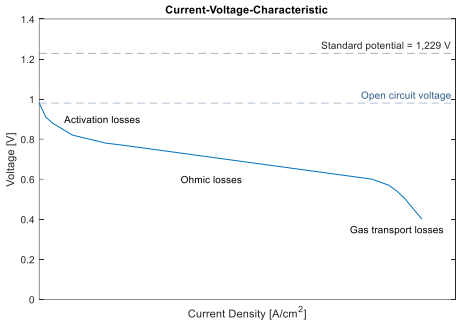


Figure 3: Fuel cell current-voltage-behaviour [7]

3. POWER REQUIREMENT

3.1. Application Scenario

In the project, a hydraulic driven construction machinery is used to test the hydraulic powerpack. For testing, a semi-stationary concrete pump is used. The main drive unit of the present concrete pump is a diesel combustion engine, which will be replaced by an electric motor. The electric motor drives the hydraulic pumps for the working cylinders. As a result, the main aim of this project is to drive the electric motor with a PEMFC. In order to calculate the required hydraulic power for the powerpack, a concrete pressure-power nomogram is utilized. The following boundary conditions are applied to the stationary concrete pump: a concrete flow of 8 m^3/h, a pumping distance of min. 60 m and a max. concrete viscosity of 450 Pa·s/m. This results a required, average net power of 12 kW for the further application, which must be provided by the electric motor. Consequently, the PEMFC must be able to provide the power constantly or partly in combination with a battery. Any power peaks that occur e. g. due to variable load requirements must be compensated by a battery storage system. For the design of an appropriate fuel cell, the characteristic of the entire fuel cell system must be considered.

3.2. Fuel Cell System Design

According to the previous chapter 2.1, the chemical reaction takes place in the reaction chambers of one single cell. Because only a low voltage is achieved with a single cell, several cells are connected into series to form a fuel cell stack. Due to the serial arrangement of the cells, the individual voltages add up and the power increases. By arranging the bipolar plates in the configuration of a stack, a compact design is achieved.

When talking about fuel cells in general, other important components are needed besides the stack. In literature, these components are often referred to balance of plant (BoP) components. Each fuel cell system can be subdivided into the following subsystems with their respective BoP components: Hydrogen path on anode side, air path on cathode side, cooling circuit and electronic control system. As major BoP power consumers the H_2 recirculation pump, the coolant pump and the air compressor must be mentioned. For the design of a fuel cell system according to (5), the self-consumption of the BoP components has to be considered. Consequently, the effective net power P_{FC} of a fuel cell system is defined as the difference between the power of the fuel cell stack P_{Stack} and the power of the BoP components P_{BoP}. [5]

$$P_{FC} = P_{Stack} - P_{BoP} \tag{5}$$

The following **Figure 4** compares the characteristic data curves of the fuel cell stack and the fuel cell system. The graph shows that the highest fuel cell system efficiency (~ 53 %) can be reached in partial load range or in the lower third of power. The power demand of the BoP components is lowest at the point of maximum system efficiency. In the range from low current densities up to the point of maximum system efficiency, the power demand of the BoP components for operating the fuel cell is high in relation to the effective power. This effect is mainly driven by the air compressor, the cooling- and recirculation pump, the heating system and the electronic control units to run the fuel cell system. The effective efficiency of the overall system has it's minimum at maximum power. At this point the efficiency of the stack is lowest and the power demand of the BoP components is highest. Therefore, for the design of a fuel cell system it is not only important to know the required net power, but also to take into account other consumers (BoP). For the complete functionality of the whole fuel cell system, a special challenge is posed to the cooling subsystem. Since in the worst case at about roughly 50 % efficiency, 50 % of the electric power is converted into thermal losses. For the hydraulic powerpack this effects a massive cooling subsystem, which must be integrated in the body of the machine. Furthermore, the hydraulic powerpack can't use natural airstream for cooling like on-highway vehicles. Instead, electric powered fans need to be placed in the machine concept.

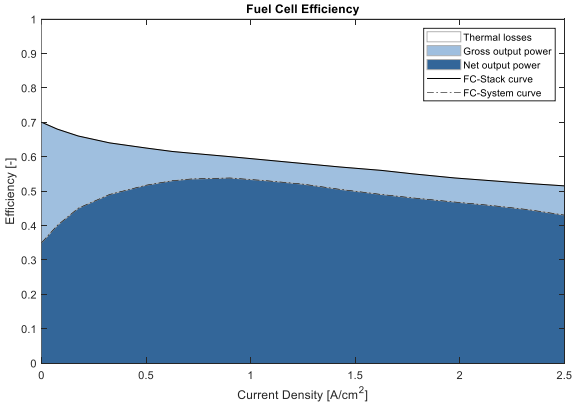


Figure 4: Fuel cell efficiency and losses [8]

Regarding the cooling subsystem it is possible to show the difference between a common internal combustion engine (ICE) and a fuel cell powered vehicle. **Table 3** is used to compare the distribution of the energy conversion in two cases by an example of an effective mechanical power of 10 kW. The internal combustion engine emits 25 % of the total energy via the cooling subsystem. The far greater percentage is dissipated to the environment through the exhaust gas enthalpy. In comparison, only 10 % of the total energy converted in the fuel cell is dissipated to the environment by the air flowing through it. [9]

$$P_{Heat} = \alpha A(\theta_{Fluid} - \theta_{Env}) \tag{6}$$

Due to the operating temperature of 60 - 80 °C of the fuel cell, the temperature gradient to the environment is significantly lower than with a conventional combustion engine, which operates in the range of 80 - 120 °C. According to (6) it means, for the same amount of heat power, a two to four times greater product of αA is required. Based on [10], this effects a two to four times greater fan power for cooling radiators, taking into account the heat transfer surface. This happens especially when the ambient temperature reaches high levels like 40 °C. The temperature range up to 40 °C is a common requirement in the development process for new machines. [5][9]

Table 3: Comparison of energy distribution in an ICE and a fuel cell [9]

	ICE [%]	Example ICE [kW]	Fuel cell [%]	Example FC [kW]
Mechanical/ electrical work	30	10	50	10
Cooling	25	8,3	40	8
Exhaust gas enthalpy	35	11,7	10	2
Friction	10	3,3	-	-
Total	100	33,3	100	20

3.3. Fuel Cell Operating Strategy

The sizing and the choice of the operating strategy of a fuel cell depends on various aspects, including the field of application, the required electrical power and the available resources. There are two main design topologies: efficiency optimized design and cost optimized design. With the help of **Figure 5**, it is shown that the system and stack efficiency of a fuel cell is dependent on the current density. Due to the characteristic curve, it is not possible to achieve the best possible system and stack efficiency at each operating point. In principle, the optimum system efficiency is achieved at low current densities; the lowest efficiency at maximum current density. The same applies to stack efficiency. As the current density increases, the stack efficiency decreases. In contrary, the fuel cell produces the maximum electrical power when the maximum current density is reached.

Efficiency optimized design focuses on achieving the maximum electrical efficiency and the lowest thermal losses of a fuel cell system. To reach this operating point, the fuel cell must be operated at partial load, as shown in the left diagram of **Figure 5**. This operating point can be achieved by either installing more cells or a larger active cell area in order to obtain the same fuel cell performance at a lower current density. The consequence of oversizing is that efficiency optimized fuel cells are more expensive regarding investment than cost optimized fuel cells. Furthermore, due to the larger number of cells and in parallel the larger BoP components, the packaging volume and the weight of the entire fuel cell module increases. In contrast, a cost optimized design focuses on minimizing the manufacturing and investment costs of the fuel cell. In this consideration the operating costs are excluded. Using a cost optimized strategy, it results that the operating point of the fuel cell system has a lower efficiency, the hydrogen consumption is less effectively and more thermal losses occur.

According to the required power, the fuel cell operates at the maximum possible power as shown in the right diagram of **Figure 5**. Contrary to the former strategy, the choice of the cost optimized strategy results in the fuel cell has no power reserves for higher performance requirements. Within the scope of the research project, the cost optimized operating strategy is applied to the hydraulic power unit.

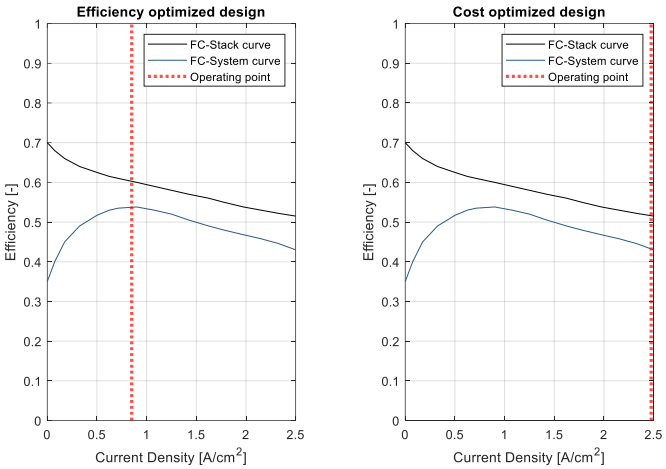


Figure 5: Fuel cell operating strategies

4. SYSTEM ARCHITECTURE

The hybridization of a fuel cell drive mainly differs in the arrangement of the electrical energy storage unit and the energy converter. This results in either a serial or a parallel arrangement of the main components. For the research project, the hydraulic powerpack is subjected to a parallel arrangement according to **Figure 6**. Due to the unavoidable energetic losses during charging and discharging of the battery, a serial arrangement of battery and fuel cell is not considered. The fuel cell system is connected to a DC/DC converter to be raised to the voltage level of the battery system. The chosen concept offers the possibility to increase the state of charge of the battery in case of a power surplus of the fuel cell. Thus, the battery system is charged with the help of the fuel cell when, for example, the powerpack does not have to supply hydraulic power. Within the scope of the research project, a fuel cell system with an effective power of 10 kW is selected. With an average power demand of 12 kW, the fuel cell is undersized, so that the battery system is used to compensate the required power difference and power peaks too.

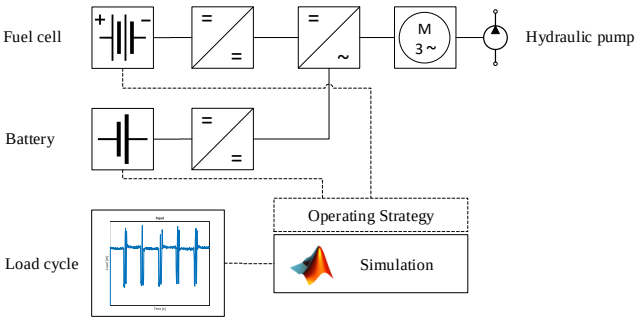


Figure 6: Block diagram of the hydraulic powerpack

The hydrogen infrastructure becomes greater and the hydrogen supply grid grows every year in Germany. However, the current hydrogen infrastructure situation does not allow the machine to be supplied with the sufficient fuel at all locations. This must be taken into account, especially on isolated construction sites. As a solution the consortium of the research project decided to use a battery system with a capacity of 35 kWh. The larger battery capacity of the hydraulic powerpack allows purely electric operation for around three hours. During this time no hydrogen is needed.

5. PROTOTYPE ASSEMBLY

In the prototype application, the powerpack is used to drive a stationary concrete pump. In the first step the goal agreed with the project consortium is to integrate the components of the powerpack into the existing housing of the concrete pump. Subsequently, the main goal is to assemble all the components as a stand-alone solution to provide hydraulic power to other hydraulically operated machines. The most challenging part of the drive modification is to integrate the required components into the existing housing.

The application scenario, stationary concrete pump, is currently driven by a diesel internal combustion engine (ICE) with an output power of 50 kW and an operating time of 8 hours. The following **Figure 7** shows the volume related comparison of switching the machine to hydrogen or battery technology following the same power output of 50 kW and the same operating time of 8 hours. For the comparison, the existing components of the original ICE machine were captured and their volume determined (factor 1). For the hydrogen and battery powered modification, the correspondingly required drive systems were indicated and then the volumes were determined too. Each colored block in Figure 7 corresponds to a converted volume of 150 l.

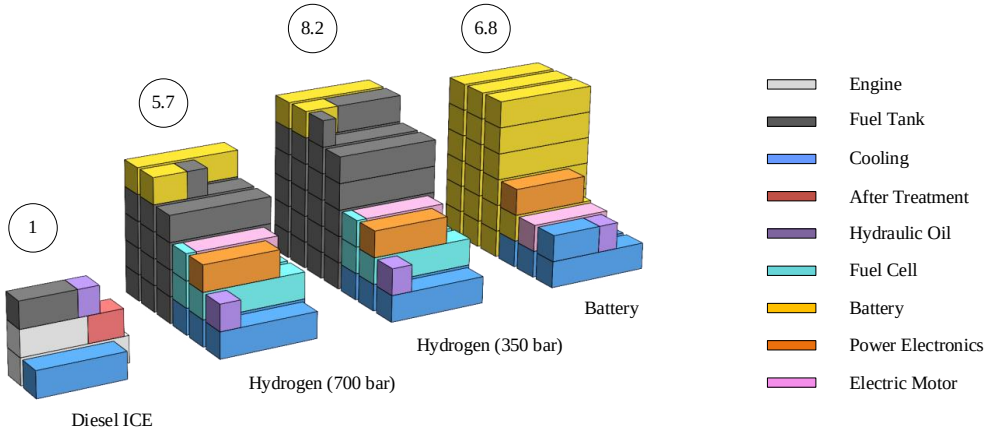


Figure 7: Comparison of installation space for ICE, H_2 350 bar, H_2 700 bar and battery technology

As shown in chapters 3 and 4, the fuel cell converts around 50 % of its power into thermal energy in the worst-case operating point. For this reason, a large cooling system must be installed. In addition, the battery must also be cooled, which is why more overall installation space must be available for the cooling system than for the ICE variant. In parallel to the CO_2 neutral working principle of a PEMFC, another advantage of using hydrogen is the 3 times greater gravimetric energy density compared to the fossil energy carrier diesel. Per kilogram of hydrogen, an energy capacity of 33.3 kWh can be achieved; in comparison, diesel achieves a gravimetric energy density of only 11 kWh/kg. The far greater challenge is to store this amount of energy volumetrically. Under normal pressure, hydrogen has a volumetric energy density of 0.003 kWh/l. To increase the volumetric

energy density, gaseous hydrogen is stored in tanks at 350 bar (0.8 kWh/l) or 700 bar (1.3 kWh/l). By comparison, conventional diesel fuel has a volumetric energy density of 10 kWh/l [11]. The storage of the energy content is the main reason why the volumetric factor increases to a factor of 8.2 compared to the ICE variant. The prototypical implementation of the system is in preparation so that initial operational experience can be gathered in the further course of the project.

6. CONCLUSION AND OUTLOOK

The present paper deals with a hydrogen driven powerpack that generates hydraulic power. Against the current developments in the construction industry, the novel machine offers a possibility to work CO_2 neutral when using green hydrogen and in parallel to produce no particles and nitrogen oxides. There are many different types of fuel cells on the market today. For mobile applications the PEMFC is used in most cases. This fuel cell type is therefore used in the research project "Hydrogen 2 Hydraulics" and its functionality is explained in this report. For the selection of a fuel cell system, the influence of the auxiliary consumers (BoP components) at different load points is shown clearly. Based on the presented operating strategies, the operating range of a fuel cell can be adjusted and either an efficiency or a cost optimized variant can be chosen. For a fuel cell with a power of 10 kW, which is used in the research project, the approach of a cost optimized variant is adopted. The volumetric energy density of hydrogen shows that the biggest challenge is the integration of the fuel tanks if a constant operating time is to be achieved as for a comparable, existing diesel application. For the application scenario presented in the paper, this means an increase in the installation space by a factor of 8 under the same boundary conditions. Since the installation space in construction machinery is not infinite, the tank volume must be reduced at the expense of operating time. It is important to note that there won't be an "all fits one solution" for all machines; the use of hydrogen must be re-evaluated for each application and one must be prepared to adapt the requirements taking into account the possibilities and limitations.

NOMENCLATURE

α	Heat transfer coefficient	$W/(m^2 \cdot K)$
Θ	Temperature	K
A	Area	m^2
E^0	Standard potential	V
F	Faraday constant	As/mol
I	Current	A
L	Length	m
p	Pressure	Pa
P	Power	W
$\dot{Q}$	Volumetric flow rate	m^3/h
$\Delta_R G_M^0$	Free enthalpy at standard pressure	J
V	Volume	m^3
W	Energy	Wh
z	Electric charge number	-

BoP	Balance of Plant
CH_4	Methane
CH_3OH	Methanol
CO_2	Carbon Dioxide
DC	Direct Current
EU	European Union

FC	Fuel Cell
GHG	Greenhouse Gas
GDL	Gas Diffusion Layer
H_2	Hydrogen
H_2O	Water
ICE	Internal Combustion Engine
Max..	Maximum
Min.	Minimum
NO_x	Nitrogen Oxides
NRMM	Non-Road Mobile Machinery
PEMFC	Proton Exchange Membrane Fuel Cell
PM	Particular Mass
TREMOD-MM	Transport Emission Model for Mobile Machinery

REFERENCES

[1] Die Bundesregierung (2023) Klimaschutzgesetz und Klimaschutzprogramm. https://www.bundesregierung.de/breg-de/aktuelles/klimaschutzgesetz-2197410. Accsessed 2 Oct 2023

[2] Eder P (2020) LECTURA Flash Umfrage Motoren. https://press.lectura.de/de/report/lectura-flash-umfrage-motoren/51. Accsessed 10 Oct 2023

[3] Heidt C, Helms H, Kämper C, Kräck J (2020) Aktualisierung der Modelle TREMOD/TREMOD-MM für die Emissionsberichterstattung 2020 (Berichtsperiode 1990-2018): 117/2020. Dessau-Roßlau

[4] Geitmann S, Augsten E (2022) Wasserstoff und Brennstoffzellen, 5th edn.. Hydrogeit Verlag, Oberkrämer

[5] Klell M, Eichlseder H, Trattner A (2018) Wasserstoff in der Fahrzeugtechnik, 4th edn.. Springer Vieweg, Wiesbaden

[6] Li C, Liu Y, Xu B, Ma Z (2019) Finite Time Thermodynamic Optimization of an Irreversible Proton Exchange Membrane Fuel Cell for Vehicle Use. Processes. 7(7):419. https://doi.org/10.3390/pr7070419

[7] Yurukcu M, Yurtsever F M, Demirel S, Saldaña J, Sari M M (2021) Advanced materials for next-generation fuel cells. In: Sustainable Materials for Transitional and Alternative Energy, Gulf Professional Publishing

[8] Höflinger J, Hofmann P, Müller H et al (2017) FCREEV A Fuel Cell Range Extended Electric Vehicle. MTZ Worldwide 78: 18–23 (2017). https://doi.org/10.1007/s38313-017-0033-x

[9] Schäfer S (2012) Modellbasierte Steuerung des Kühlkreislaufes einer Brennstoffzelle mit automatisiertem Test der Software. Dissertation, TU Darmstadt

[10] Masten D A, Bosco A D (2010) System design for vehicle application (GM/Opel). In: Handbook of Fuel Cells, John Wiley & Sons Ltd

[11] Lehmann J, Luschtinetz T (2014) Wasserstoff und Brennstoffzellen. Springer-Verlag Berlin Heidelberg

HIGH-PRESSURE SHUT-OFF VALVE SUITABLE FOR HYDROGEN APPLICATIONS

Peter Tappe*, Peter Baumgartl, Johann Weiß

[1]*Magnet-Schultz GmbH & Co KG, Allgäuer Straße 30, 87700 Memmingen, Germany*

* Corresponding author: Phone: +49 8331 104-392; E-mail address: peter.tappe@magnet-schultz.de

ABSTRACT

It is well known that the direct or indirect use of hydrogen in mobile applications plays a decisive role in the decarbonisation of this sector. Regardless of the use by means of fuel cell or modified internal combustion engine, various valve functions are required along the functional chain.
In mobile applications, the hydrogen is stored in high-pressure tanks with up to 1050 bar. For dispensed extraction, the shut-off solenoid valve size 32 mm is used as part of an on-tank valve.
The article describes the construction design and explains the design of the sub-functions, in particular the pilot-controlled valve function.

Keywords: hydrogen, shut-off valve

1. INTRODUCTION

For all necessary valve functions inside of a mobile fuel cell, Magnet-Schultz developed an entire valve platform. In those applications, the hydrogen is stored in high-pressure tanks with up to 1050 bar. For dispensed extraction, the shut-off solenoid valve (size 32 mm) is used in an on-tank valve. For the medium pressure range of up to 25 bar, a shut-off valve is also required (SOV). The modular system for this pressure range is completed with a flow control valve. Purge and drain valves are required at the outlet of the fuel cell to manage the emerging water.

As part of the Magnet-Schultz valve platform, this article presents the High-Pressure Shut-Off Valve with the pilot control inside the valve.

2. MEDIA CYCLE OF A FUEL CELL

Fuel cells are known to convert the chemical oxidation reaction of oxygen and hydrogen directly into electrical energy with a high degree of efficiency. In addition to electrical energy, the reaction products are heat and water. Consequently, the technical operation of such fuel cells requires the supply and discharge of the media hydrogen and water.

The media flows are shown schematically in **Figure 1** below. The necessary valves are also shown in this figure. This also includes the so-called on-tank valve, which in turn is an assembly with the High Pressure Shut Off Valve installed in it.

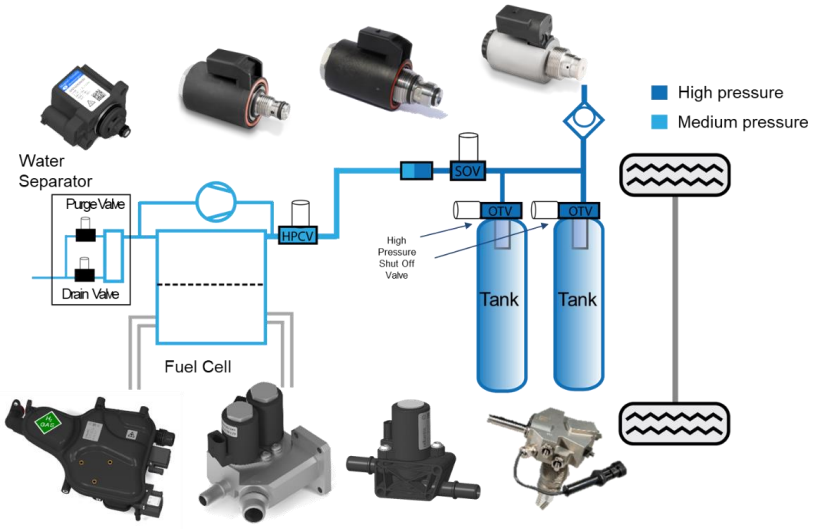


Figure 1: Media flows in a fuel cell

Because gaseous hydrogen has a low volume-specific energy density, it is provided at the input of fuel cells by high-pressure tanks. To increase efficiency, storage pressures of up to 1,000 bar are aimed for. The hydrogen must be supplied at this pressure level through the on-tank valves. A subsequent pressure-reducing valve depressurises to an average pressure level of approx. 25 bar. In addition, a shut-off valve separates the high and medium pressure cycle. A flow control valve is required at the inlet of the fuel cell to meter the hydrogen.

The water produced at the outlet of the fuel cell is deposited using a separator and discharged into the environment through a drain valve. The nitrogen produced as a by-product is discharged via a purge valve.

Magnet-Schultz offers a suitable solution for all required valve functions.

The focus of this article is on the On-Tank Valve, the basic structure of which is shown in **Figure 2** below.

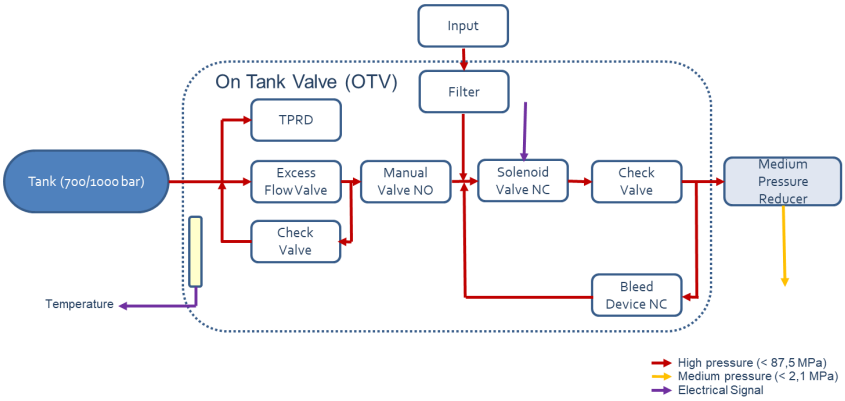


Figure 2: Structure of the On-Tank-Valve

The 'On-Tank-Valve' assembly consists of various components including the High Pressure Shut Off Valve, which is described below.

3. HIGH-PRESSURE SHUT-OFF VALVE

The High-Pressure Shut-Off Valve (**Figure 3**) has the following technical data:

- Rated working pressure (NWP): 0 ... 700 bar
- Maximum pressure (MAWP): 1050 bar
- Nominal width: 2.7 mm
- Type of voltage: DC, PWM
- Protection class: IP6K9K
- Testing according to EC79



Figure 3: High Pressure Shut-Off Valve

The patented design of the valve takes into account the increasing requirements from the application in mobile fuel cells and also shows functional advantages compared to the general state of the art:

- Increasing the nominal operating pressure
- Switching function independent of the counter pressure on the downstream circuit
- Constant switching times
- Low leakage

3.1. Functional description

Due to the extremely high working pressure, the state of the art usually works with a pilot control, as shown in **Figure 4**.

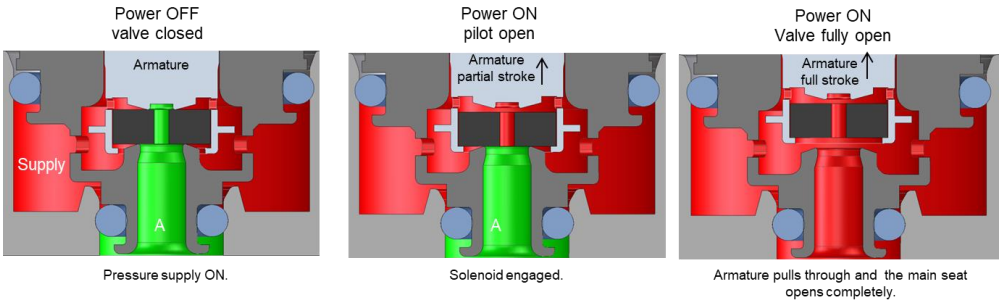


Figure 4: Opening process with back pressure support

When the valve is closed, both the armature and the sealing element in the valve are pressed onto the seat by the working pressure. If the solenoid valve is now energised, the armature opens the pilot control. This starts a pressure equalisation between the inlet side and the outlet. Only when a sufficiently high-pressure level is reached at the outlet, the valve element can fully open the main seat due to the magnetic force that is still applied.

The description of the process makes it clear that the opening behaviour is clearly dependent on the filling and thus the pressure build-up at the outlet. This means that the switching process depends on the design conditions in the overall system. There is also a dependency on the filling level of the storage tank.

The opening process of the technically optimised valve, which is also supported by a pilot control, is shown in **Figure 5**.

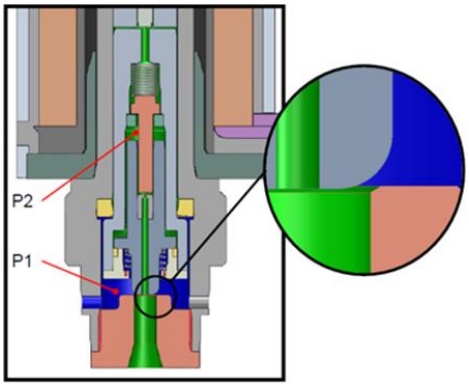


Figure 5: Opening process with optimised valve

In the initial state, the valve seat is also pressed onto the seat by the working pressure. The resulting sealing force results from the interaction of the working pressure at the inlet, the pressure at the outlet and the spring forces at the armature, the main plunger and the pilot valve.

When the solenoid valve is energised, the pressure-equalising armature moves against the spring force up to the mechanical contact with the pilot control. As a result, the pilot control nozzle is now opened as soon as the solenoid force overcomes the pressure difference at the pilot nozzle as well as the resulting spring forces.

As the process continues, the armature moves up to the mechanical contact with the main plunger and the seal installed therein. Due to the opening of the pilot plunger, pressure equalisation can now take place between the outlet and the armature chamber. This creates an opening pressure force so that, together with the resulting armature force, the main plunger opens slightly.

The pressure in the armature chamber is then equalised again so that the main plunger can open completely using spring force. The valve is now fully open.

3.2. Static interpretation

The description of the opening process illustrates the complex functional sequence in which the resulting spring forces and pressure build-up processes act and influence each other due to the pilot and main controls.

Proven FEM simulation programmes can therefore be used for the basic interpretation. Since the construction provides for a pressure-balanced armature, the design of the magnetic circuit is primarily focussed on the ability to overcome the resulting spring forces. The corresponding results can be seen in **Figure 6**.

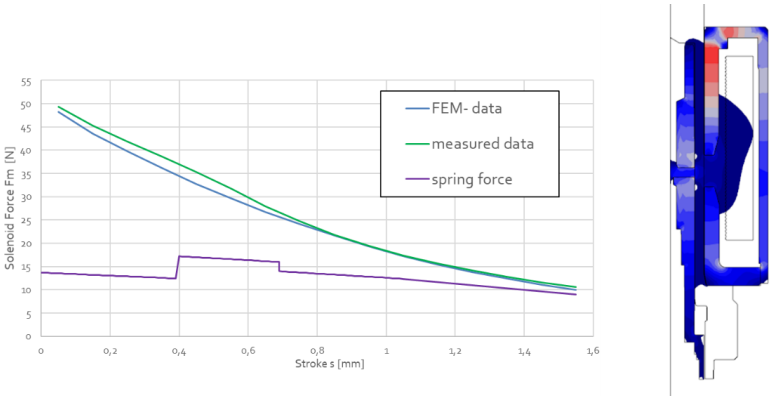


Figure 6: Force-stroke behaviour (solenoid)

The mass flow of the main seat was calculated using CFD simulation (**Figure 7**). The simulation result checks which variation arises starting from the selected stroke point (0 mm). The current interpretation is almost optimal, taking into account the tolerances that occur

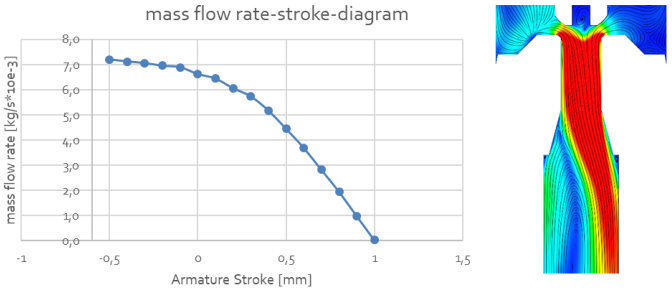


Figure 7: Mass flow over main valve seat

3.3. Transient layout

Due to the complexity of the functional processes, a dynamic simulation is also suitable in addition to the static calculations. At the same time, the complexity requires a systematic approach with the final goal of a multi-domain simulation. In preparation for the complete transient simulation, a 2D model of the construction was first derived. Due to the rotationally symmetrical design, such a derivation is possible (see **Figure 8**).

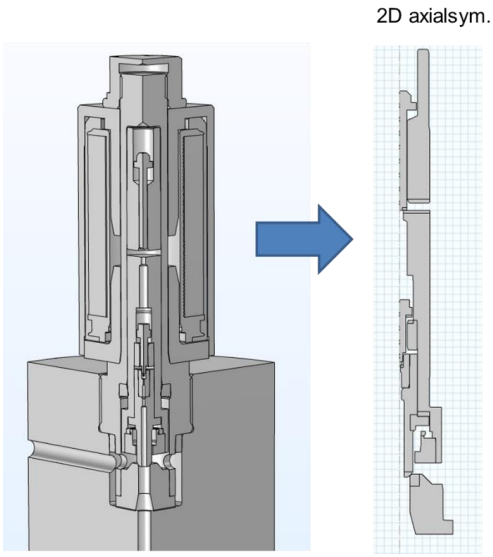


Figure 8: Derivation of the 2D model for transient magnetic force simulation

Thus, in a first step, the magnetic circuit could be calculated transiently in isolation (see **Figure 9**).

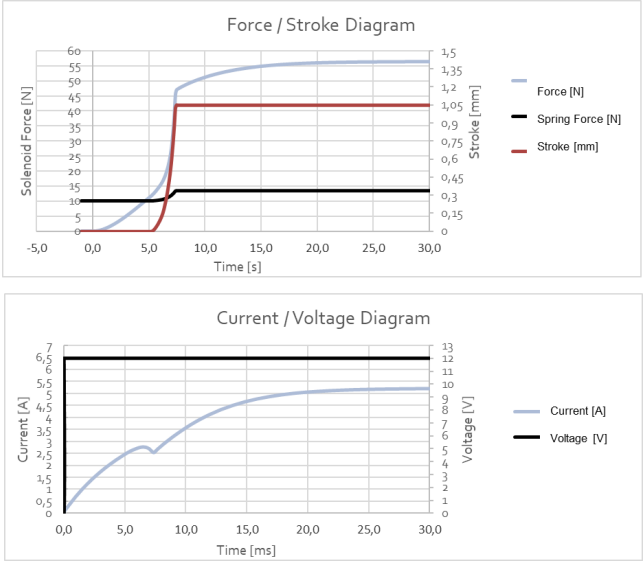


Figure 9: Transient magnetic behaviour

The complete 3D model is required for the CFD calculation. In order to gain experience with the transient flow simulation, water was used as an alternative medium instead of the compressible gas (see **Figure 10**).

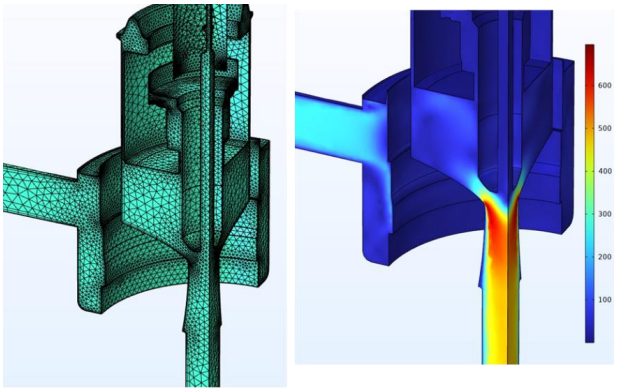


Figure 10: CFD calculation alternatively with water

The multi-domain calculation requires specific interventions in the FEM model. In the case of mechanical contact points, only partial areas of the functionally relevant individual parts should be modelled using structural mechanics in order to reduce the calculation time significantly. In the case of relevant air gaps in the magnetic circuit, the mesh structure must be suitably adapted manually (see **Figure 11**).

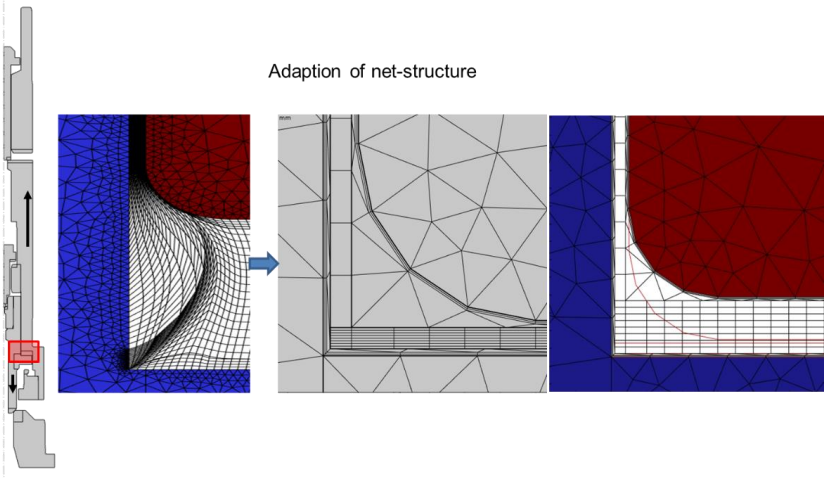


Figure 11: Adaptation of the network structure in magnetically relevant air gaps

The preparatory work described herein for the full transient simulation of solenoid valves has not yet been completed and represents a current intermediate step.

4. CONCLUSION AND OUTLOOK

The patented High Pressure Shut-Off Valve developed by Magnet-Schultz is able to switch a usable mass flow for mobile fuel cells with a very limited solenoid size. The special feature is the design of the pilot control, which ensures reliable valve opening with a constant switching time regardless of

the back pressure.

For support and further optimisation, a complete transient calculation of the valve is aimed for. The preparatory steps for the static and dynamic simulation of individual functional elements described in this article were carried out for this purpose.

HOLISTIC APPROACH TO ELECTRO-HYDRAULIC DRIVE SOLUTIONS FOR HYDROGEN PISTON COMPRESSORS

peer reviewed

Nicolas Englert*, Henning Noack, Johannes Schwacke, Andreas Günder

Bosch Rexroth AG, Zum Eisengießer 1, 97816 Lohr a. Main, Germany

* Corresponding author: Tel.: +49 9352 18-2177; E-mail address: nicolas.englert@boschrexroth.de

ABSTRACT

Compressor drives with hydraulic drive technology are showing good market growth due to the trend towards hydrogen. The application seems very simple at first due to its repetitive cycle, but a closer look reveals some technical challenges. However, with a deep understanding of the end user's point of view as well as the physical principles, it is possible to design a hydraulic system that meets the target cost but also achieves robustness, high flow rate and good energy efficiency. However, no general recommendation can be given for a specific system structure. Though in general, it can be seen that variable displacement pumps should be the favoured solution for motion reversal in a wide target range.

Keywords: Industrial hydraulics, Hydrogen, Compressor drive, Efficiency

1. INTRODUCTION

To meet climate targets, many sectors must switch to renewable energies. The use of hydrogen as an energy carrier ranks among the key technologies for sectors where electrification has failed so far or for long-term energy storage. Due to its low volumetric energy density, hydrogen for mobile applications (cars, trucks, construction machinery, railroads, shipping, aviation) must be compressed so that storage units can be dimensioned as small as possible. During the refuelling process of these vehicles, a powerful compressor is required on site because the pressure in the supply line is very low or the pressure in the supply storage tank decreases with increasing extraction.

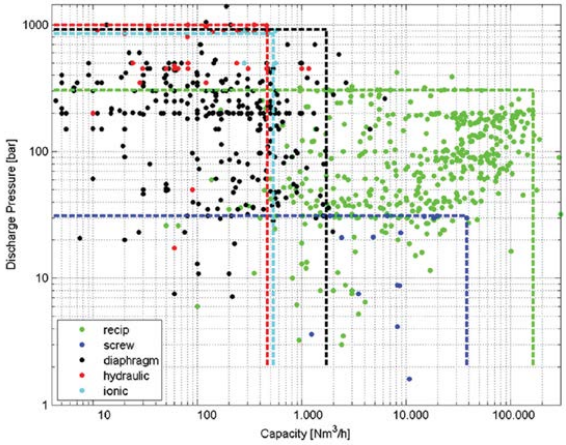


Figure 1: Compressor map with current industrial references (dots) and performance ranges (lines) for pure hydrogen. [1]

In addition to this end use, many of the pressurized components need to be validated during development and often 100% tested at the end of production, which is why industry also has a need for compressors.

Compressors for these applications can be designed using several functional principles. Due to high end pressure of up to 1000bar, large pressure differences of up to 900bar and small to medium flow rates, diaphragm and piston compressors are particularly suitable for these applications (see figure 1). Due to further requirements such as the ability to start-stop operation and low space requirements, hydraulically driven piston compressors have gained great relevance.

While hydraulic piston compressors generally exist already for a long time, the design of the hydraulic system is often simple and not contemporary, because this application didn't get a lot of attention. That's why the aim of this paper is to explain the special requirements and the basic approach for optimizing hydraulic systems for this use case.

2. SYSTEM STRUCTURE

First, the typical structure of a reciprocating compressor is explained using Figure 2. It consists of three main components: the drive, the compressor, and the system control. The task of the drive is to convert the electrical energy into mechanical energy while generating as little power loss as possible, which will be dissipated in the form of heat. In addition, it is also responsible for controlling the mass flow since the compressor usually acts only passively. In contrast to the drive unit, the heat generated at the compressor is mostly due to the physical laws of compression and not just to losses due to inefficiency. It is also often unavoidable that part of the heat output of the compressor flows back into the drive unit.

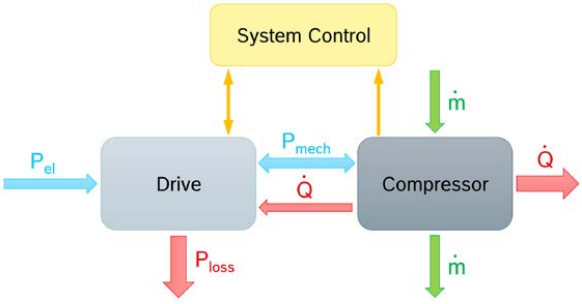


Figure 2: Basic structure of reciprocating compressors.

In addition to this very simple design, several compressor stages are usually used, which can be connected either in parallel or in series. Figure 3 and figure 4 show the typical design of a hydraulic double rod cylinder connected to single gas pistons on both sides. But the same applies to reciprocating compressors with a crank shaft.

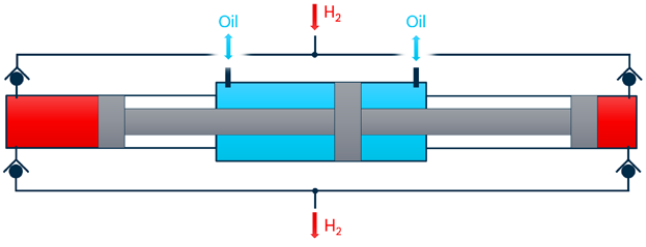


Figure 3: Typical design of a hydraulic piston compressor with parallel gas pistons.

In the case of a parallel connection, the individual pistons are moved out of phase, like in a combustion engine, so that the force of the inflowing gas can be used for compression on the other piston during the return stroke, because hydrogen already has an increased pressure in the inlet in the applications mentioned. In this way, the mass flow can be increased with reduced electrical input power but increased energy requirements. Furthermore, the flow rate is stabilized.

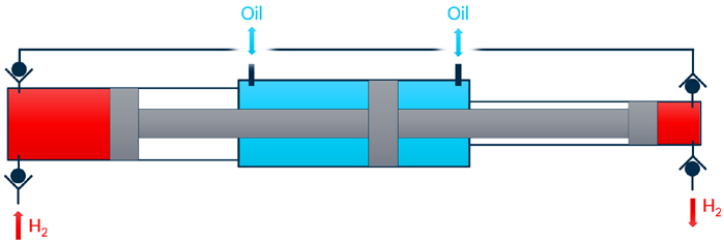


Figure 4: Typical design of a hydraulic piston compressor with serial gas pistons.

With the series connection, the sometimes very large pressure ratio is broken down into smaller steps. This primarily improves efficiency, since the dead volume is not as significant, and the compression curve approaches the isotherm through intermediate cooling. In addition, if the pressure ratio is too high, the heating of the gas can be so severe that the service life of the compressor is massively reduced.

Although multistage compressors are always required in practice for the reasons mentioned above, they are not considered below because the findings are transferable, and the complexity can be reduced.

3. CUSTOMER REQUIREMENTS

To be able to design the system in the best possible way, the customer's point of view and its requirements must first be understood. At first there are global requirements, which must always be met, such as the purity of the hydrogen or compliance with safety requirements. But in addition to them there are also various specific requirements and performance data that differ from customer to customer. The following explanation is not universal but should be a valid simplification of the complex requirements. The target is to list only requirements that can be provided with hard numbers and no soft factors.

Table 1: Specific boundary conditions of the application.

Boundary condition	Parameter	Unit
Electrical connection rating	Electric power	kW
Footprint	Area	m²
Country specific certification	-	-
Environmental condition	Temperature, dirt, corrosion protection, vibration, noise level	°C, dB, …
Max. initial cost	Price	€ / $

First, there are some constraints or filters that are imposed by the application. For each of the items listed in Table 1, there is a limit that must not be exceeded. However, it is of limited added value to the end customer if the system is better than the limit value. For example, the electrical power supply may be limited to 80kW by local restrictions. A compressor with an average power requirement of 120kW cannot be operated, but it is only of secondary importance how far one is below the limit

value. The maximum initial costs in Table 1 are only intended to limit the total budget, but the economic decision should not yet be made here.

After the available systems have been pre-selected with the help of the boundary conditions, the relevant ones can be compared based on their KPIs (= Key Performance Indicator). The most important KPI is the flow rate, as it determines how quickly refueling processes can take place or how long pressure tests will last. However, the value is not static since it depends strongly on the inlet pressure and the inlet temperature of the hydrogen (see formula (5)). Therefore, standardized conditions should be used here. In the best case, different extreme points are selected that show the behavior in the relevant operating range.

Table 2: Key Performance Indicator of a compressor system.

KPI	Parameter	Unit	Quantification
Flow rate	Mass of H2 per hour	kg / h	Higher is better
Initial cost	Cost per flow rate	€ / (kg / h)	Lower is better
Energy efficiency	Energy per mass of H2	(kW × h) / kg	Lower is better
Service Life	Time interval	h	Higher is better

The initial costs are naturally very important. Since different systems rarely have the identical flow rate, the costs should therefore only be compared specifically, as acquisition costs per flow rate.

At present, energy efficiency is often not as highly weighted as the other three aspects, but it will become increasingly important in the future. Like the flow rate, the energy requirement per kilogram of compressed hydrogen can only be determined as a function of a specified operating point.

Finally, the time interval between inspections is very important, as this can seriously shift the cost structure in retrospect. Especially at refueling stations, where consumers depend on refueling their vehicles, because no other refueling station can be reached with the remaining range, a maximum availability must be realized.

4. PHYSICAL PRICIPLES OF THE COMPRESSOR

For a better understanding of the optimization approaches, the basic physical process of polytropic compression and the positive and negative influencing factors will be explained first. For this purpose, the ideal process of isothermal compression without dead volume with ideal gas is explained first, and then the individual harmful influences and their effects are shown.

In practice, hydrogen behaves only like an ideal gas near atmospheric conditions (1 bar, 20°C). At higher temperatures and pressures, the repulsive forces between the molecules become increasingly significant. This means that the actual density of hydrogen is lower than calculated with the ideal gas equation. As a result, the pressure builds up earlier, similar to the difference between isothermal and isentropic compression, increasing the work required for compression and thus also increasing the heat generated. Furthermore, the mass flow rate at high inlet pressure is lower because the compressor chamber is filled with less mass. In turn, the negative effects of dead volume are slightly reduced since there is less gas mass present in the same dead volume. All these effects change the quantitative design of the compressors, but not the qualitative behaviour. Therefore, it is acceptable to use the simplified approach with ideal gas for optimization, if exact numerical values for the flow rate are not required.

The pressure build-up can be described with formula (1), where the polytropic factor n depends on the type of compression. The isothermal compression ($n = 1$) assumes constant temperature, which can only be achieved if the cooling is very intense, and the compression speed is infinitely low. The

opposite would be the isentropic or adiabatic process ($n = \kappa$, for hydrogen $\kappa = 1{,}4$), where no cooling or infinitely fast compression is assumed.

$$p(V) = p_0 \times \left(\frac{V}{V_0}\right)^n \tag{1}$$

Real compressors do not follow either process, so a polytropic change of state must be used here. The pressure curve stays between the previous two mentioned but tends more towards isentropic. Since heat transfer is faster at larger temperature differences, the polytropic exponent changes continuously during the process. For simplification, an empirically determined value between 1 and κ can be chosen.

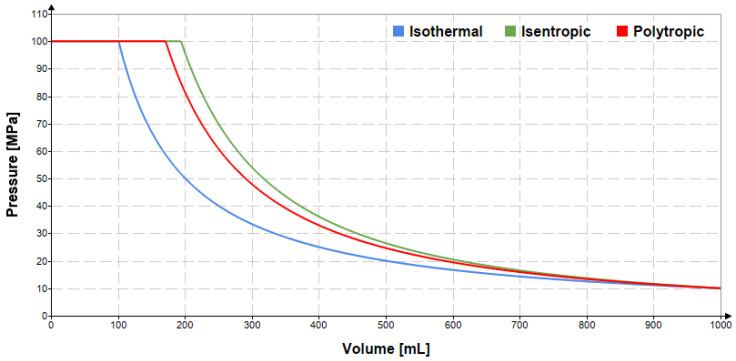


Figure 5: Pressure curve of the compression process from 10MPa to 100MPa.

In continuous operation, a cyclic process takes place, which is shown in Figure 6. The work done for compression can be calculated using formula (2) or figuratively from the area enclosed by the cycle. The waste heat, which is transferred to the cooling system, corresponds to the compression work reduced by the enthalpy change of the hydrogen (see formula (3)). The enthalpy of ideal gas only increases due to temperature, for real gas it would be temperature and pressure depended. In total the whole temperature increase must be cooled since the hydrogen can't be used while it is hot in the most applications. But for the heat dissipated at the compressor itself, it is only relevant how hot the gas is during discharge. Since the pressure increase is given, the formula thus establishes the relationship between the initial temperature of the hydrogen and the cooling capacity needed to lower the discharge temperature below the isentropic process to the polytropic process. Therefore, the discharge temperature must also be calculated with the assumed polytropic factor and is included in the formula (4).

If the maximum temperature and required cooling capacity are to be calculated for dimensioning of the system it is strongly advised not use the ideal gas formulas shown here, but rather look up the real gas values, e.g. in REFPROP. [2]

$$W_{comp} = -\oint p\, dV = \int_3^1 p_{out}(V)\, dV - \int_3^1 p_{in}(V)\, dV \tag{2}$$

$$Q_{loss} = \Delta W - \Delta H = W_{comp} - m_0 \times c_{p.H2} \times (T_{out} - T_0) \tag{3}$$

$$T_{out} = T_0 \times \left(\frac{p_{out}}{p_{in}}\right)^{\frac{n-1}{n}} \tag{4}$$

The ideal compressor without dead volume follows the dotted line in Figure 6. In real machines,

however, dead volume is always present, which distorts the cycle according to the solid line. At the end of the stroke **3**, there is still volume present with the output pressure. During the return stroke, this volume must therefore first be expanded to the inlet pressure **3-4** before further gas can flow in via the inlet valve **4-1**. The greater the pressure ratio, the more noticeable the dead volume. That's why an extreme case with pressure intensification factor 10 has been chosen as an example.

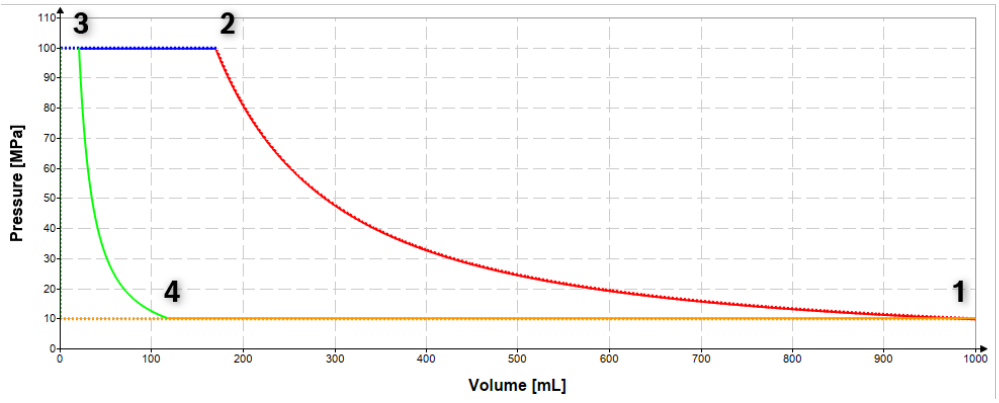


Figure 6: Cycle for continuous operation.

The flow rate per stroke is determined only by how much gas flows into the compressor during the return movement after the expansion of the dead volume **4-1** (formula (5)). For maximum efficiency, it is therefore also important to fill the compressor chamber with gas as cool as possible since the higher density leads to more mass flow for the same stroke volume.

$$m_0 = \rho_{H2} \times V_0 = \frac{p_0}{R_{H2} \times T_0} \times V_0 \tag{5}$$

Furthermore, it is also evident that the dead volume reduces the flow rate, since the volume available for the inflowing gas no longer corresponds to the displacement volume (ΔV between **3-1**), but only a reduced effective volume (ΔV between **4-1**) remains, which can be calculated with formula (6). It is important to understand that the polytropic coefficient in this equation refers to decompression and will therefore have a different value than during compression. For simplicity, identical values are used in the diagrams.

$$V_0 = V_{stroke} - V_{dead} \times \left[\left(\frac{p_{out}}{p_{in}} \right)^{\frac{1}{n}} - 1 \right] \tag{6}$$

The energy efficiency is not affected by this, if the work of the expanding gas can be used, since the area of the cycle is reduced in proportion to the flow rate.

The continuous mass flow will then be determined as a function of the stroke frequency f or the mean piston speed $\bar{v}$ according to formula (7).

$$\dot{m} = m_0 \times f = m_0 \times \frac{\bar{v}}{2 \times l_{stroke}} \tag{7}$$

The consideration of further losses such as the pressure losses at the inlet and outlet valves as well as leakage via the piston seal plays no role for the optimization of the drive and are therefore not taken into account.

As a last step, the power curve of the process should be calculated to design the required drive unit. For this purpose, the pressure curve (1) must be multiplied by the area of the gas piston and the velocity profile of the drive unit (8). Since the pressure curve is dependent on the volume, but the formula for power should only be dependent on time, (9) is used to establish the time dependence of the volume.

$$P_{comp}(t) = p(V(t)) \times A_{piston} \times \dot{x}_{drive}(t) \tag{8}$$

$$V(t) = V_0 - A_{piston} \times x_{drive}(t) \tag{9}$$

Figure 7 shows two different motion profiles as examples.

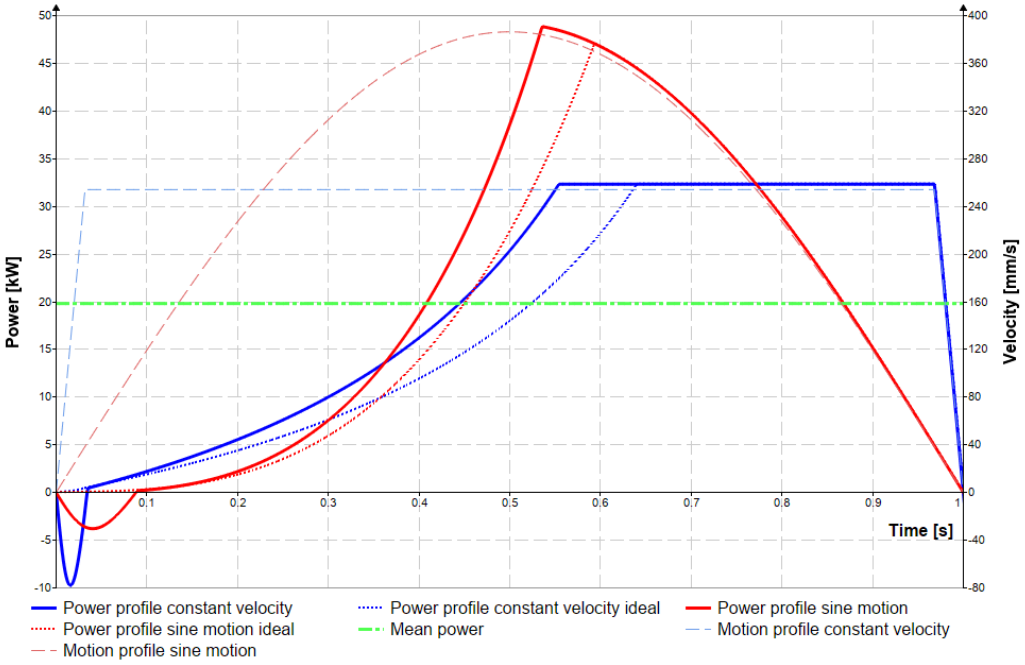


Figure 7: Power curves of different motion profiles.

In red a sine motion is visualized whereas in blue a typical motion of a hydraulic cylinder with short acceleration phases and constant velocity over a major part of the stroke is shown. To illustrate the losses in the compressor, the dotted line shows the ideal compressor with isothermal change of state and without dead volume. Two significant deviations can be seen. On the one side there is an area with negative power at the beginning of the stroke during the decompression of the dead volume, on the other side the maximum power is reached earlier, because the target pressure is reached faster due to the heating of the gas. Likewise, peak power of both motion profiles is very different, with both being significantly larger than the average power. This is due to the fact that the counterforce only builds up in the course of the stroke, and the speed is controlled independently of this.

For hydraulically driven systems, the trapezoidal profile is most important, as it requires a lower peak power and thus a smaller power unit. Furthermore, the maximum speed of the hydraulic cylinder is lower for the same mass flow, which has an advantageous effect on flow losses and seal life. Section 5 therefore assumes a trapezoidal profile only. For the acceleration and deceleration high values can be achieved, so the smoothening effect on the profile is quite small. Since the moving mass is low, no damage by pressure spikes should occur.

5. OPTIMIZATION APPROACHES

In this section, the possibilities of drive technology to optimize the customer's KPIs will be shown. It is assumed that the basic design has already been carried out according to the end application and that the boundary conditions from Table 1 are therefore already fulfilled.

5.1. Flow Rate

To maximize the flow rate, the influencing factors of formulas (5) to (7) should be analysed. The drive unit has the possibility to achieve an increase of the mass flow rate via a high average piston speed as well as a precise switching near the end position, in order to minimize the dead volume.

The high average piston speed can be achieved with a short stroke and high switching frequency as well as with a long stroke and low frequency. However, since the negative influence of the dead volume is proportional to the stroke volume (see formula (6)), a larger stroke volume, with identical amount of dead volume, can reduce the negative effects, which is why the longest possible stroke is preferable from this point of view.

To achieve precise end position switching, a highly dynamic and high-resolution measuring system is required in addition to a suitable control concept. This is particularly challenging for the design of the cylinder since double rod cylinders are state of the art and the common displacement transducers cannot be applied there. This is further complicated by the explosion protection requirements, which lead to severe restrictions in the design of the sensor electronics.

The most common solution is therefore the classic inductive limit switch, which should ideally have an adjustment option so that the manufacturing tolerances can be compensated. Encoders on the outside of the cylinder tube are more difficult to implement and are therefore not widespread, but they can be implemented in principle. However, the weaknesses of the displacement transducer can be compensated by sophisticated software since the work cycle is highly repetitive and predictable in this application.

5.2. Cost

As already shown in section 3, the costs should be considered as a specific parameter in relation to the flow rate. Therefore, the primary goal should be to maximize the flow rate with as few additional costs as possible. Taking the hydraulic cylinder as an example, one should therefore try to run it as fast as possible with a high working pressure, since this results in a cylinder that is as small as possible and therefore as inexpensive as possible for the same delivery rate.

Saving costs by choosing components with low quality or a system structure with poor energy efficiency is not recommended, as these result in worse KPIs for energy efficiency and low service life, which are also important parameters of the overall product.

5.3. Energy efficiency

In addition to the powerful cooler in the compressor, which is not influenced by the drive solution, it is primarily the main control element of the movement inversion and the power control that are decisive for the energy efficiency of the drive. Of course, the pressure losses in the piping should be kept as low as possible. The friction of the cylinder seals can usually be neglected.

Due to the increased power density of compressor drives compared to average hydraulic solutions and long continuous operating periods, there is a high cooling demand, especially at low energy efficiency. This is further complicated by the fact that, in the case of strong heat input due to power dissipation, it is difficult to transport the heat evenly to the cooling unit, which can result in highly

unevenly distributed temperatures in the system. As a result, large additional costs for cooling are added and even short disturbance can damage the system due to overheating. In addition to the pure energy costs, which in some cases are still of little importance, the other advantages of efficient hydraulics should therefore be considered.

As mentioned in previous sections, compressors are mostly used in multi-stage systems. When selecting the main control element, therefore, consideration should be given to how well the flow rate and power can be controlled, so that multi-stage operation, possibly even a modular system, is possible.

Movement inversion by valve

Figure 8 shows a simplified schematic with a valve-controlled compressor drive. The motor and pump operate at constant speed and thus constant volume flow since the cylinder moves at constant speed except for the change of direction. When the end position is reached, the valve is switched over to reverse the movement. Due to the constant rotation speed, the pressure of the pump depends on the load on the cylinder. If the flow rate of hydrogen is to be reduced, the volume flow and thus the cylinder speed must be reduced, for example via the motor speed.

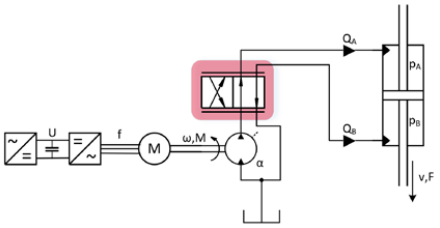


Figure 8: Simplified schematic of a valve-controlled system.

The advantages of this system structure include low complexity and low cost. Furthermore, the switching of the end position is very precise and suitable for high frequencies, which means that the flow rate of the compressor can be maximized. To evaluate the efficiency of this drive concept, the power curves of the electric motor and the hydraulic cylinder are simulated with SimulationX in Figure 9. In doing so, many assumptions must be made for the system sizing, which are not listed in detail, as they only affect the scale of the result but not the behaviour itself. An idealized behaviour of a compressor, as explained in chapter 4, was assumed as load.



Figure 9: Motor and cylinder power in the valve-controlled system.

It can be seen that the power curves are different from each other. First, an offset during the pressure build-up can be recognized. This is primarily caused by the pressure loss at the valve and general piping losses as well as the efficiency of the motor and pump. Since the cylinder runs at a constant

speed and the pressure losses depend on the flow rate, the offset is constant and is therefore more significant at low working pressures. For optimization, a valve with a low pressure loss can be selected, but a non-negligible loss always remains, otherwise the control quality is affected.

The second major deviation is that the motor power drops with a delay, meaning that the motor runs at full power for longer. This is due to the fact that the valve is briefly closed during the changeover, but the pump continues to deliver. To prevent the pressure from rising uncontrollably at this point, a pressure relief valve or hydraulic accumulator must be used. For the simulation, an accumulator was assumed which is filled while the valve is closed and keeps the pressure at an elevated level for some time after the reversal of direction.

These two basic issues lower the efficiency of the hydraulic system, in the case of the performed simulation to a value of 40%. Depending on the degree of optimization of the system and the load, this efficiency underlies a tolerance and is not constant, but without a fundamental change in the schematic, a significantly better value is not achievable.

Movement inversion by pump

In the system approach with movement reversal using the pump, the valve is not required, and the movement is carried out by the pump displacement adjustment. In this assumption, the motor rotates at a constant speed. The simplified schematic is shown in Figure 10.

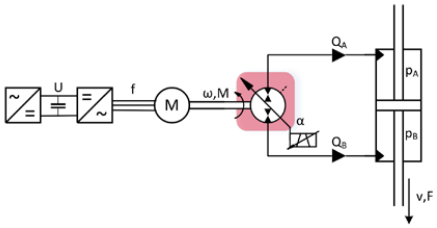


Figure 10: Simplified schematic with the pump as control component.

Compared to the valve-controlled system, the adjustment with the pump is less precise and slower. The lower precision can be compensated with suitable software, but high switching frequencies are not possible, especially with larger pumps. In order to still achieve a good flow rate, the stroke of the cylinder must be increased so that a good average piston speed is still possible.

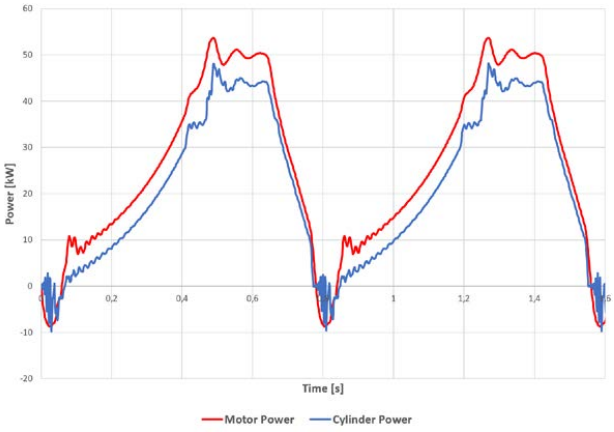


Figure 11: Motor and cylinder power in the pump-controlled system.

A system simulation was performed also for this system with the same boundary conditions as for the valve and is shown in Figure 11. It can be clearly seen that the motor power follows the cylinder power with a slight offset. The offset is due to motor and pump efficiencies and piping losses. However, since the valve from the previous section is removed, the level of offset is significantly lower. In addition, the engine power decreases equally with the cylinder power in this system structure.

Another characteristic is that the negative power reaches the pump and thus also the electric motor through the decompression of the dead volume. This can be used regeneratively with various concepts, e.g. charging a capacitor in the DC link or in the form of mass inertia of the motor shaft. However, it should be weighed up whether the additional expense is worth the energy saved. In any case, this circumstance should be taken into account and the energy should be used or dissipated in a targeted manner so that no uncontrolled over- or underpressures occur.

In the simulated example, with the same boundary conditions as in the previous section, an average efficiency of about 80% could be achieved. This increase of 40% compared to the valve-controlled system makes it clear that in many applications it can be worthwhile to invest the extra effort and cost in a variable displacement pump.

Movement inversion by motor

For the completeness, in addition to the two previous options, the adjustment with the motor should also be shown. Figure 12 shows the simplified schematic, which does not differ much from the pump-controlled system.

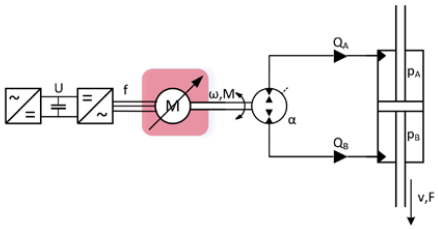


Figure 12: Simplified schematic with the pump as adjustment component.

There is a very strong similarity to the pump-controlled system in terms of advantages, disadvantages, and performance curve. The decisive difference, however, is that in the case of motion reversal, the mass inertia of the motor and pump shafts acts against the changeover. The extent to which this effect is significant depends on the dimensioning of the drive unit and can therefore be a decision criteria for or against this approach. Small drives are more suitable for changing the direction of rotation in a highly dynamic manner. It should therefore be verified for the case at hand whether this concept can be implemented. Of course, the energy stored in the rotation can also be used for the process, but this is only possible if the speed is continuously reduced towards the end of the movement. This fact therefore shows restrictions with regard to frequency maximization with this system structure, which is why it is probably only suitable in exceptional cases, like particularly small compressors with only a few kW.

Combination of movement inversion concepts

As is so often the case in technology, none of the extremes explained are the single best solution, but rather a suitable mixture of system structures that balance each other's disadvantages. Since these special solutions can create an USP, their implementation can’t be shown in detail.

As a first example, consider a valve-controlled system with a variable displacement pump. The offset due to pressure losses mentioned cannot be improved in this case, but the flow rate of the pump can be reduced during switching, when the valve is closed. With this measure, the efficiency can be significantly improved, but the advantage in cost and simplicity of the system are reduced.

Alternatively, it is possible to provide a system with variable displacement pump and variable speed motor. This makes it possible to select the most efficient operating point in terms of speed and displacement angle to maximize the efficiency of the motor-pump unit. Furthermore, the system allows the decompression energy to be stored in the form of shaft inertia without having to dissipate it in a motion reversal.

These two examples show that there is a large solution space of suitable system structures, but the characteristic advantages and disadvantages of the adjustment component for the motion reversal remain and can only be improved within certain limits.

5.4. Service Life

As a rough estimate 10,000 hours of operation between service intervals should be the target. This period of time is normally well within the capabilities of the pump, valves and motor, as long as the system is well implemented. This implies that there should be no unacceptable pressure peaks or cavitation. Likewise, the temperature and oil quality should remain within the specified limits of the individual components.

Although cylinders are known to be very robust components, a rough estimate of mileage shows why they can often be the critical component in this application. The average piston velocity $\bar{v}$ should be designed between 0.5 m/s and 1 m/s. In this range, the highest possible flow rate is achieved, but the sealing system should not yet be overloaded. With this specification and the time interval of 10,000 hours, it is estimated that 18000km to 36000km will be achieved. For industrial cylinders, 3000km is usually a rough guideline for when the sealing system might be worn out. One of the main challenges is therefore to design a sealing system that is optimized for the typical operating cycle and can thus seal off above-average mileage.

In addition to the mileage, the load cycles may also be estimated. These range from 10 million (0.3Hz) to 100 million (3Hz) between service intervals. Of course, only the seals and, if necessary, the running surfaces should be renewed during servicing and there should be no mechanical damage. This results in the requirement that all force-carrying components on the cylinder must be absolutely fatigue-proof.

6. CONCLUSION

Hydraulics are a suitable drive solution for small to medium compressors. However, to fully utilize its strengths, the physical principles of a gas compressor must be understood in addition to a deep understanding of the end application. Already when determining the system structure, there are several possibilities, which are more or less suitable depending on the target product. In the following dimensioning of the components, it is then essential to know how their properties contribute to the KPIs of the overall system. Even if this application seems quite simple at first due to the repetitive cycle, it is possible to create a great value added through specially developed components, system architectures and software, which contributes to the success of the OEM.

From the point of view of Bosch Rexroth AG, purely valve-controlled system structures are only a suitable solution in rare cases due to their low efficiency. Motor-controlled drives may be suitable for very small compressors in test technology, but a variable displacement pump will certainly be the

best choice for the broad range of applications with higher flow rates. Furthermore, it is advisable to design the system together with the OEM in order to point out the specific characteristics and to benefit from the advantages of hydraulics. Particularly in the case of cylinders, the use of standard products should be avoided, as these are not suitable for the loads involved and may result in premature fails.

NOMENCLATURE

A_{piston}	Piston surface area	m^2
$c_{p.H2}$	Constant pressure heat capacity of hydrogen	$J \times kg^{-1} \times K^{-1}$
f	Movement frequency	s^{-1}
ΔH	Enthalpy difference	J
l_{stroke}	Stroke length	m
m_0	Mass intake of stroke	kg
$\dot{m}$	Mass flow	$kg \times s^{-1}$
n	Polytropic exponent	-
p	Pressure	MPa
p_{out}	Discharge pressure	MPa
p_{in}	Inlet pressure	MPa
P_{comp}	Compression power	W
P_{el}	Electrical power	W
P_{loss}	Power loss	W
P_{mech}	Mechanical power	W
$\dot{Q}$	Heat flow	W
R_{H2}	Gas constant of hydrogen	$J \times kg^{-1} \times K^{-1}$
T_0	Starting temperature of hydrogen	K
T_{out}	Discharge temperature	K
$\bar{v}$	Mean piston velocity	$m \times s^{-1}$
V	Volume	m^3
V_0	Effective stroke volume	m^3
V_{dead}	Dead volume	m^3
V_{stroke}	Total stroke volume	m^3
ΔW	Work difference	J
W_{comp}	Compression work	J
x_{drive}	Piston position	m
$\dot{x}_{drive}$	Piston velocity	$m \times s^{-1}$
ρ_{H2}	Density of hydrogen	$kg \times m^{-3}$

REFERENCES

[1] European Forum for Reciprocating Compressor (2022) Hydrogen Compression Boosting the Hydrogen Economy.

[2] NIST Applied Chemicals and Materials Devision (2018) REFPROP Database 23. United States

[3] Eifler W, Schlücker E, Spicher U, Will G (2009) Küttner Kolbenmaschinen. Wiesbaden, Germany

[4] Langenheinecke K, Kaufmann A, Langenheinecke K, Thieleke G (2020) Thermodynamik für Ingenieure. Wiesbaden, Germany

Chapter 24

Actuators and Sensors-II

AUTOMATION OF PNEUMATIC THROTTLE CHECK VALVES BY USING NOVEL MULTI-STABLE SOLENOIDS

peer reviewed

Thomas Kramer*, Jürgen Weber

Chair of Fluid-Mechatronic Systems (Fluidtronics), Institute of Mechatronic Engineering, Technische Universität Dresden, Helmholtzstrasse 7a, 01069 Dresden

* Corresponding author: Tel.: +49 351 463-31964; E-mail address: thomas.kramer@tu-dresden.de

ABSTRACT

Multi-stable solenoids are novel energy-efficient actuator structures, which can hold any armature position powerless. Power is only required to change the armature position. They combine the continuous adjustability of proportional solenoids with the energy efficiency of polarised magnetic circuits. Thus, they are well suited for automating pneumatic throttle check valves, in order to set the throttle cross section and thus the cylinder piston velocity to a specific value and hold it for a certain time. This is useful in industry 4.0 context to produce on demand with various required piston velocities or for gradual compensation for increasing frictional forces during cylinder lifetime.
The focus of this paper is the application of novel multi-stable solenoids for replacing the todays widely used manual adjustment of throttle check valves. Therefore, two different throttle valves are designed: a spool valve and a poppet valve. They are based on the transfer characteristic of a conventional throttle check valve. The flow behaviour of both valves is investigated to show the principal transfer behaviour and disturbances. The valves are equipped with a multi-stable solenoid for demonstrating the adjustability of the throttle's volume flow rate and on that basis the adjustability of the piston velocity in a pneumatic cylinder drive.

Keywords: multi-stable solenoid, solenoid, pneumatic, throttle check valve, automation

1. INTRODUCTION

Electro-mechanical transducers are key components of fluid power systems. Especially as valve drives, solenoids based on the reluctance principle are widely used. For such solenoids, various design and operating principles have been established [1]. These can be arranged concerning the output behaviour (switching, continuous) and the magnetic circuit design (neutral, polarised) according to **Table 1**. Stable states are defined by no current flowing.

Table 1: Classification of existing solenoid principles

	switching (limited position number)	**continuous** (proportional)
neutral magnetic circuit	mono-stable switching solenoid	mono-stable proportional solenoid
polarised magnetic circuit	bi-stable, tri-stable, … switching solenoid	

Neutral magnetic circuits as mono-stable solenoids with their simple design are widely used. They consist of a coil, a yoke as fixed soft magnetic part and a movable soft magnetic armature, as seen in **Figure 1** left. The coil generates a magnetic field in the magnetic parts depending on the coil current i. The resulting magnetic flux Φ leads to a magnetic force F_m, which acts on the armature. In combination with a spring, an armature position x_a can be set according to the force balance of spring force F_s and magnetic force F_m. Such solenoids require a continuous power supply to hold any position except for the stable one given by the spring force (maximum air gap $x_\mathrm{a,max}$).

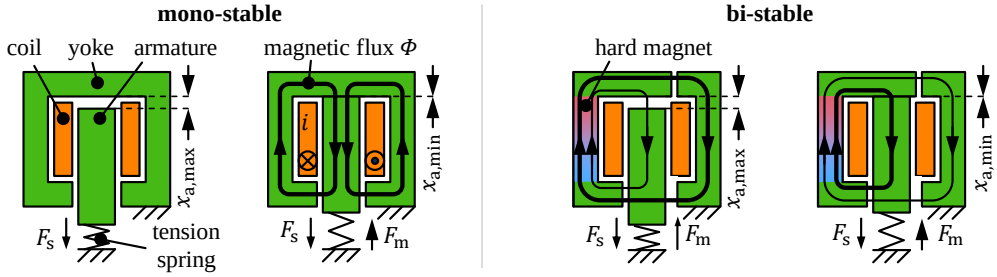


Figure 1: Principal design of neutral mono-stable and polarised bi-stable switching solenoid (based on [2])

Bistable solenoids [3–5] or solutions with more than two, but a limited number of stable switching positions [6, 7], are an efficient alternative to neutral mono-stable switching solenoids. Therefore, different design principles are known [2, 3]. Additional to the soft magnetic parts, they consist usually of a hard magnetic part, as seen in **Figure 1** right. This leads to a magnetic flux even without exciting the coil. The magnetic circuit is polarised. Thus, two holding forces can be realised, which can be hold without power supply: a spring force and a permanent magnetic force. The two holding forces result in two stable armature positions $x_{a,min}$ and $x_{a,max}$. The switching between the positions is realised only by a short current pulse through the integrated coil. The existing solutions of polarised magnetic circuits aim only at a limited number of switching positions.

Multi-stable solenoids

In order to combine both, the energy efficiency and the continuous adjustability, multi-stable solenoids are currently a focus of research. In [8], an approach for such multi-stable solenoids is published. The working principle utilises the inner magnetic hysteresis behaviour of semi-hard magnetic material. Such materials have a high energy storage capability and can be polarised by the solenoid's integrated coil. **Figure 2** illustrates the working principle with the material behaviour $B(H)$ and the basic design of multi-stable solenoids.

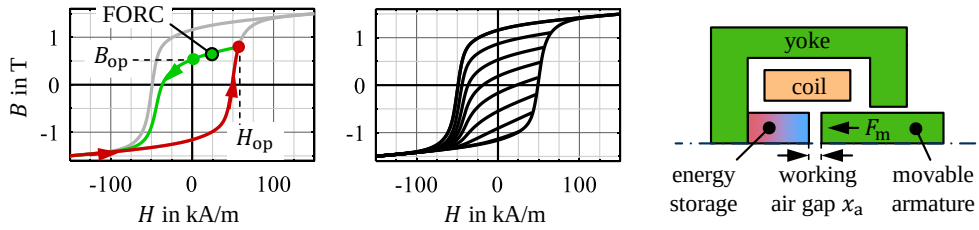


Figure 2: Magnetisation procedure for FORC generation (left), set of FORCs (middle), both for anisotropic CROVAC 12 from measurements using a self-made hysteresisgraph for hard magnetic materials (principle according to [9]), principal design of multi-stable solenoids (right)

The inner magnetic hysteresis behaviour can be accesses by so-called first order reversal curves (FORC). Therefore, the material is first magnetised along an outer hysteresis branch starting at saturation as reference (red line). After field reversal at any operating point H_{op}, the associated FORC arises (green line). If the external field H is shut down, the remanent magnetic flux density B_{op} remains. Depending on the reversal point H_{op}, any remanent point in the hysteresis loop can be reached as shown in **Figure 2** middle. Thus, the semi-hard magnetic material acts as variable magnetic energy storage.

The design of the proposed multi-stable solenoids is as simple as possible. The magnetic circuit consists of the energy storage, a yoke and an armature. The coil is in the center. With a pulsed coil excitation, a magnetic field is generated mostly in the energy storage for polarising it to a certain level. After shutting off the pulse, the remanent operating point according to the $B(H)$-behaviour is set. This leads to the proposed actuator functionality, where any operating point can be set and hold without additional power supply.

For detailed designing, an analytical approach was developed. The inputs of the approach are the maximum force potential $F_{\mathrm{m,max}}$ at a given air gap width $x_{\mathrm{a,max}}$, a semi-hard magnetic material behaviour, an air gap design and power supply parameters. Outputs are the size of the energy storage and the coil parameters. On that basis, demonstrators with three different semi-hard magnetic materials were build up. The behaviour of the corresponding materials is shown in **Figure 3** left. The demonstrators with anisotropic CROVAC 12 and AlNiCo 9 are designed for $F_{\mathrm{m}} = 50$ N at $x = 2$ mm. The demonstrator for hard ferrite Y35 has a lower force potential $F_{\mathrm{m}} = 15$ N at $x = 2$ mm due to significant reduced remanent flux density B_{r}. **Figure 3** middle shows the maximum magnetic force of the demonstrators.

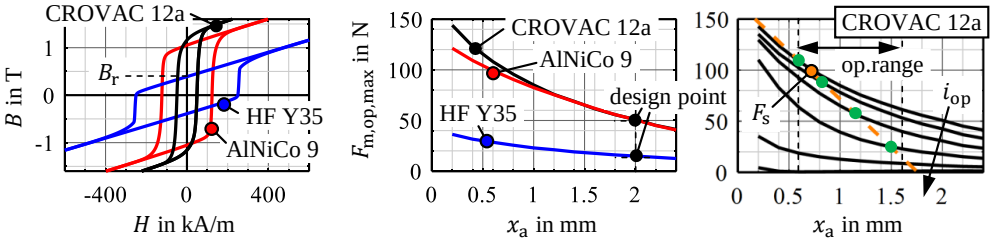


Figure 3: Magnetic material behaviour $B(H)$ of used energy storage materials (left), measured output behaviour $F_{\mathrm{m}}(x_{\mathrm{a}}, i_{\mathrm{op,max}})$ of the designed demonstrators at full polarisation (right), measured output behaviour $F_{\mathrm{m}}(x_{\mathrm{a}}, i_{\mathrm{op}})$ for setting remanent armature position at force balance points

For setting a remanent armature position, the magnetic force F_{m} has to act against a spring with the force F_{s}. Depending on the polarisation current i_{op}, any position can be set and held without power according to the force balance

$$F_{\mathrm{m,op}} = F_{\mathrm{s}} + F_{\mathrm{l}} = F_{\mathrm{s0}} - c x_{\mathrm{a,op}} + F_{\mathrm{l}}. \tag{1}$$

Therefore, **Figure 3** right shows the magnetic force characteristic for selected polarisation currents i_{op} and the resulting balanced operating points in green (see also [8]). Due to nonlinear force behaviour, the operating range is set by the spring to approximately 1 mm. It has to be taken into account that additional load forces F_{l} influence the force balance and thus the resulting armature position $x_{\mathrm{a,op}}$.

Applications of multi-stable solenoids

The potential fields of application for such novel multi-stable structures are manifold [8].

Especially in fluid power, new flexible and energy-efficient possibilities open up with this novel actuator principle. The manual adjustment of throttle check valves can be replaced, in order to adjust their opening cross-section automated during operation. This allows, for example, cylinder velocities to be set as required (industry 4.0: produce on demand) or a gradual compensation for increasing frictional forces during cylinder lifetime as a result of increasing seal wear. In addition, pressure relief, pressure reducing and proportional valves can be equipped with this novel actuator principle

to maintain any armature force or armature position without power.

In this paper, the focus is on automating pneumatic throttle check valves for demonstrating the novel actuator type in a real application. For this purpose, two throttles in form of a spool and a poppet valve are designed. The transfer characteristics of the throttles are measured to analyse the adjustability of the flow rate and to identify disturbance forces. The throttles are equipped with a suitable multi-stable solenoid demonstrator for investigating the setting of the throttle opening area and the resulting flow rate. The automated throttles are attached to a cylinder drive as meter-out flow control to demonstrate the setting of different cylinder piston velocities.

2. REFERENCE THROTTLE

Pneumatic cylinder drives are often equipped with a meter-out throttling to adjust the cylinder velocity $\dot{y}$ nearly independent from an external load force and from the supply pressure p_s as well as to increase the stiffness of the mechanical system. **Figure 4** left shows a corresponding circuit.

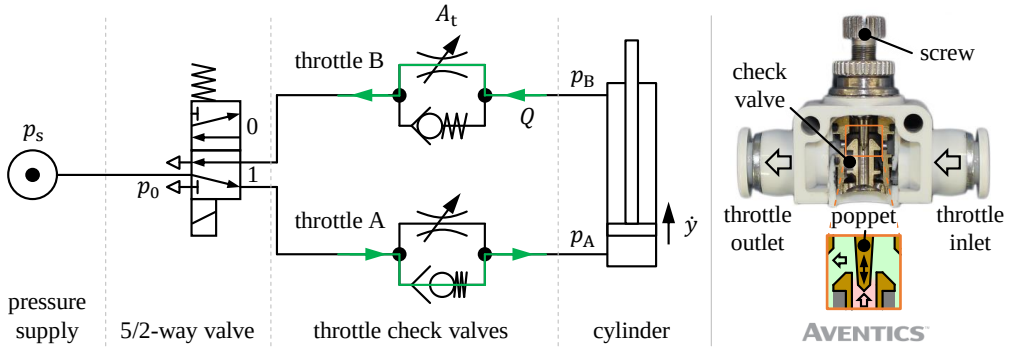


Figure 4: Pneumatic cylinder drive with meter-out control (left), reference throttle check valve as section model and principal cross section of throttle (right)

The metering is realized by two throttle check valves, one for each direction of movement. For moving the cylinder piston out, the mass flow goes through throttle B, as illustrated in the schematic. In contrast, throttle A flows through for moving in. By adjusting the throttle with the cross section A_t, the volume flow rate Q is set.

A common throttle for a typical cylinder size of 32 mm is an Aventics QR1-DBS-DA08, which is shown in **Figure 4** right. This throttle type is used in the paper as reference. The thumb screw serves for fine adjustment of the throttle cross section by moving the poppet. A key parameter of the throttle is the sonic conductance C at maximum opening. It is derived from measurements according to ISO 6358 [10], simplified for direct outlet to atmosphere with the pressure p_0. **Figure 5** left shows the measurement setup. It consists of a supply with variable pressure p_s, a volume flow rate sensor and a pressure sensor at the throttle inlet.

The supply pressure p_s is increased and decreased between 0 and 6 bar (relative) by manual adjustment of a usual pressure reducing valve. Meanwhile, relative inlet pressure p_1 and volume flow rate Q of the throttle are recorded. From the resulting $Q(p_1)$-behaviour in sonic range, the conductance is determined from the orange-coloured regression line with

$$C = \frac{\mathrm{d}Q}{\mathrm{d}p} = 39.4 \, \frac{\mathrm{Nl}}{\mathrm{min \cdot bar}} \tag{2}$$

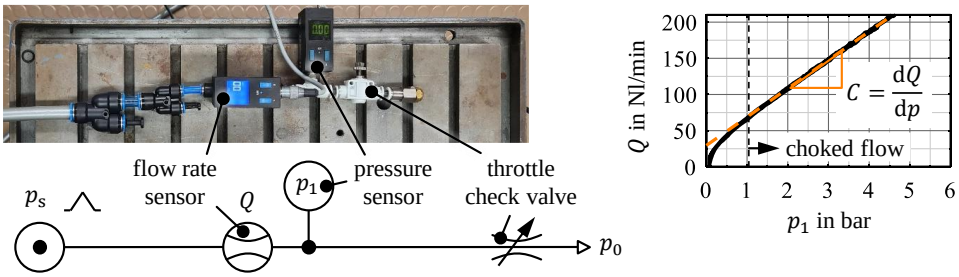


Figure 5: Measurement setup for characterising reference throttle Aventics QR1-DBS-DA08 (left), throttle transfer behaviour $Q(p_1)$ for determining sonic conductance (right)

3. THROTTLE DESIGN

On the base of the reference, two different throttles are designed and investigated: a spool valve and a poppet valve design. The spool valve is an ideal load case for the multi-stabile solenoid, since it has theoretically no relevant disturbing forces. The poppet valve is more a real load case for the multi-stable solenoid with its additional pressure force. Specific valves are necessary due to adaption to the multi-stable solenoid demonstrator.

For designing both throttle valves, the maximum cross section A_t has to be calculated from the reference sonic conductance. According to ISO 6358, the maximum mass flow rate is

$$\dot{m} = C p_1 \rho_0 \sqrt{\frac{T_0}{T_1}}, \tag{3}$$

considering a given pressure p_1 for choked flow and a given temperature T_1. In addition, the mass flow through a technical pneumatic resistance can be calculated by

$$\dot{m} = \alpha_\mathrm{d}\, A_\mathrm{t} \sqrt{2 p_1 \rho_1}\, \psi \tag{4}$$

[11]. The values are throttle parameters as cross section area A_t and discharge coefficient α_d, the air parameters pressure p_1 and density ρ_1 as well as the maximum value $\psi = 0.484$ of the discharge function considering choked flow.

By equating both, eq. (3) and (4), and using the ideal gas law, it results for the throttle cross section

$$A_\mathrm{t} = \frac{C p_0}{\alpha_\mathrm{d} \psi \sqrt{2 R_0 T_0}}. \tag{5}$$

The discharge coefficient α_d depends on the throttle design and is here assumed to a mid-size value $\alpha_\mathrm{d} = 0.7$ [11]. By applying eq. (5) with the sonic conductance from eq. (2), the maximum cross section of the reference throttle is $A_\mathrm{t,ref} = 4.7$ mm². This opening area serves as base for designing a comparable spool and poppet valve.

Spool design

For simple manufacturing, the throttle of the spool valve is realised by n_h cross holes with the diameter d_h. **Figure 6** left depicts the resulting throttle area exemplary at one hole.

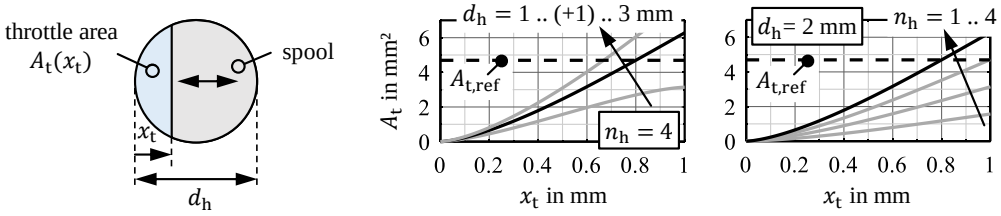


Figure 6: Principal design of spool valve based on cross hole (left), transfer characteristics of spool valve by varying hole diameter and hole number (right)

By moving the spool, the cross section can be enlarged or reduced. The position-dependent opened circular segment is calculated by

$$A_t(x_t) = n_h \left(\frac{d_h^2}{4} \arccos\left(1 - \frac{2x_t}{d_h}\right) - \left(\frac{d_h}{2} - h\right) \sqrt{d_h x_t - x_t^2} \right). \tag{6}$$

With variations of hole diameter d_h and hole number n_h, the behaviour $A_t(x_t)$ can be adjusted to fit the required cross section as shown in **Figure 6** right. The stroke of the full opening should be less than 1 mm, since it is the stable working range of the multi-stable demonstrators, as described above. A suitable combination is $d_h = 2$ mm and $n_h = 4$ with a resulting stroke of 0.8 mm as depicted in **Figure 6** in black. The resulting spool valve design is shown in **Figure 7**.

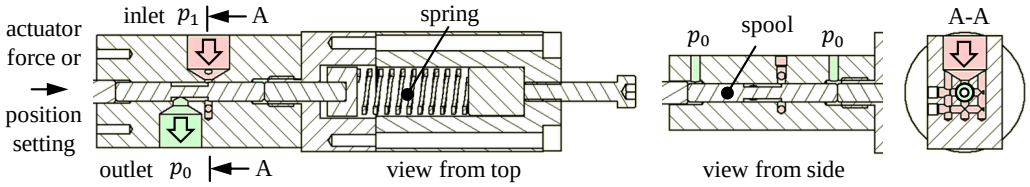


Figure 7: Design of spool valve

The drill holes are the inlet side. Here, tightness between throttle and fitting is possible. The spool is the outlet side. Tightness is here not important due to connection with the ambient pressure p_0. In order to avoid disturbance forces, the spool ends are unloaded with pressure equalisation holes to p_0.

Poppet design

The design of the poppet valve is illustrated in **Figure 8**.

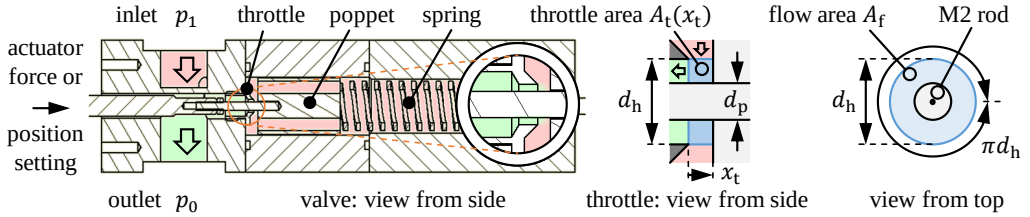


Figure 8: Design of poppet valve (left), principal design of its throttle (right)

The throttle cross section corresponds to the cylindrical surface

$$A_t(x_t) = \pi d_h x_t \tag{7}$$

with the hole diameter d_h according to **Figure 8** middle. The speciality is here the pin through the hole to actuate the poppet by pressing from the left side. It has to be taken into account that the pin reduces the hole area. Thus, it must be ensured that the resulting flow area

$$A_\mathrm{f} = \frac{\pi}{4}\left(d_\mathrm{h}^2 - d_\mathrm{p}^2\right) \tag{8}$$

is greater than the throttle area A_t. By using a M2 rod with a measured diameter $d_\mathrm{p} = 1{,}8$ mm, the hole diameter is set to $d_\mathrm{h} = 3{,}1$ mm resulting in a flow area $A_\mathrm{f} = 5$ mm². By applying eq. (7), the stroke for reaching the reference cross section $A_\mathrm{t,max} = A_\mathrm{t,ref}$ is $x_\mathrm{t,max} \approx 0{,}5$ mm.

4. THROTTLE TRANSFER BEHAVIOUR

Both throttles are investigated concerning its transfer behaviour. First, the real sonic conductance values are determined from measurements to validate the analytical throttle designs. Therefore, the measurement setup according to ISO 6358 in **Figure 9** is used (side view of solenoid test rig in [8]).

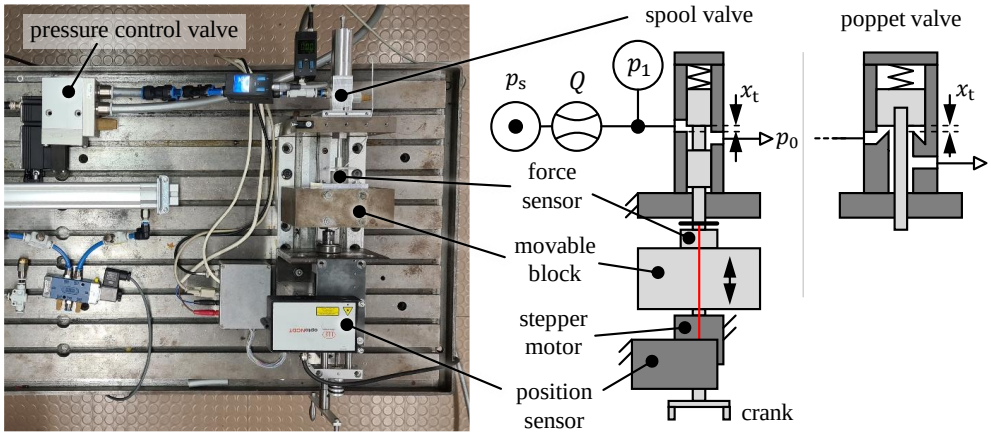


Figure 9: Measurement setup for characterising transfer behaviour of spool and poppet valve

The throttles are opened to the calculated stroke values for $A_\mathrm{t,ref}$ by moving the block to the corresponding position. Then, the pressure difference across the valve is increased and decreased. Therefore, a manual pressure reducing valve is used instead of the installed pressure control valve, shown in **Figure 9**. This is due to control discontinuities of the pressure control valve, especially at low pressures. During the supply pressure variation, the inlet pressure p_1 and volume flow rate Q are measured. **Figure 10** depicts the measurement results $Q(p_1)$ for both valves.

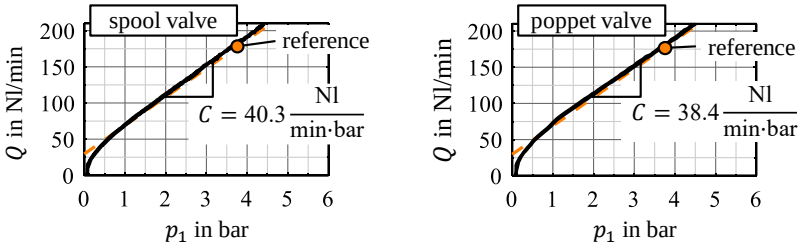


Figure 10: Transfer behaviour $Q(p_1)$ of spool and poppet valve for determining sonic conductance

From the curves, the sonic conductance is evaluated according to eq. (2). The values show a very good match with the reference throttle. Therefore, the designs are validated.

With the same test setup, here with pressure control valve, the transfer characteristic $Q(x, p_s)$ and force behaviour $F(x, p_s)$ are measured. The throttle cross section is increased and decreased by moving the spool and poppet, respectively forward and backward. **Figure 11** shows the resulting behaviour for different supply pressures p_s, which are set by the pressure control valve.

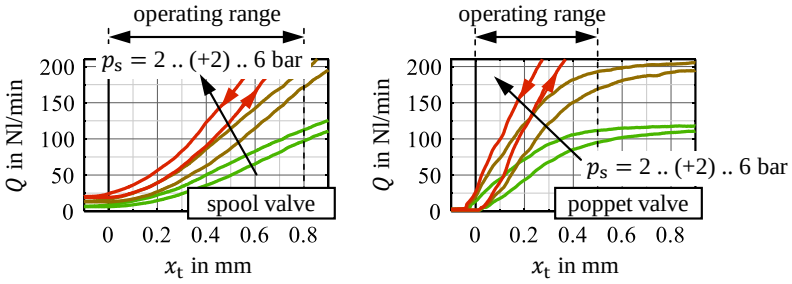


Figure 11: Transfer behaviour $Q(x_t)$ of spool and poppet valve

The curves show principally the expected behaviour. By opening the throttle, the volume flow rate Q increases. At the spool valve, the curves follow the cross section behaviour according to **Figure 6** with its non-linear range at small opening. The behaviour $A_t(x_t)$ of the poppet valve is linear until the flow is dominated by the limiting inner flow area A_f. At the calculated stroke, there is already a throttle effect shown. However, it has no significant influence on the sonic conductance according to **Figure 10**. The real hole diameter is possibly slight larger to compensate the inner throttle effect. Furthermore, the left graph shows at $x_t = 0$ that the spool valve is not tight due to the movement-related gap. This is less a problem in real applications, since the cross section is set to $A_t > 0$ for ensuring cylinder piston movement. The hysteresis effects in the measurements can, for example, be attributed to delays of the flow rate sensor due to its thermal heat-loss measurement principle.

In order to quantify possible disturbances such as friction and pressure force, the measured force behaviour acting to the pin is evaluated. **Figure 12** illustrates the behaviour $F(x_t)$ at the spool valve. It has to be considered that the measurements are done including the spring, whose force can be assumed as constant increasing behaviour.

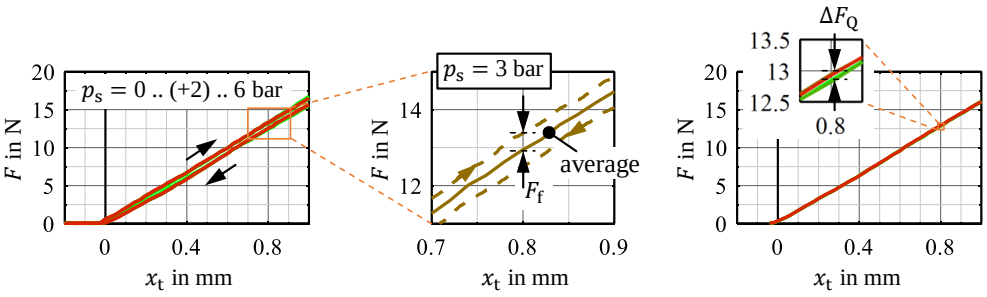


Figure 12: Measured force behaviour at various supply pressures for spool valve

The raw curves at the left side show only small hysteresis, mainly induced by friction forces $F_f \approx 0.4$ N. For investigating the pressure influence, average curves are generated as illustrated in **Figure 12** middle. The comparison of the mean curves at the right side shows only very small flow forces due to different pressures with $\Delta F_Q < 0.15$ N. Thus, the throttle can be assumed as an ideal

load case for the multi-stable solenoid.

The measured force behaviour of the poppet valve is depicted in **Figure 13**.

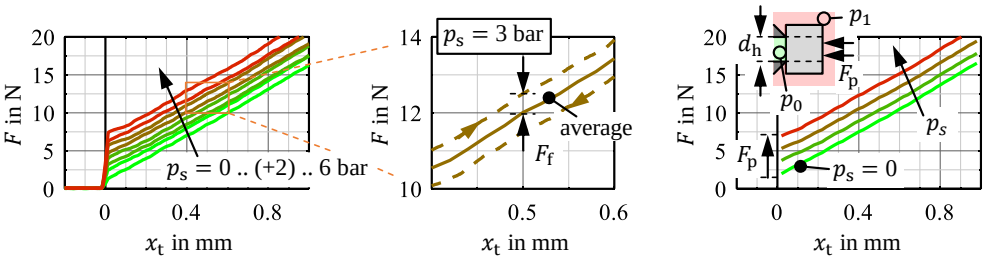


Figure 13: Measured force behaviour at various supply pressures for poppet valve

The hysteresis coming from friction force $F_f \approx 0.5$ N is comparable to the spool valve. The significant difference to the spool valve is the supply pressure dependency. This results from a part area A_p at the poppet's top, which is not pressure-compensated as shown in **Figure 13** right. With increasing pressure, the pressure force

$$F_p = p_1 A_p = p_1 \frac{\pi}{4} d_h^2 \tag{9}$$

increases as well. The theoretical value from the hole diameter $d_h = 3.1$ mm at closing throttle and $p_s = 6$ bar is $F_p = 4.5$ N. The measured value of 5 N differs a bit, possibly caused by the above addressed larger hole area. The pressure force decreases slightly with increasing throttle opening due to different pressure distribution at poppet's bottom.

5. MULTI-STABLE THROTTLE BEHAVIOUR

On the base of the principal transfer characteristics, a multi-stable solenoid is attached to the throttles for demonstrating the setting of any throttle opening and thus any volume flow rate. Firstly, a demonstrator has to be chosen, which roughly matches the force requirement. In order to demonstrate the influence of disturbance forces, the actuator force should be small. The demonstrator with hard ferrite Y35 has the smallest force potential $F_m = 15$ N at $x_a = 2$ mm (see **Figure 3**) and is chosen for the experiments. **Figure 14** illustrates the measured actuator behaviour and the chosen spring characteristic as well as the design of the demonstrator.

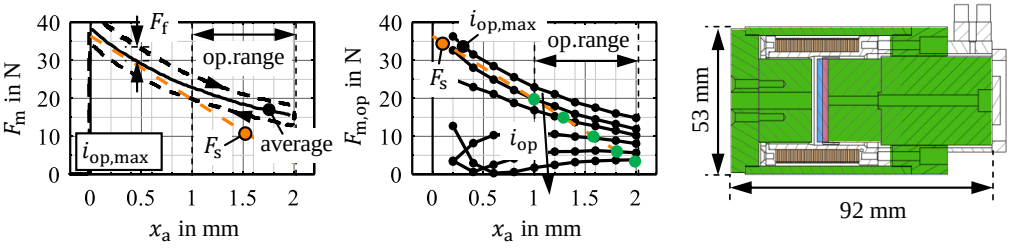


Figure 14: Measured actuator behaviour of the multi-stable demonstrator based on hard ferrite Y35 at full polarisation by moving the armature out and in (left), measured actuator behaviour at fixed positions and variations of polarisation current with remanent operating points at force balance points (middle), half section of used demonstrator (right)

The left side shows a typical force position characteristic with its striking hysteresis, here mainly coming from friction force F_f at the armature. For measuring these curves, the actuator is fully polarised first. Then, the armature is moving out and in, meanwhile the force is measured. This behaviour is utilised to set the spring preload. In addition, **Figure 14** middle shows the remanent magnetic force behaviour at fixed armature. Therefore, the armature is fixed to various positions x_a first. The actuator is then polarised by the current i_op. It results the operating point specific remanent magnetic force $F_\mathrm{m,op}(x_\mathrm{a}, i_\mathrm{op})$, which the figure shows. Depending on the polarisation current, different force curves arise as expected from the actuator principle. In combination with a spring, the remanent armature positions $x_\mathrm{a,op}(i_\mathrm{op})$ result according to the force balance from eq. (1) (green dots).

With this configuration, a working range of approximately 1 mm can be covered. This is sufficient for both, the spool and the poppet valve. It has to be noted that the throttle opening length x_t is inverse to the actuator's air gap (see **Figure 15** right):

$$x_\mathrm{t} = x_0 - x_\mathrm{a}. \tag{10}$$

The throttle is closed approximately at an armature position $x_0 = x_\mathrm{a,max} = 2$ mm.

On that basis, the multi-stable behaviour is investigated. Therefore, the measurement setup seen in **Figure 15** is used. This is based on the measurement setup for throttle characterisation according to **Figure 9**, which is extended by the multi-stable solenoid demonstrator. The actuator is equipped with a position sensor to measure the armature position x_a.

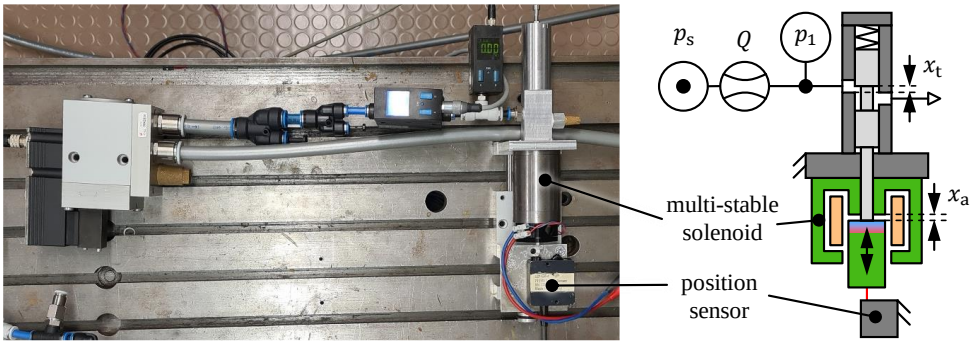


Figure 15: Measurement setup for characterising multi-stable throttle behaviour

With this setup, measurements are done for setting various remanent operating points. Therefore, a polarisation procedure is applied to the solenoid's coil, as exemplary illustrated in **Figure 16** for one operating point (similar to [8]).

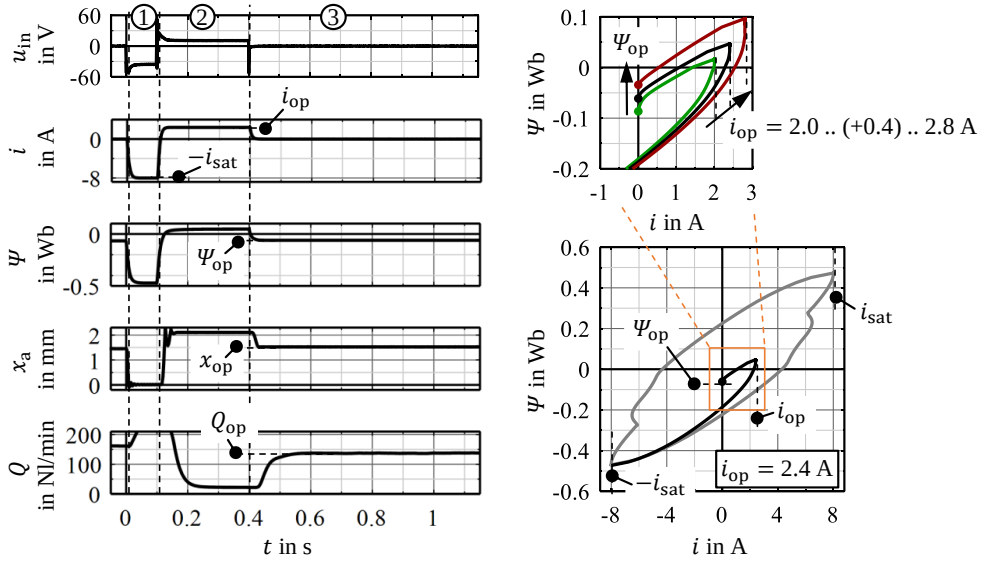


Figure 16: Polarisation procedure for setting the remanent operating point, time-dependent behaviour (left), electro-magnetic behaviour (right)

In the first step, the semi-hard magnetic material is excited to negative saturation, in order to enable a reproduceable reference point. Therefore, the coil current is regulated to $i = -i_{sat}$. In the second step, the material is polarised to a defined level by regulating the current to i_{op}. At last, the power supply is switched off, the current goes to zero and the remanent operating point sets in. For an inside in the electro-magnetic behaviour, the flux linkage Ψ as summed magnetic flux Φ over all windings w is calculated from input voltage u_{in} and coil current i according to the law of induction with

$$\Psi = w\Phi = \int u_{in} - Ri \, dt \tag{11}$$

(see also [8]). **Figure 16** shows at the bottom right side the setting of the remanent flux linkage Ψ_{op} in the electro-magnetic map $\Psi(i)$. The magnetic flux leads directly to a remanent magnetic force, which is simplified yielded according to the Maxwell's pulling force equation

$$F_m = \frac{\Psi_{op}^2}{2\mu_0 w^2 A_m} \tag{12}$$

with A_m as air gap area. In combination with the spring, the armature position x_a and thus the throttle position x_t is set to a remanent state according to eq. (1) and (10). After the polarisation process, the throttle cross section holds open at the corresponding level without additional power supply. Depending on the inlet pressure p_1 of the throttle, a volume flow rate Q_{op} is established.

By varying the polarisation current i_{op}, the remanent flux linkage Ψ_{op} and thus the remanent magnetic force $F_{m,op}$ can be set to any level in the operating range as depicted in **Figure 16** top right. Consequently, the volume flow rate Q_{op} can be adjusted as follows. **Figure 17** illustrates the setting of various armature positions $x_{a,op}$ and volume flow rates Q_{op} depending on the polarisation current i_{op} for the spool valve. Every operating point (i_{op}, p_s) was applied three times to show reproducibility. The shown curves correspond to the average of these three measurements.

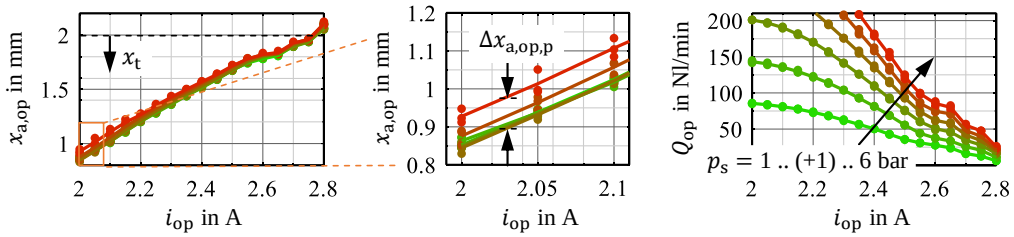


Figure 17: Multi-stable behaviour of armature position (left and middle) and volume flow rate (right) for spool valve

The graphs $x_{a,op}(i_{op})$ and $Q_{op}(i_{op})$ illustrate the basic working principle. Depending on the polarisation current i_{op}, the armature position x_a and thus the volume flow rate Q can be set to any value. The position behaviour is nearly linear as expected from the force balance points in **Figure 14**. Multiple measurements of the same operating point show small deviations up to $\Delta x_{a,op,n} = 0.1$ mm, especially at large throttle opening. Furthermore, the position setting shows a pressure dependency mainly on high pressure levels and large throttle opening, which is in the same size $\Delta x_{a,op,p} \approx 0.1$ mm. Slight flow effects according to **Figure 12**, small deviations in the friction and probably pressure dependent friction conditions in combination with the reduced load stiffness leads to these position deviations. As shown in **Figure 14**, with decreasing armature position, the angle and thus the stiffness between spring and magnetic force curves decrease as well.

In the behaviour of the volume flow rate, deviations can only be slightly seen due to limited measurement range of the sensor. However, it can be noted that the deviations are mainly in the range of large volume flow rates with high supply pressure and large throttle opening. Here are further investigations necessary, for example, with optimised friction behaviour.

The same measurements were done for the poppet valve. The results $x_{a,op}(i_{op})$ and $Q_{op}(i_{op})$ are depicted in **Figure 18**.

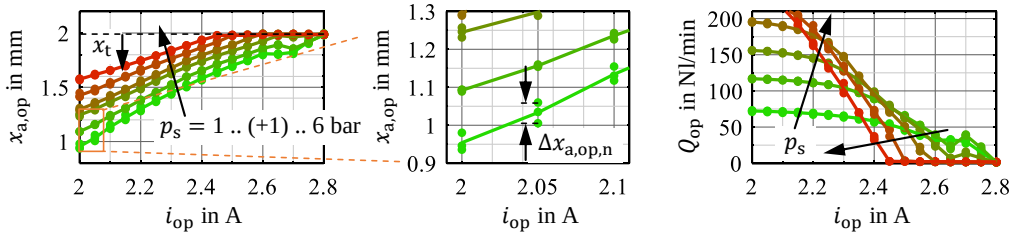


Figure 18: Multi-stable behaviour of armature position (left and middle) and volume flow rate (right) for poppet valve

In the position behaviour, the significant influence of the pressure force can be seen. According to the force balance

$$F_{m,op} = F_s + F_p = F_{s0} - cx_{a,op} + p_1 A_p, \tag{13}$$

the air gap $x_{a,op}$ increases with higher inlet pressure $p_1 \approx p_s$ for an assumed constant magnetic force $F_{m,op}$. This leads to a pressure-dependent throttle cross section and a corresponding influence on the volume flow rate. The effect is inherent for usual poppet valves and can be reduced by designing a valve with a pressure-compensated poppet or by using an actuator with a higher force

potential. With the second, the stiffness against load forces can be increased. Another possibility is the consideration of the pressure dependency in the control law $x_{\mathrm{a,op}}(i_{\mathrm{op}}, \Delta p_{\mathrm{v}})$ by measuring the pressure difference Δp_{v} across the valve during operation.

6. MULTI-STABLE ADJUSTMENT OF CYLINDER PISTON VELOCITY

The throttles with their multi-stable actuator are integrated in a simple pneumatic cylinder drive to demonstrate the adjustability of the piston velocity. The corresponding test setup is illustrated in **Figure 19**.

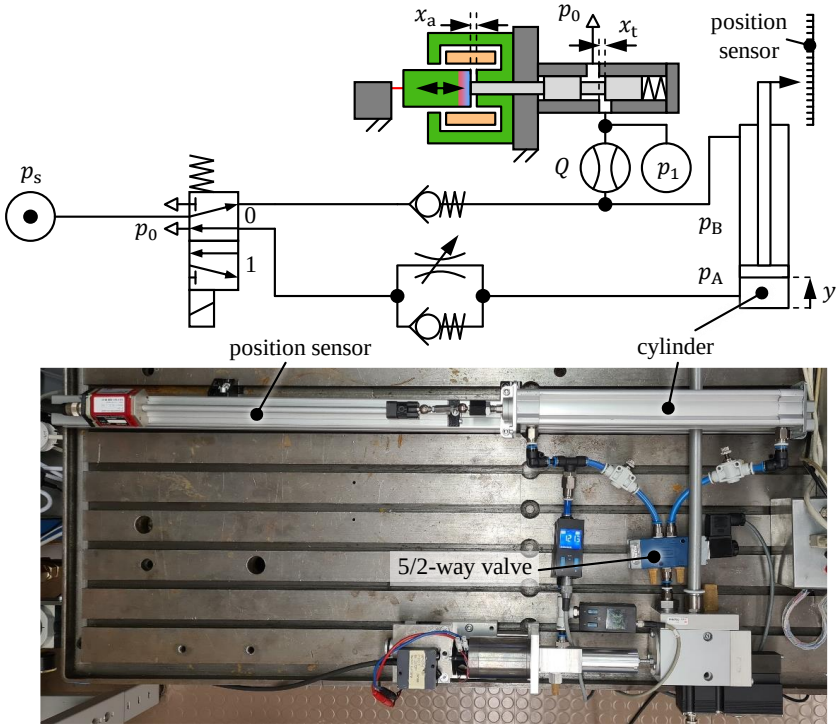


Figure 19: Measurement setup for characterising multi-stable adjustment of cylinder piston velocity

The setup is based on a conventional meter-out control system as shown in **Figure 4**. The throttle check valve for moving out is closed and bypassed by the presented spool or poppet valve. The meter-out has here the advantage of an outlet to ambient. This is important, since the design of the multi-stable solenoid demonstrator is not tight. The cylinder rod is attached to a position sensor. This measures the cylinder piston position y for determining the piston velocity $\dot{y}$. The cylinder has a diameter of 32 mm and a stroke of 250 mm.

With this setup, measurements are done for different supply pressures p_{s} and polarisation currents i_{op}, for every operating point three times. The procedure is as shown in **Figure 20**.

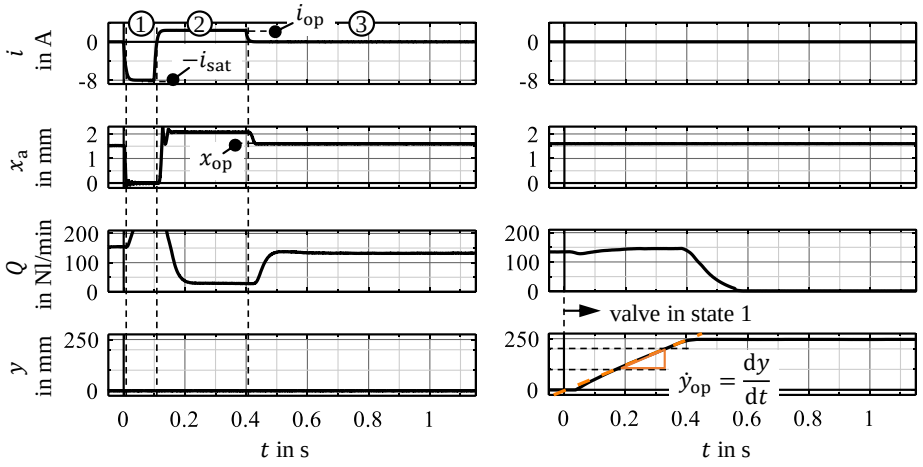


Figure 20: Procedure for setting defined cylinder piston velocity by using the spool valve, polarisation procedure (left), cylinder piston movement (right)

Initially, the 5/2-way valve is in state 0 (move in). The multi-stable solenoid is polarised with the current i_{op} according to the procedure in **Figure 16**. Afterwards, the valve is switched to state 1 for three seconds to move the piston out. At the end, the valve is reset to state 0 for reaching the start condition. During moving out, the velocity is determined by using a regression curve in the range $100\ \text{mm} < y < 200\ \text{mm}$. It is calculated from the derivation

$$\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}t}. \tag{14}$$

The results for the remanent armature position $x_{a,op}$ and mean piston velocity $\dot{y}_{op}$ depending on the polarisation current i_{op} by using the spool valve are shown in **Figure 21**.

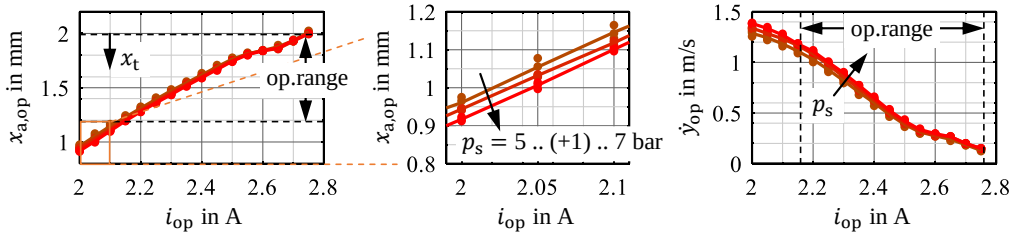


Figure 21: Multi-stable behaviour of armature position (left and middle) and cylinder piston velocity (right) for spool valve

The armature position behaviour is similar to the investigations without cylinder. Here it is noticeable that the pressure dependency of the remanent armature position has the opposite direction in contrast to **Figure 17**. Obviously, the friction conditions has no monotonic dependency from the pressure. This can already be guessed in **Figure 17**. As described above, further investigations are necessary to understand the behaviour in detail for its optimisation.

The resulting piston velocity shows the expected behaviour. Depending on the polarisation current i_{op}, any piston velocity can be set in a nearly linear dependency. The deviations for different supply pressures are slightly higher as expected from the variations of the armature position. This can

be attributed to pressure dynamics after valve switching.

The same measurements were done with the poppet valve. **Figure 22** shows the results.

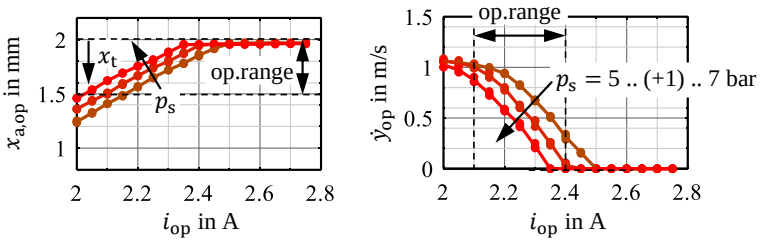


Figure 22: Multi-stable behaviour of armature position (left) and cylinder piston velocity (right) for poppet valve

In principle, the velocity can be set nearly in linear dependence to the polarisation current i_{op} as well. However, the behaviour of the armature position and thus of the piston velocity is influenced by the pressure difference through the valve as described above. Thus, the operating range depends on the pressure and is illustrated in the graphs for $p_s = 6$ bar. As result of the pressure dependency, the armature moves during piston movement as shown in **Figure 23**.

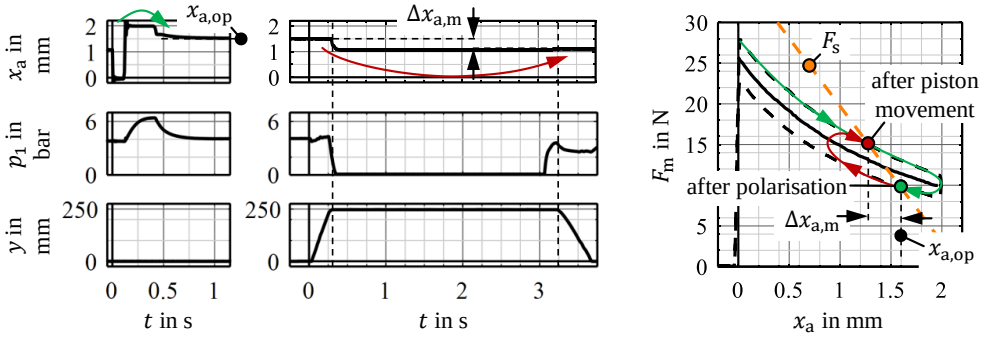


Figure 23: Time-dependent behaviour during polarisation and cylinder piston movement for poppet valve (left), exemplary force behaviour during polarisation and cylinder piston movement (right)

The problem is here a misalignment $\Delta x_{a,m}$ between start and end of piston movement cycle. This comes from armature friction, which depends on the movement direction, as illustrated in the force behaviour in **Figure 23** right. During the polarisation process, the armature comes from above the force balance point $x_{a,op}$ (green line), while it comes after piston movement from below (red line). This is caused by different pressure conditions during piston movement, which acts on the poppet. It results in different remanent armature positions according to armature movement direction due to friction. For solving this issue, the bearing concept has to be improved. For example, friction forces can be reduced nearly to zero by using membrane springs.

The investigations show that multi-stable solenoids are suitable for automating throttle check valves, in order to vary the piston velocity. Therefore, spool valves are preferred due to their very low disturbances. Poppet valves are only usable with additional effort, either by reducing friction and pressure forces or by considering these forces in the control.

7. CONCLUSION AND OUTLOOK

In this work, the automation of pneumatic throttle check valves by using novel multi-stable solenoids is investigated. Based on a conventional throttle check valve, a spool and a poppet valve are designed. Analytical geometrical approaches are used for dimensioning the throttle functionality of the valves. The approaches are validated with measurements. On that basis, the throttles are measured concerning its flow characteristics, which show the expected adjustability by changing the throttle opening section. The spool valve exhibits no significant disturbances and is therefore an ideal load case for multi-stable solenoids. The poppet valve has a significant pressure dependency and is more a real load case for the novel actuator type. The valves are connected to a demonstrator of a multi-stable solenoid, in order to investigate the setting of remanent throttle cross sections. The principal functionality is proven. As expected, the pressure conditions at the poppet valve have an influence on the remanent armature position and thus on the resulting volume flow rate. At the end, the automated throttles are integrated in a pneumatic cylinder drive for demonstrating the setting of different piston velocities. The measurements show that the spool valve in combination with a multi-stable solenoid is well-suited for the automated adjustability of the piston velocity. The poppet valve has too many influences coming from friction and pressure forces and is in this state inappropriate for this task.

Further investigations are dealing with the influence and reduction of friction for optimising the transfer characteristics of the valves. In addition, it is planned to develop a specific multi-stable actuator for the spool valve, which is significantly smaller than the demonstrator used for the investigations here. By using AlNiCo 5 as energy storage material for a force potential $F_{\mathrm{m}} = 18$ N at $x_{\mathrm{a}} = 1$ mm, the actuator volume can be reduced to ca. 30 %. Beside the application of multi-stable solenoids for driving pneumatic throttle check valves, there are a variety of other applications to be investigated. These include, for example, applications in hydraulics or processing technology.

ACKNOWLEDGEMENT

The research presented in this paper is based on the project “Novel multi-stable electro-magneto-mechanical actuators based on variable remanent operating points of semihard- and hard-magnetic materials”, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 434232806.

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

NOMENCLATURE

A	area	m^2
B	magnetic flux density	T
d	diameter	m
c	stiffness	N/m
C	sonic conductance	Nl/(min·bar)
F	force	N
H	magnetic field strength	A/m
i	current	A
$\dot{m}$	mass flow rate	kg/s
n	number	-
p	pressure	bar
Q	volume flow rate	Nl/min

R	gas constant	288 (N·m)/(kg·K)
t	time	s
T	absolute temperature	K
u	voltage	V
w	winding number	-
x	position	m
y	piston position	m
$\dot{y}$	piston velocity	m/s
α_d	discharge coefficient	-
μ_0	vacuum permeability	$4\pi \cdot 10^{-7}$ (V·s)/(A·m)
ρ	density	kg/m³
Φ	magnetic flux	Wb
ψ	discharge function	-
Ψ	flux linkage	Wb

REFERENCES

[1] Kallenbach E, Eick R, Ströhla T, et al (2018) Elektromagnete: Grundlagen, Berechnung, Entwurf und Anwendung, 5. Auflage. Springer Fachmedien Wiesbaden, Wiesbaden

[2] Roschke T (2004) Potenzial bipolarer Magnete in Verriegelungs- und Hubanwendungen. In: Innovative Klein- und Mikroantriebstechnik: Vorträge der ETG-/GMM-Fachtagung. Margret Schneider, Darmstadt, Germany

[3] Krauß L (1984) Ein Beitrag zur Auswahl, zum Entwurf und zur Ansteuerung von bipolaren Elektromagneten, insbesondere für Magnetventile. Dissertation, Technische Hochschule Ilmenau

[4] Burmeister LC (1967) NASA Contributions to Advanced Valve Technology - A survey, revised and enlarged edition. National Aeronautics and Space Administration

[5] Johnson BG, Massey SE, Sturman OE (2001) Sturman Digital Latching Valve. Linköping, Schweden, pp 299–314

[6] Olbrich M, Schütz A, Bechtold T, Ament C (2021) Design and Optimal Control of a Multistable, Cooperative Microactuator. Actuators 10:183. https://doi.org/10.3390/act10080183

[7] Uusitalo J-P (2010) A novel digital hydraulic valve package a fast and small multiphysics design. Dissertation, Tampere University of Technology

[8] Kramer T, Weber J (2023) An approach for novel energy-efficient multi-stable solenoids and the demonstration of its fundamental behaviour. In: 18th Scandinavian International Conference on Fluid Power (SICFP 2023). Tampere, Finland

[9] Czichos H, Saito T, Smith LR (2006) Springer handbook of materials measurement methods. Springer, Germany

[10] (2013) ISO 6358-1:2013-05 - Pneumatic fluid power - Determination of flow-rate characteristics of components using compressible fluids - Part 1: General rules and test methods for steady-state flow. ISO

[11] Beater P (2007) Pneumatic drives: system design, modelling and control. Springer, Berlin, London

LOW ENERGY CONSUMPTION HIGH FLOW CONTROL SYSTEM USING SPOOL-IN- SPOOL DESIGN OF PROPORTIONAL VALVE

Dipl.-Ing. Jan Koudelka[1]*, Jaroslav Dvořáček[1], Dipl.-Ing. Filip Vlach[1],

Dipl.-Ing. Pavel Jech[1], Dipl.-Ing. Jan Bjelka[2]

[1] *Design Department, ARGO-HYTOS s.r.o., Dělnická 1306, Vrchlabí, Czech Republic*
[2] *Sales Department, ARGO-HYTOS s.r.o., Dělnická 1306, Vrchlabí, Czech Republic*

* Corresponding author: Tel.: +420 771 272 891; E-mail address: j.koudelka@argo-hytos.com

ABSTRACT

Recently, in many mobile applications, great emphasis has been placed on reducing energy consumption, whilst maintaining the high hydraulic performance of the system. At the same time, a small size and, if possible, a low weight must be maintained if the system is used to be used on moving parts of machines.

All these requirements are met by our new high flow control system using the proven spool-in-spool proportional valves principle. This principle has now been enhanced using LS functionality directly on the spool incorporating a brand-new seven chamber design. Our spool-in-spool valve principle enables control of high flow rates and pressures through the valve with a very small solenoid. The large hydraulic power transferred by the valve and low-pressure losses of the solution, enable high dynamics and proportional controllability of the entire system.

The sandwich construction of the control block increases the variability of the entire device whilst space requirements are significantly reduced.

The function of the pressure compensator, together with the LS control of the hydraulic pump, eliminate the pressure drop in individual sections during simultaneous use and increase energetic efficiency of the solution. Everything is included in the individual sections.

Keywords: Proportional, Spool-in-spool, LS, High flow, Pressure drop, Small solenoid

1. SPECIFICATION OF PROJECT

Abbreviations:

PSVS: Proportional section valve system

PMVS: Proportional modular valve system

MPC: Main pressure compensator

(E)-LS: (Electronic) Load sensing

HIC: Hydraulic integrated circuit

PIB: Parts in body

PRV: Pressure relief valve

SIC: Screw in cartridge

1.1. Critical targets definition

- Fine proportional control in the beginning of flow scale. This practically means smooth and gentle control of heavy loads (p=150 bar at Q=0-20 l/min). The movement of the load is driven by a hydraulic motor and linear hydraulic cylinder, and both are in operation at the same time.
- The changes of the load or sudden reaching of cylinder end position must not influence the rotation part of the movement and vice versa. This leads to a solution with pressure compensating elements with every single proportional valve in the system.
- Fast acceleration of the load movements (to Q=80 l/min at Δp=10-14 bar) is required as well and hydraulic system must be able to provide that with full proportional control without excuses and interruptions.
- PSVS must work with common Load sensing concept which must fully cooperate with hydraulic drives incorporated in the base machine platform. PSVS can share pressure load information with the machine basis in two ways – hydraulically or electronically. And actually, the second way seems to have more benefits so all needed service ports for full E-LS control are included. The critical condition for LS system here is that when the proportional valve is in neutral position and actuator gets the external hit which causes pressure peak in the channels then this pressure peak must not be transmitted to main LS line. This will secure the LS pump drive is not disturbed by external influences and general pressure supply to the subsystem work calmly.
- Essential part of PSVS LS or E-LS system is also main pressure compensator (MPC). The purpose of that element is to unload the main pressure feeding line to tank line locally in PSVS central block to save as much energy as possible in shortest possible reaction time. MPC regulates the amount of available energy and saves the energy during all operations based on actual system demands as well.
- Space demand on complete solution is crucial. The PSVS must be light and compact as much is possible within physical laws.
- The overall system must work with the highest possible efficiency. This means the manifold design must respect the optimal oil flow path and internal channel shape and channels connections must be optimized.

1.2. PSVS conceptual art:

If we put the pure functional points of this system to the side and focus just on given space and given interface between basic machine platform and PSVS we see the traditional hydraulic components which are accessible on the market cannot do the job. An underlined point is the fixed displacement of ports on the machine platform this means PSVS must be in line with all details without exclusions. Considered options below:

Solution with SIC valves (HIC)

Screw in cartridges provides flexible basis for efficient mobile outdoor HIC in many cases. So mechanical valves like compensators or PRV seems to be the best option. Unfortunately, when we speak about proportional direction control functions 4/2, Q=80 l/min flow the SIC is not ideal. There is lack of performance or solenoids would need to be too big for the given space. Generally, the pressure losses are higher than other solutions.

Solution with CETOP valves (HIC)

CETOP proportional flow control valves could handle the flow so we could find valves with enough performance, but for the given space are also too big and they are less suitable for mobile outdoor

applications. PRV or pressure compensators in CETOP housings provides very low-pressure losses on the other side they are space demanding too.

Solution combination of SIC and CETOP (HIC)

Combination of SIC and CETOP valves are the very efficient solutions for many kinds of applications. Unfortunately, the target flow demands Q=80 l/min makes this option also too big and too heavy.

Solution with section mobile bankable valves

The space issue could be solved by a bankable valve system. The given hydraulic interface can be additionally solved with a special interface plate but here we find problematic proportional flow control function with LS feature. The prop. control both directions Q=80 l/min at Δp=10-14 bar built in bankable section is known on the market but unfortunately it must be pilot operated, and every producer provides their own design and solution. Most of them have a design which cannot be easily adapted to the given interface and space. There is also an important question about the response of pilot operated proportional controls and its working characteristics.

HIC based on taylor-made casted housing with parts in body (PIB) valve versions. Bankable sections concept with special central inlet casted block.

According to the above analysis the concept has been evolved. HIC based on taylor-made casted housing with parts in body (PIB) valve versions. Bankable sections concept with special central inlet casted block. Proportional direction flow control pilot operated valve with floating spool concept and with direct load sense via the proportional spool´s 7th chamber architecture.

2. HISTORICAL EVOLUTION OF THE VALVE (COMPARISON)

To meet the definition of critical targets, we were inspired by development of our PRM8-10 CETOP 5 valve. We developed a new 7-chamber valve by extension with LS channels and tested the function together with external pressure compensator as modular assembly (PMVS). Based on the test results, we continued with development of a 7-chamber section valve with integrated pressure compensator in one housing. In the sub-chapters below, the individual phases of development are listed.

2.1. PRM8-10 valve

Figure 1: Pilot operated proportional directional control spool valve with high hydraulic power [1].

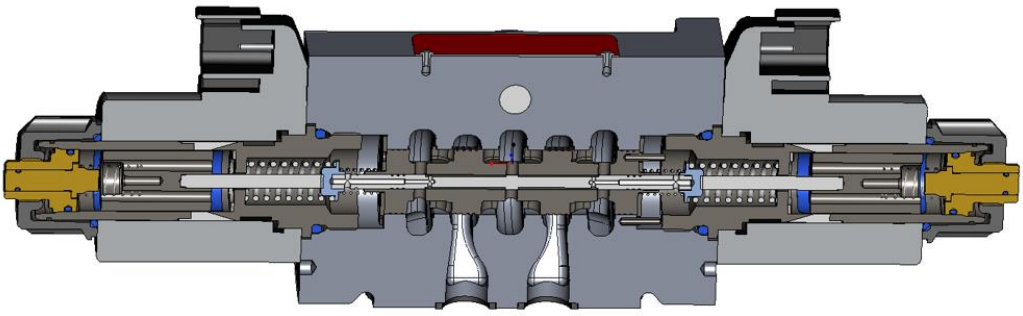


Figure 1: PRM8-10

2.2. 7-chamber valve

PRM8-10 valve extended by LS signal brings the valve with 7-chamber housing shown in **Figure 2**. The housing was produced by additive manufacturing technology. Thanks to that we have achieved a significant time reduction for producing the functional sample. It allows us to use a bigger chamber, channels and spool diameter as well, which cause lower pressure drop. It caused higher flow at the same inlet pressure as it is shown in **Figure 5**.

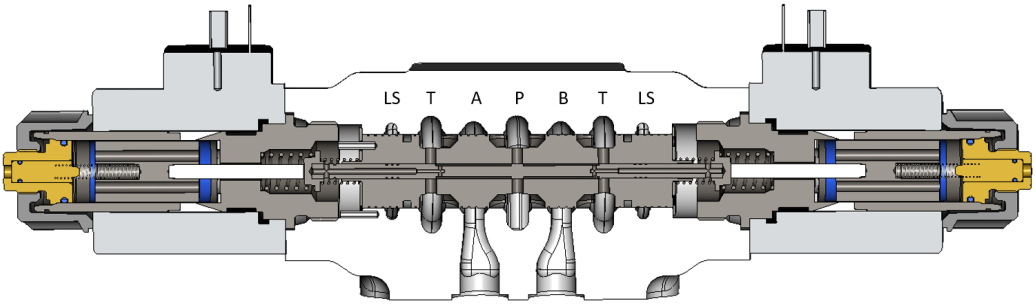
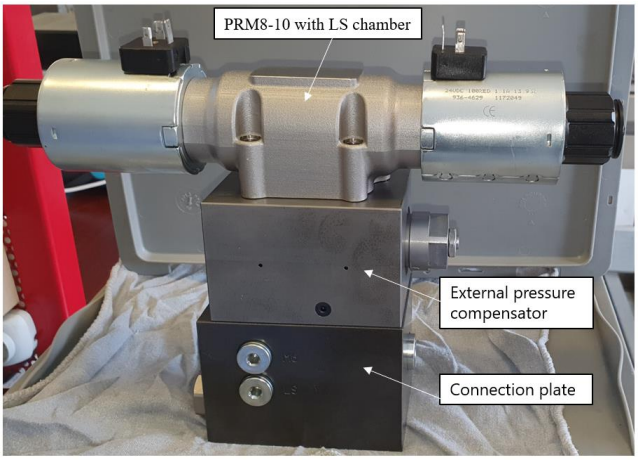


Figure 2: 7-chamber valve

This valve together with external two-way pressure compensator and connecting plate makes a modular assembly for initial testing to see, how the system works independently on the load. Load level was set up by pressure relief valve placed between channels A and B shown in **Figure 3.** During the measurement, we recorded the pressure values in the individual channels to know the pressure drop of the valve itself and all modular assembly.



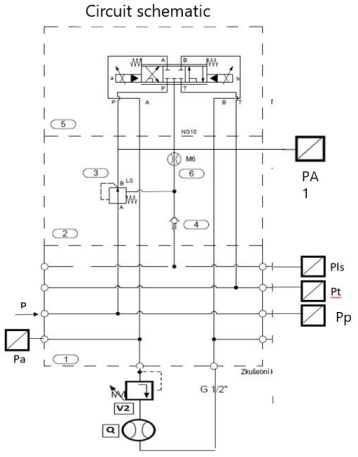


Figure 3: 7-chamber valve modular assembly - PMVS

2.3. 7-chamber section valve with integrated pressure compensator - PSVS

Figure 4: PSVS solution connected the benefits of the 7-chamber valve and pressure compensator (which is parts-in-body here) in one casted housing. Casted housing has even bigger channels then previous one shown in **Figure 2**. Here, we achieve even smaller pressure drops and even higher flow at the same inlet pressure as it is shown in **Figure 5**. At the same time, we reduced the overall size and mass as it is shown in **Table 1.**

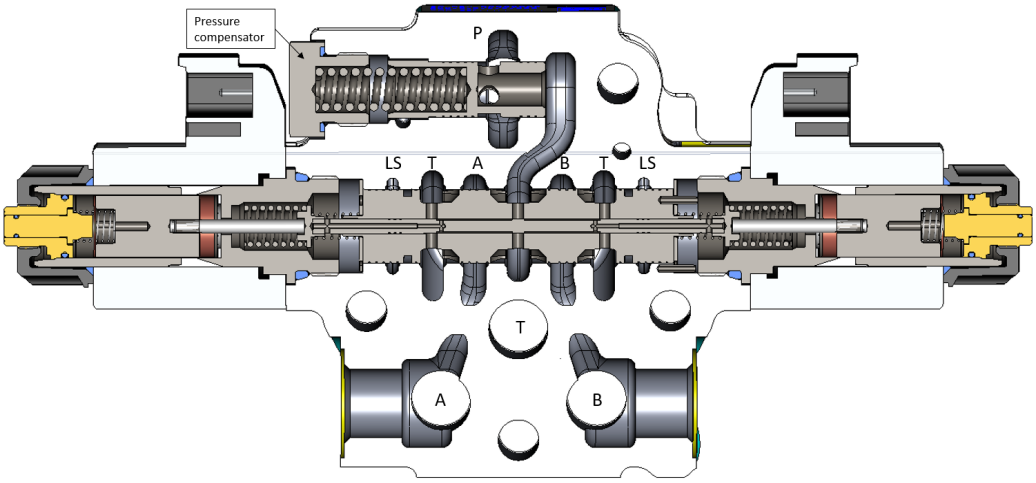


Figure 4: 7-chamber section valve assembly - PSVS

2.4. Measurement comparison

Figure 5 describes the flow through the valve as a function of actuating current at inlet pressure p_{in}=250 bar, oil viscosity ν=32mm^2/s, temperature T=40°C. Proportional function is controlled by our electronic control unit EL7-E [2] with frequency f=90 Hz and amplitude A=15 %. To be able to compare the results properly, the PSVS version was tested with no function of pressure compensator. That means the pressure compensator was fully open.
Measurement describes the flow increase of 7-chamber valves against the standard PRM8-10. With the PSVS the flow is more noticeable, Q=210 l/min. In percentage the value is 31 %.

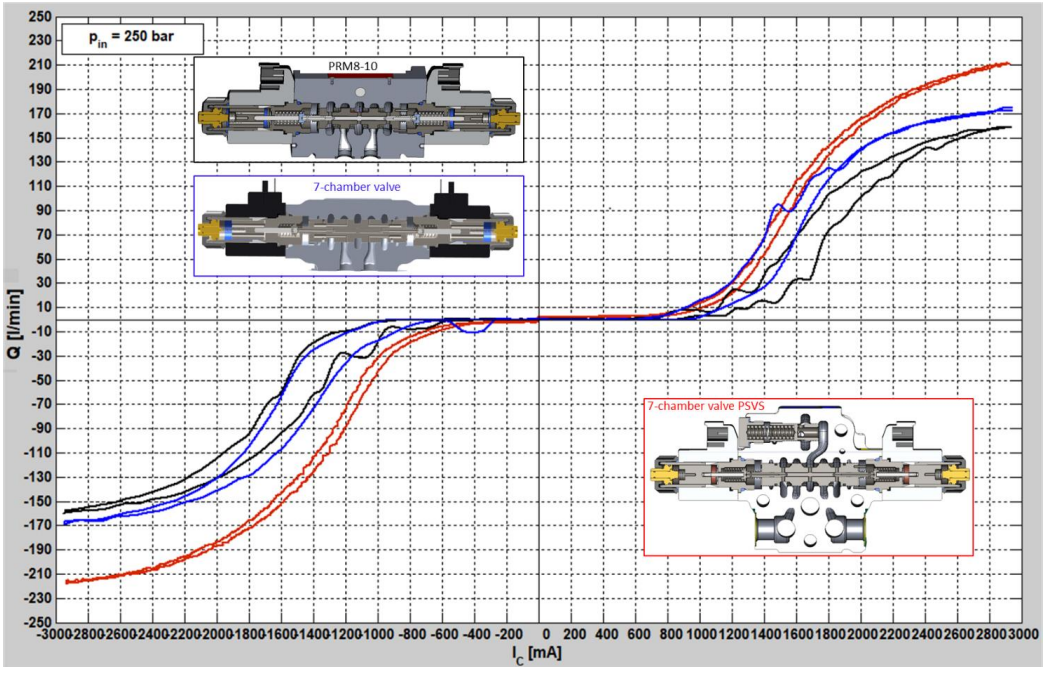


Figure 5: Flow characteristics comparison

Table 1 describes the size and the mass differences between Modular and Section solution. Here you can see massive improvement at both parameters with the Section solution.

Table 1: Modular and sectional comparison

parameters	Size HxWxL [mm]	mass [kg]
Modular solution	222x70x278	12
Section solution	127x48x278	4,4

3. PSVS - FINAL HYDRAULIC RESULTS

Figure 6 describes the flow through the valve as a function of actuating current at different inlet pressure. Pressure compensator is fully open. Other test conditions are the same as described in the chapter 2.4. The flow is limited by the flow value Q=220 l/min at the pressure p_{in}=300 and 350 bar. This is due to our test laboratory pump maximum.

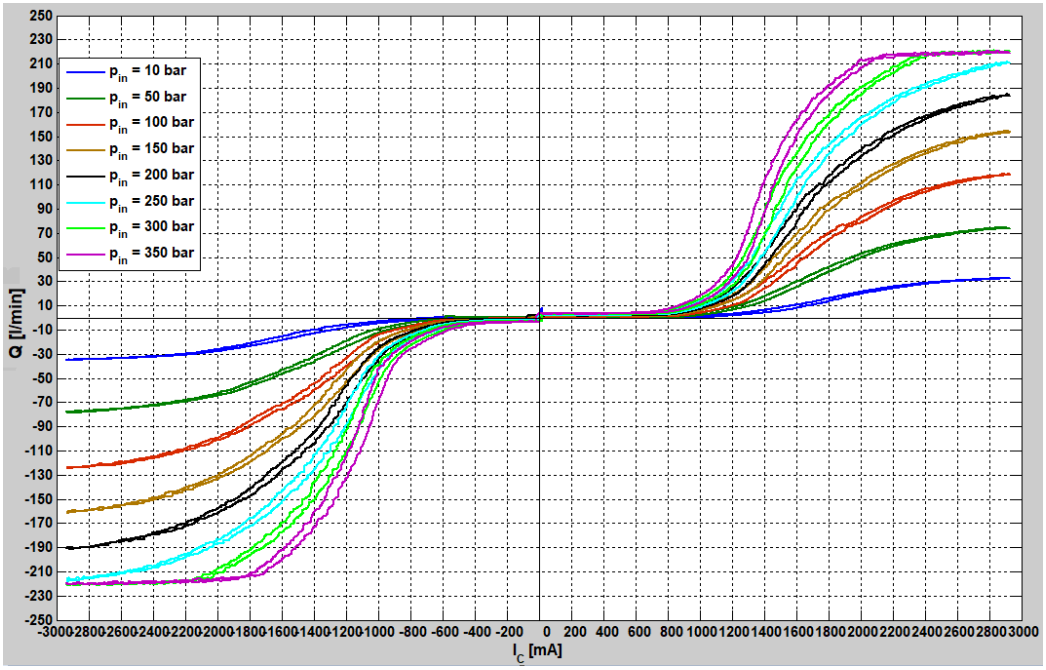


Figure 6: Flow characteristics with fully open pressure compensator

Figure 7 describes the flow through the valve with the different load between the channels A and B, as a function of actuating current at inlet pressure p_{in}=200 bar. Pressure compensator is fully open. Other test conditions are the same as described in the chapter 2.4.

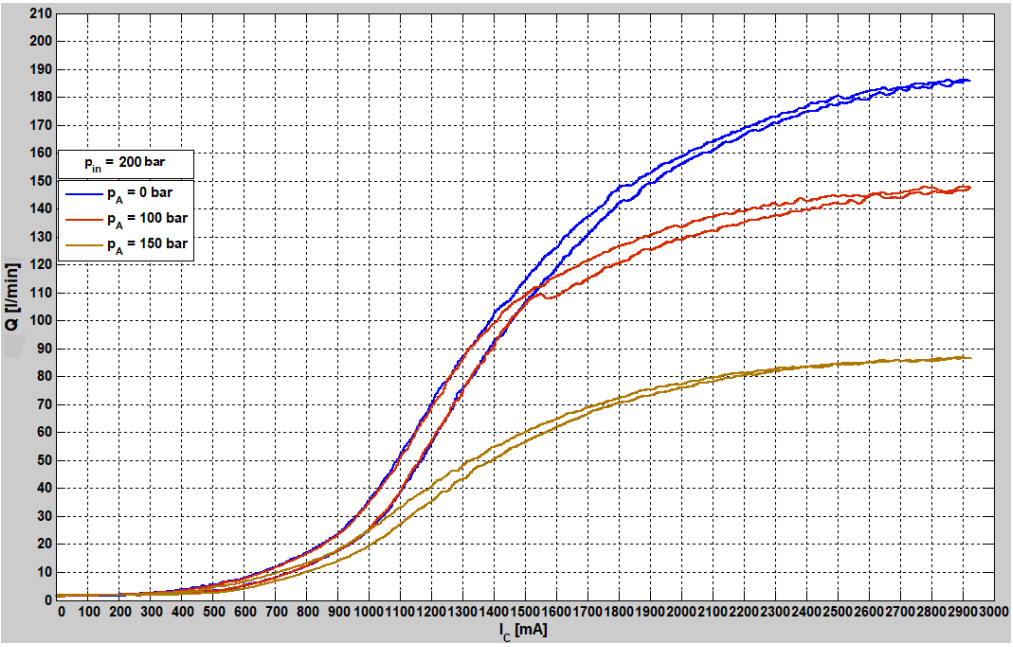


Figure 7: Flow characteristics with load and fully open pressure compensator

Figure 8 describes the flow through the valve with the different load between the channels A and B, as a function of actuating current at inlet pressure p_{in}=200 bar. Other test conditions are the same as described in the chapter 2.4. Pressure compensator is here in the function, set on the value Q=80 l/min and works the same with different loads.

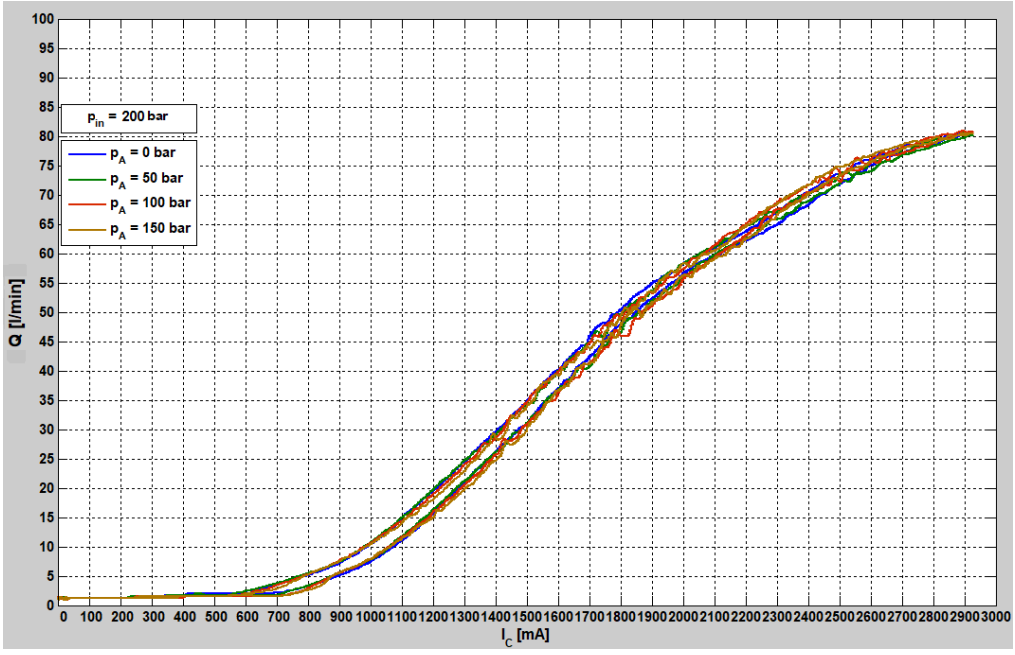


Figure 8: Flow characteristics with load and working pressure compensator

4. SUMMARY

In our paper, we describe the development and measurement results of a proportional, indirectly controlled spool valve for high flow rates. The basic requirement for a solution consisting of a high-flow valve with a pressure compensator, which are implemented in one small sandwich casting, are met. Due to the use of the electromagnet size NG06 and high transmitted power, the ratio of electromagnetic control to controlled hydraulic power is high.

By increasing of the spool diameter and optimizing the casting channels, we made the valve more efficient compared to the CETOP version PRM8-10 by 31% at p_{in}=250 bar, while maintaining the proportional function.

By implementing a pressure compensator and LS control in each sandwich, we have reduced flow fluctuations when several sections are operating at the same time.

We can use two different spool connections in sections. The spool that is closed without the signal -- 3Z11 and the spool that has channel P closed and channels A and B led to the tank - 3Y11.

The ability to connect multiple sections with two different spool connections allows us to cover a wide range of functions in mobile applications.

NOMENCLATURE

Q	Flow	l/min
p_{in}	Inlet pressure	bar
p_A	Load pressure	bar
Ic	Actuating current	mA
H	Height	mm
W	Width	mm
L	Length	mm
M	Mass	kg
f	Frequency	Hz
A	Amplitude	%
T	Temperature	°C
ν	Kinematic viscosity	mm^2/s

REFERENCES

[1] https://www.argo-hytos.com/ - PRM8-10 Data sheet

[2] https://www.argo-hytos.com/ - EL7-E Data sheet

REAL-TIME MODELS OF VALVE SOLENOIDS: AN EVALUATION OF MEASUREMENT AND SIMULATION-BASED PARAMETER IDENTIFICATION

peer reviewed

Simon Hucko[1]*, Xiaosha Tao[1], Katharina Schmitz[1]

[1]*Institute for Fluid Power Drives and Systems, RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen*

* Corresponding author: Tel.: +49 241 80 47744; E-mail address: simon.hucko@ifas.rwth-aachen.de

ABSTRACT

This paper examines the challenges of real-time modeling solenoid valve actuators. Usually, more complex real-time models are executed as lumped parameter models, often simulatively parameterized, using finite element method (FEM) models. The quality of the simulative parameterization heavily depends on the accuracy of the FEM model. The accuracy, in turn, is largely determined by the material parameters used to create the FEM simulation. Variations within the material can render the primary material parameters inaccurate, thereby reducing the accuracy of the FEM model and consequently the derived parameters. The quality of the real-time model thus largely depends on the quality of the material parameters.

To eliminate uncertainties introduced by subpar material parameters and to enable precise real-time models, this work showcases the possibility to reconstruct required material parameters simulatively from flux measurements. The more accurate material data obtained using this new approach could also improve the accuracy when adapting components to new requirements. The paper describes a model with lumped parameters, as well as an FEM simulation model, a test rig, and the process of B(H)/initial magnetization curve calculation, followed by validation with new operating conditions.

Keywords: Model Based Development, Softsensor, Solenoid, Valve

1. INTRODUCTION

1.1. Motivation

For decades, solenoids have been widely used as valve actuators in fluid power technology, due to their low manufacturing costs, high force density, speed and efficiency. Growing demands on functionality and the quest for better control account for the need for more detailed modeling of actuators. Applications include model-based design, condition monitoring, and the acquisition of additional data through soft sensors [1]. However, most of these applications require an accurate and comprehensive real-time model of the considered system.

1.2. Challenge and Approach

As part of the modeling process, it is necessary to find numerical representations for various coupled effects of the electrical, magnetic and mechanical domains of the solenoid, as these have significant influence on system behavior. A particular challenge is posed by the nonlinearities of the electromagnetic subsystem. In general, the nonlinearities can be categorized into rate-independent effects, such as magnetic hysteresis, and rate-dependent effects, such as eddy currents or friction [2].

In order to enable broad application, it is necessary to not only create a suitable representation of the effects, but also a cost-effective parameterization. Additionally, the models must be computable on common microprocessors within the given cycle time. Magnetic hysteresis in particular is often neglected due to the time required to parameterize mathematical hysteresis models. There are several main options for electromagnetic modeling, such as the use of black-box models, FEM-driven models, or analytical models with lumped parameters. The latter option is far less complex than the others and can be used in control applications as, for example, in model-based valve controls [3–5]. The models used can be parameterized with the help of measurements as shown in [3]. Due to the required complexity of comprehensive data acquisition and processing, such models are increasingly parameterized with simulated data. Parameterization often relies on data from magnetostatic FEM simulation-models [6]. The accuracy of these FEM models is highly dependent on the material parameters, which are usually provided by the solenoid manufacturers or material suppliers. However, uncertainties in material parameters arise from steel processing as well as variations within production batches, making it difficult to estimate the accuracy of characteristics derived from FEM models [7]. An independent determination of the assembled magnet's material parameters or initial magnetization curve is thus advisable. However, even if all flux-carrying parts of the magnetic circuit are made of the same material, it is often not feasible to analytically determine the initial curve from the measured excitation current and the magnetic flux generated by the actuator coil.

The aim of this work is therefore to investigate the possibility of simulatively reconstructing the initial magnetization curve with flux measurements, which were recorded using the assembled solenoid. By implementing an optimization algorithm, taking into account physical constraints, the initial curve is adjusted in such a way that the simulated magnetic flux converges to the measured magnetic flux.

This novel approach is validated in multiple steps: First, different operating points are simulated with the newly determined material parameters using the FEM model. The results are then compared to measured data. Subsequently, parameters for a real-time capable actuator model are derived from the FEM model. The results of the actuator model using non optimized parameters and those generated using optimized parameters are in turn compared with measurement data.

1.3. Structure of the Paper

The present study is structured as follows: First, a lumped-parameter model of a valve actuator is described. Then, the test rig for the metrological investigation of magnetic actuators is presented. Next, the structure of an FEM simulation model of an exemplarily considered actuator is explained. This FEM model is then used in conjunction with the measurement results to calculate the initial curve and parameters for a real-time capable model are derived. Finally, the simulation is validated and compared to measured data.

2. MODELING OF VALVE SOLENOIDS

In this section, the design of a lumped-parameter and an FEM model is described. The actuator under consideration is a commercially available proportional solenoid which is used in hydraulic control valves of nominal size 6. The sectional view of the solenoid can be seen in **Figure 1**.

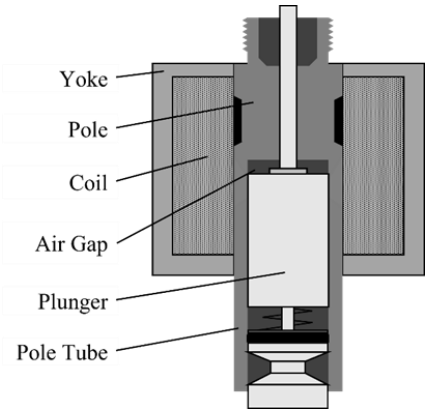


Figure 1: Sectional View of a Proportional Valve Solenoid

2.1. Lumped-Parameter Model

Lumped-parameter models are simplified mathematical representations in which the spatial distribution of the system is neglected. It is assumed that all components of the system are concentrated in one point. This reduces the partial differential equations describing the system to a topology or ordinary differential equations.

For modeling, the considered system is commonly divided into different domains. In case of solenoids, usually the electrical, the magnetic and the mechanical domains are taken into account. Each domain itself is represented by a network of discrete interconnected elements with concentrated parameters.

These elements are interconnected by system equations, enabling the simulation of the solenoid's response to various input signals.

The electrical system can be described by Equation 1, where U is the voltage applied to the coil, which is equal to the sum of the voltage across the winding resistor R and the induction voltage [8]. The magnetic system can be described by Equation 2. The magnetomotive force Θ constitutes the sum of the field-generating current i_f and the eddy current i_{eddy}. The field-generating current i_f corresponds to the product of the magnetic resistance R_m and Ψ. When modeling the eddy currents in a simplified way by assuming a constant flux distribution, i_{eddy} can be calculated from the eddy current parameter L_{eddy} and the induction voltage [8]. w equals the number of the coil's windings, i is the electric current, and Ψ the flux linkage. According to [8], the force generated by the solenoid depends on the flux linkage and on the position of the armature x, as described in Equation 3. For simplification, the generated force as well as the field-generating current are approximated by a characteristic map dependig on Ψ and x in the further course. For the calculation of the eddy currents, they are assumed to be proportional to the induction voltage.

$$U = R \cdot i + \frac{d\Psi}{dt} \tag{1}$$

$$\Theta = w^2 \cdot i = \Psi \cdot R_m + L_{eddy} \frac{d\Psi}{dt} \tag{2}$$

$$\mathrm{F_m} = \frac{\mathrm{d}}{\mathrm{dx}} \int_{\mathrm{I_0}}^{0} \Psi \, \mathrm{di} \tag{3}$$

By using these equations, the simple, computationally inexpensive solenoid model shown in **Figure 2** can be formulated.

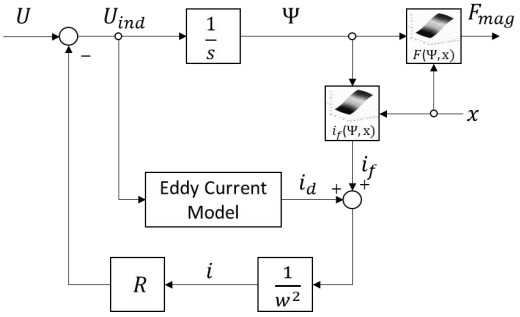


Figure 2: Lumped Network Solenoid Model.

2.2. Finite-Element-Modeling

Finite element modeling is a numerical simulation technique which divides the solenoid into small elements and calculates the electromagnetic field interactions within each element. FEM models provide a detailed and accurate representation of the solenoid's behavior, accounting for complex geometries and materials. They can simulate both static and dynamic responses of the solenoid, including factors like magnetic saturation and eddy current losses.

Solenoid valve actuators typically feature rotational axisymmetry, making 2D axisymmetric FEM simulations cost-effective. However, some plungers/armatures introduce non-axisymmetric elements. For example, in fluid-filled pole tubes, the armature is provided with round or rectangular channels. These prevent or reduce a pressure-induced counterforce when the armature moves. In the exemplarily considered magnet, the fluid is equalized by two holes in the plunger. Such non-rotationally symmetric elements can be transformed into an equivalent rotationally symmetric geometry. This requires the iron cross-section effective for the magnetic flux to remain the same [5, 8].

To allow for an automated multiple execution of the model in a manageable time, a parameterized 2D model of the considered magnet was created. The model should allow the variable adjustment of the armature displacement and the excitation current. **Figure 3** shows the 2D solenoid model and meshing.

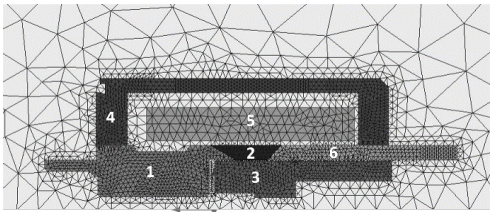


Figure 3: 2D FEM Model, 1. Pole, 2. Non-Magnetizable Area, 3. Plunger, 4. Yoke, 5. Coil, 6. Pole Tube.

The coil has an orthocyclic winding. In order to calculate the fill factor of the coil, the area of the conductor was divided by the area of the winding [8].

A variable air gap as shown in **Figure 1**, was implemented to account for the armature movement. In the area of the air gap, the automatically generated mesh was manually revised to increase precision. The non-magnetic area between pole and pole tube has been simulated as air.

The non-magnetic area between pole and pole tube as well as between pole and armature was simulated as air. A polytetrafluoroethylene sheet is used between the armature and the pole tube to reduce friction. To determine the parasitic air gap between the plunger and the pole tube, the diameter of the plunger and that of the pole tube were measured and subtracted from each other. The same has been done for the air gap between yoke and pole and between yoke and pole tube. Parts 1,3,4 and 6 are part of the flux-carrying iron circuit made of the same soft magnetic material. Material properties were added manually using data from manufacturers, material producers, or actual measurements.

The 2D model allows to calculate the force generated by the actuator under various current levels and armature strokes as well as the magnetic flux magnitude under different currents.

2.3. Nonlinear Material Characteristics

Soft magnetic materials are used for electromagnetic actuators. They are characterized by a narrow hysteresis curve and thus low hysteresis losses during magnetization.

The initial magnetization of the material follows the so-called initial curve depicted in **Figure 4**. The proportionality factor μ between the magnetic flux density B and the magnetic field strength H is called permeability. The permeability describes the slope of the B(H)curve and characterizes the influence of different materials on the magnetic field.

With the relative permeability μ_r, the absolute permeability can be expressed as a multiple of the induction constant μ_0. Beyond the saturation flux density B_S, the B(H) behaves linearly and approaches μ_0 [8].

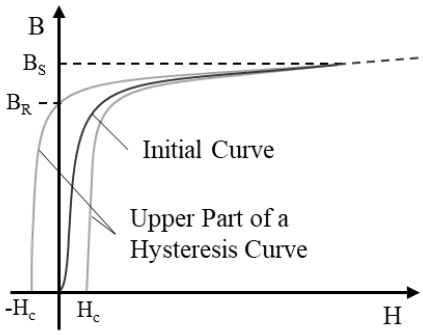


Figure 4: B(H) Characteristic Curve with Initial Curve.

3. TEST RIG

The test rig described below is used to investigate solenoids. For the present work, the magnetic flux and force generated by the solenoid were determined at different excitations and positions.

The setup shown in **Figure 5** and **Figure 6** was designed to allow for precise measurement of a wide variety of valve solenoids. **Figure 5** shows the mechanical part of the test rig.

On the left side, the test-solenoid is mounted on a rigid frame. A movable carriage driven by a servo motor allows for precise, dynamic positioning of a mechanical stop against which the armature presses. The force generated by the armature is recorded by a force sensor (MES-KM40). Both the position of the armature and the position of the carriage are measured to precisely control the armature position and to eliminate possible position deviations, for example due to backlash or the limited stiffness of the force sensor.

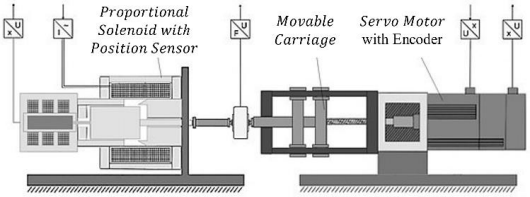


Figure 5: Schematic Depiction of the Mechanical Setup for Measuring Solenoids [1].

In **Figure 6**, the electrical part of the experimental setup is schematically shown. The circuit consists of the solenoid itself, an amplifier and a low-side measurement shunt (3 Ω ± 0.02%) for current measurement.

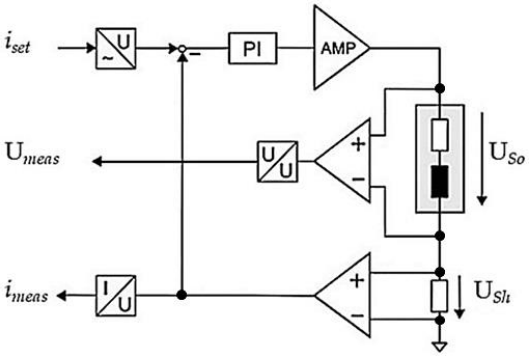


Figure 6: Electrical Schematic for Measuring Solenoids [1].

The amplifier (Kikusuiu PBZ80-5) allows, in push/pull configuration, the application of arbitrary currents and voltages. The solenoid is highlighted by a grey rectangle and is represented by an inductor and an ohmic resistor. The temperature change in the measurement shunt is small due to sufficient cooling. The resulting change in resistance is therefore negligible. The voltage is measured across the solenoid. The synchronized acquisition of voltage, current and force is carried out using a measuring amplifier (MC USB 404-60).

The flux is calculated according to Equation 4 using the measured voltage U and the current i.

$$\Psi(x,i) = \int_0^t (U - R \cdot i)\ dt \tag{4}$$

The flux curves determined in this way are shown in **Figure 7** for excitations of 0.5 A, 1 A and 1.5 A.

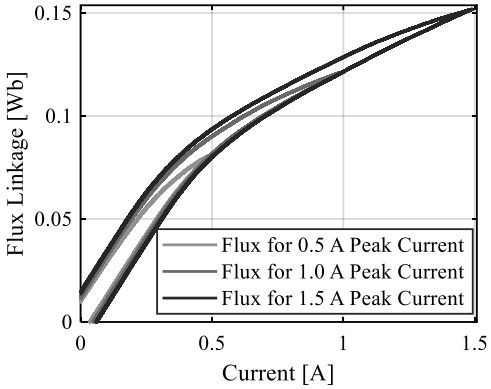


Figure 7: Measured Flux Characteristics of a Solenoid for a Triangular Excitation with a Peak Current of 0.5 A, 1.0 A and 1.5 A at a Stroke of 1.5 mm.

The measured force characteristics are shown in **Figure 8**. In order to avoid static friction at a constant current, the armature was moved quasi-statically between the stops using the servo motor shown in **Figure 5**. The position was detected using the actuator's own position sensor.

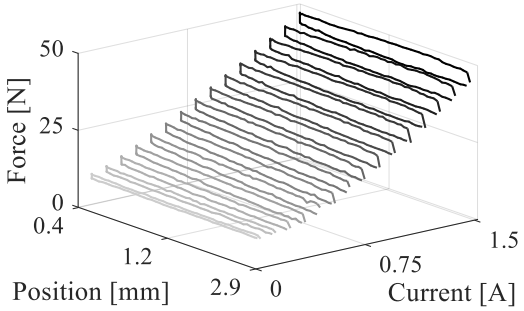


Figure 8: Measured Force Characteristics of a Solenoid.

The equally shaded lines on top of each other represent the force progression with ascending and descending directions of movement for each current step. The force hysteresis is clearly visible. The solenoid's air gap is reduced with decreasing voltage of the position sensor. Without using a dither signal, the force hysteresis for this solenoid is 1–5 N. However, modeling this relationship would exceed the scope of this paper. Therefore, the individually recorded hysteresis curves were averaged and interpolated, providing the magnetic force for a given current and position.

4. NEW METHOD FOR B/H CURVE ESTIMATION

To simulate magnetic actuators, common FE programs use initial magnetization curves (henceforth referred to as initial curve) up to saturation to account for nonlinear material characteristics.

The initial curve is practically identical with the commutation curve which can be constructed from completely recorded hysteresis curves [9]. Considering a sample with simple geometry and homogeneous flux density, the hysteresis curves and from them the initial curve can be determined by measuring the magnetic flux and the electric current using Equations 5 and 6.

$$\oint_l H\, dl = \sum I = w \cdot I \tag{5}$$

$$\Phi = \int_A B\ dA \tag{6}$$

Due to the complex geometry of a valve solenoid as well as the inconsistent flux density and field strength, the previously described procedure is not applicable. Even though the iron/magnetic circuit can be broken down into a number of simpler sub geometries, the respective magnetic field strengths and flux densities of these sub geometries are usually unknown. A simple calculation is therefore not feasible.

The new method presented in the following, enables the determination of an initial curve from measured flux/current curves by coupling an FEM program with an optimizer.

4.1. Idea of the Optimization Procedure

Using an initial curve from literature and a given valve solenoid geometry, a FEM simulation can be performed. This simulation calculates the magnetic flux for an entered excitation. Subsequently, the difference or error between the simulated magnetic flux and the measured flux is calculated as given in Equation 7, where n represents the number of measurement points or simulations.

$$e = \frac{1}{n} \cdot \sum_{i=1}^{n} \sqrt{\left(\Psi mess(i)\right)^2 - \left(\Psi sim(i)\right)^2} \tag{7}$$

If the accuracy of the measurement, the geometric dimensions and the quality of the meshing are appropriate, this error can be attributed to the initial curve used for simulation. Consequently, the error can be reduced by adjusting the initial curve. Hence, the main objective of this method is to reconstruct the initial curve of the material from measured magnetic fluxes. This reconstruction is done by iteratively performing static FEM simulations of the magnet. In each step, the calculated magnetic flux is compared to the measured one to calculate the error. With the help of an optimizer, the initial curve is then adjusted to reduce the error. This optimization is performed at a fixed armature position for different excitations.

4.2. Optimization

Figure 9 shows the conceptual structure of the optimization.

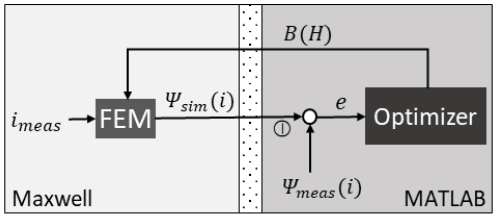


Figure 9: Block Diagram of the Optimization Procedure.

The algorithm used for the procedure forms a feedback loop which enables adaptive automated simulations through the integration of ANSYS Maxwell and MATLAB. As far as the authors are aware, there is no direct interface between MATLAB and ANSYS Maxwell allowing such an optimization. Therefore, a suitable interface was developed to facilitate the required data transfer.

In **Figure 9** the left side represents the simulations of the FEM model in ANSYS Maxwell, while the right side illustrates the optimization algorithm developed in MATLAB.

The initial curve, once optimized, is returned to the FEM simulation model in the form of a data table. Within the FEM simulation model, the new curve and equivalent measured excitation current i_{mess} are used to calculate new magnetic fluxes $\Psi_{sim}(i)$. Subsequently, $\Psi_{sim}(i)$ is passed back to MATLAB, where the error e is calculated. This process iterates automatically until the specified optimization conditions are met, for example by reaching a maximum number of computation cycles or by achieving the required accuracy.

In the following, the selection of the optimization algorithm is explained.

The definition of favorable boundary and initial conditions can help to minimize the number of necessary iteration steps. The chosen conditions, the resulting requirements for the algorithm and the choice of the algorithm are explained.

As depicted in **Figure 9**, the initial curve is the output parameter of the MATLAB module and serves as the optimization objective. The input parameter is the magnetic flux $\Psi_{sim}(i)$, derived from the FEM calculation. The error function to be optimized, is defined by Equation 7.

As defined in [8], the initial magnetization curve shown in **Figure 4** exhibits nonlinear, steadily increasing properties.

The curve has an inflection point, prior to which the permeability increases strictly monotonically. After the inflection point, it decreases strictly monotonically, until complete saturation. When exceeding the saturation flux density, the initial curve behaves linearly and approaches μ_0.

Based on these characteristics, constraints on the optimized curve are established.

Aside from introducing constraints, the number of iterations can be reduced by choosing suitable starting parameters. Typically, there is reference data available for the initial curve to be optimized, provided by literature, the manufacturer or material supplier.

To summarize, when selecting an algorithm to minimize the error function, the following aspects must be considered:

- Nonlinearity of the optimized parameter
- Possibility to define initial values
- Ability to set continuity conditions or create nonlinear constraints

Based on the given requirements, the augmented Lagrangian pattern search algorithm was selected. The pattern search method is an optimization algorithm which does not require derivatives of the objective function during computation. The augmented Lagrangian method is a mathematical optimization technique commonly used in pattern search algorithms. It addresses constrained optimization problems, where both equality and inequality constraints exist. It does so by introducing a penalty term into the objective function, which quantifies the violation of constraints and adjusts this penalty based on Lagrange multipliers. The method iteratively optimizes the augmented Lagrangian function, gradually reducing the constraint violations. The process allows the algorithm to search for optimal solutions while accounting for constraints.

4.3. Evaluation of Results

The aim of the optimization is to minimize the error between the measured and simulated fluxes. **Figure 10** shows the measured magnetic flux and the flux calculated with the FEM model for a piston stroke of 1.0 mm. After optimization, the two lines essentially coincide.

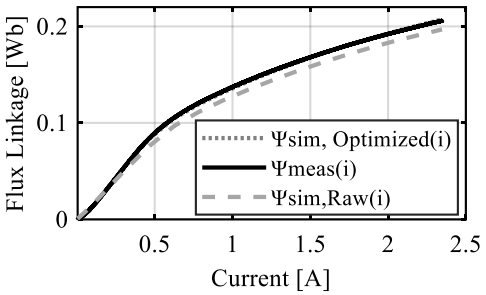


Figure 10: Comparison of Measured $\Psi_{\text{meas}}(\text{i})$, $\Psi_{\text{sim,Optimized}}(\text{i})$ with the Optimized and $\Psi_{sim,Raw}(\text{i})$ with the Unoptimized Initial Curve at 1.0 mm Stroke.

Figure 11 shows the initial curve of the material used. The dashed line is the original, non-optimized curve. It was obtained from the work of [7] and served as a starting point for the optimization.

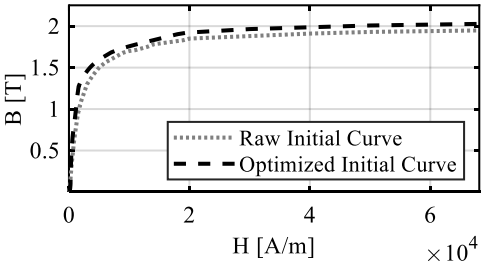


Figure 11: Unoptimized (Dotted) and Optimized (Dashed) Initial Magnetization Curve.

After about two thousand iteration steps, the initial curve, shown as a solid line in **Figure 11,** is obtained. It is evident, that saturation of the optimized curve occurs at higher field strengths.

5. EVALUATION OF SIMULATION-BASED PARAMETER ESTIMATION FOR REAL-TIME MODELS

To parameterize the model shown in **Section 2**, $F(\Psi, x)$, $i_f(\Psi, x)$ and L_{eddy} are determined with the aid of the FEM model and stored in the form of characteristic diagrams.

The relation $F(\Psi, x)$ is shown in **Figure 12**. It describes the force F generated depending on the flux linkage Ψ and the armature position x.

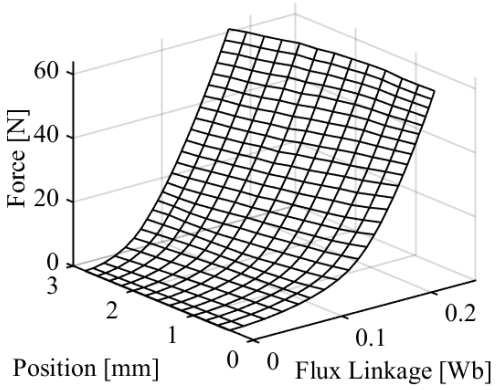


Figure 12: Force F Generated Depending on the Flux Linkage Ψ and the Armature Position x.

The relation $i_f(\Psi, x)$ is shown in **Figure 13**. The characteristic diagram describes the relationship between the field-generating current i_f, the flux linkage Ψ and the armature position x.

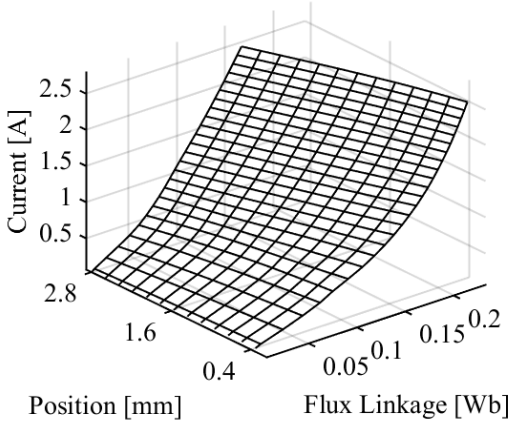


Figure 13: Field-Generating Current i_f, in Dependence of the Flux Linkage Ψ and the Armature Position x.

To evaluate occurring eddy currents, various signals were simulated with different current change rates.

As the focus of this work is on the parameterization of real-time capable models, extensive modeling of the eddy currents, for example using a shell model as shown in [8], is not carried out. Instead, the approach described in **Section 2** is used. A simplified linear dependence between the eddy currents and the induced voltage is assumed.

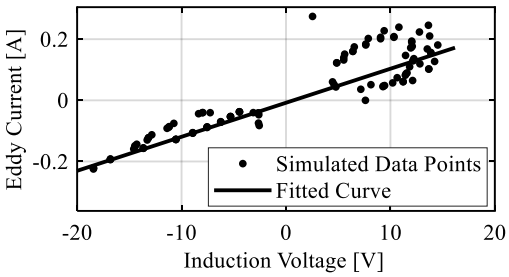


Figure 14: Fit of the Occurring Eddy Currents Depending on the Induction.

Figure 14 shows an example of the eddy currents occurring as a function of the induced voltage for an excitation with triangular current signals of different maximum heights and current change rates. In the investigations, current change rates to a maximum of 125A/s were considered.

As the rate of current change increases, eddy currents increase as well and require a more complex model to describe them. In **Figure 14**, this can be recognized by the fact that the eddy currents scatter further away from the linear regression line with increasing current change rates.

6. VALIDATION

The validation of the new approach for optimizing the initial curve is carried out in two steps. In the first step, the optimized material parameters are used to simulate operating points with the FEM simulation, which were not part of the optimization. The results are then compared with measured data. In the second step, the optimized parameters are used to investigate the dynamic behavior of the solenoid during a ramped excitation. The results of the simulations are compared with those obtained with non-optimized parameters.

First, the results of two FEM simulations with and without an optimized initial curve were compared. **Figure 15** illustrates the force versus stroke at a current of 1.5 A.

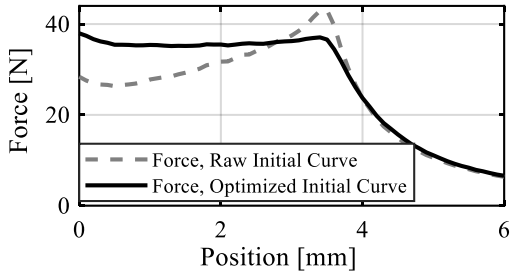


Figure 15: Magnetic Force as a Function of Stroke at a Current of 1.5 A.

The dashed line represents the force, calculated with an unoptimized initial curve. The force progression exhibits poor linearity in the working range of 0.5-2.8 mm. The solid line represents the force characteristic calculated using the optimized initial curve. It is evident, that the linearity within the operating range has been significantly improved after optimization, which is more consistent with the measured characteristics of the actuator.

For further validation, first, the simulated and measured magnetic flux, and subsequently the force is considered.

The magnetic flux at one armature position has been used to determine or optimize the initial curve.

The course and magnitude of the magnetic flux usually changes significantly depending on the armature stroke. To verify the optimized initial curve, the magnetic flux at other positions is simulated and then compared with measured data in **Figure 16**.

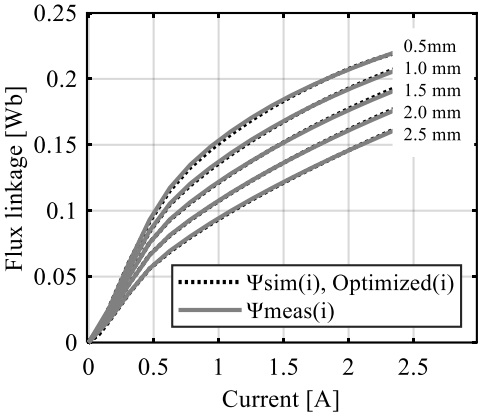


Figure 16: Measured and Simulated Flux using the FEM Simulation with Optimized Initial Curve as a Function of Current for different Strokes.

As can be seen in **Figure 16**, the simulated and measured fluxes are close to each other for all positions. The shape and curvature of the characteristics have been reproduced. The deviation between simulated and measured values increases with increasing excitation.

In Figure 17, the measured magnetic force is compared to the magnetic force calculated by the FEM simulation with an optimized initial curve.

The measurement of the axial force is performed as described in Section 3. The force is plotted at different excitations of 0.5 A, 1 A, 1.5 A, 2 A in the working range of 0.5 to 2.8 mm. The simulated force-stroke characteristics are depicted by marked lines. The measured ones by solid lines. The measured force curves show a hysteretic behavior typical for this kind of solenoid. It can be assumed that the hysteresis is caused by frictional forces acting with the same magnitude in both directions of movement. The force generated by the magnet can therefore be assumed to lie between the upper and lower hysteresis curves.

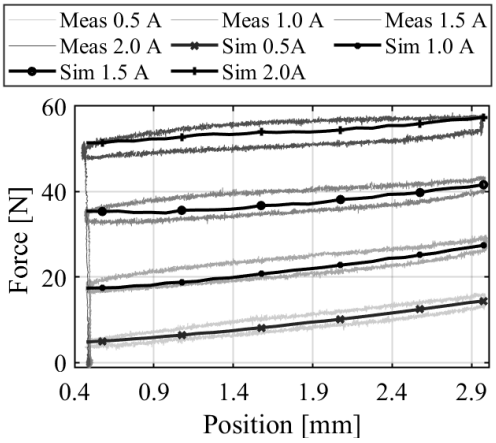


Figure 17: Measured Force and Simulated Force using the FEM Simulation with Optimized Initial Curve.

It should be taken into account that measurement errors in the flux measurement as well as an insufficient quality of the meshing lead to a distortion of the initial curve. These distortions would lead to deviations between simulated and measured force values at operating points not used to optimize the initial curve. In the present case, however, the simulation results agree well with the measurements for all operating points considered. Deviations that affect or offset the flux and force characteristics equally over the entire excitation in all positions cannot be identified by this validation method. If available, a measured initial curve could be used for this purpose.

For a final validation, the dynamic behavior respectively the force build-up of the solenoid during a ramp-shaped excitation with a current change rate of 10 A/s is investigated. For this purpose, the simulation results with and without an optimized initial curve, were compared to a measured force build-up in **Figure 18**.

The solid curve shown in **Figure 18** describes the simulated force with the unoptimized initial curve. The dashed curve describes the simulated force with optimized and the dotted curve the measured force.

The deviation of the simulation with non-optimized parameters is more than 20 N. The large deviation can be explained by the fact that the initial curve used was taken from generally available material data. Changes in the composition and the annealing process can have a significant effect on the magnetic properties and vary from manufacturer to manufacturer and sometimes from batch to batch.

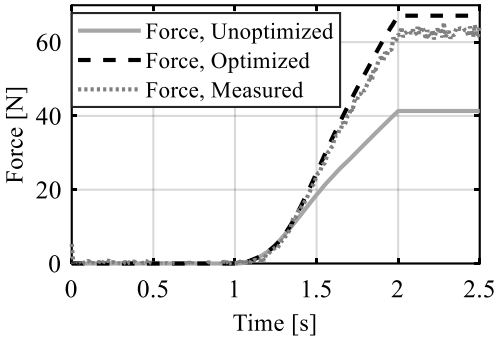


Figure 18: Measured and Simulated Force Build-up of the Solenoid With and Without Optimization.

Simulations based on measured initial curves of the unprocessed material show less pronounced but nevertheless often intolerable deviations as shown in [5].

It can be clearly seen that the simulated force curve with an optimized initial curve is closer to the measured force curve. The remaining deviation is smaller than 3 N and corresponds to the hysteresis of the force characteristics shown in **Figure 17**.

The magnetic and friction hysteresis has not been considered in the modeling. Comparing the measured and simulated force in **Figure 17**, the measured force is lower than the simulated force when the current increases due to the hysteresis. Thus, the simulated (optimized) force shown in **Figure 18** is greater than the measured force by half the hysteresis width. Apart from the expected deviation due to the non-observed hysteresis, the simulation represents the force build-up well.

Therefore, without the knowledge of exact material parameters, the presented method allowed not only to significantly improve the FEM, but also the real-time capable dynamic model.

7. CONCLUSION AND OUTLOOK

In the present work, the simulative parameterization of real-time capable solenoid models was investigated. For this purpose, the structure of the used lumped-parameter model, the FEM model as well as the test rigs for the metrological investigation were described. Subsequently, a new method was presented which makes it possible to determine the material parameters of the magnetic circuit required for the FEM simulation. As validation, the force generated by the magnet at different operating points was simulated with the newly determined material parameters and compared to the measurements. The simulated data agreed well with the measured data. The accuracy of the FEM simulation was significantly improved. With the aid of the improved FEM simulation, the parameters for a real-time model were then derived. Finally, the dynamic behavior respectively the force build-up of the solenoid was considered. It was found that the simulation with optimized parameters showed a smaller deviation from the measured force build-up than the simulation with unoptimized parameters. In planned future work, a hysteresis and an extended eddy current model will be presented that allows better representation for higher current change rates. In addition, the method presented here offers the future possibility of detecting deviations in the material parameters caused by production or variations between batches. These can be taken into account in the simulation of magnetic actuators to improve model accuracy. A more accurate simulative representation of the solenoid would also simplify the adaptation of existing solenoids.

In terms of real-time capable models, more accurate FEM models allow real-time capable solenoid models to be parameterized inexpensively, quickly, and with high accuracy. Such models could then be used for closed-loop control or in soft sensors as well as for condition or power monitoring applications.

NOMENCLATURE

U	Voltage
R	Resistance
i	Current
Ψ	Flux linkage
Θ	Magnetomotive force
w	Number of windings
R_m	Magnetic resistance
F	Force
x	Position
H	Field strength
B	Flux density
μ	Permeability

ACKNOWLEDGEMENTS

This research was funded by Forschungskuratorium Maschinenbau e.V. FKM, grant number 7052200.

REFERENCES

[1] S. Hucko, H. Krampe, and K. Schmitz, "Evaluation of a Soft Sensor Concept for Indirect Flow Rate Estimation in Solenoid-Operated Spool Valves," *Actuators*, vol. 12, no. 4, p. 148, 2023, doi: 10.3390/act12040148.

[2] S. Rosenbaum, *Entwurf elektromagnetischer Aktoren unter Berücksichtigung von Hysterese*. Zugl.: Ilmenau, Techn. Univ., Diss., 2011. Ilmenau: Univ.-Verl. Ilmenau, 2011. [Online]. Available: https://www.db-thueringen.de/receive/dbt_mods_00018835

[3] N. D. Vaughan and J. B. Gamble, "The Modeling and Simulation of a Proportional Solenoid Valve," *Journal of Dynamic Systems, Measurement, and Control*, vol. 118, no. 1, pp. 120–125, 1996, doi: 10.1115/1.2801131.

[4] S. S. Tørdal, A. Klausen, and M. K. Bak, "Experimental System Identification and Black Box Modeling of Hydraulic Directional Control Valve," *MIC*, vol. 36, no. 4, pp. 225–235, 2015, doi: 10.4173/mic.2015.4.3.

[5] Albert Wolfgang Schultz, "Simulationsgestützter Entwurf elektromagnetischer Linearaktoren für fluidtechnische Ventile," p. 119. [Online]. Available: https://d-nb.info/988163764/34

[6] C. Krimpmann, A. Makarow, T. Bertram, I. Glowatzky, G. Schoppel, and H. Lausch, "Simulationsgestützte Optimierung von Gleitzustandsreglern für hydraulische Wegeventile," *at - Automatisierungstechnik*, vol. 64, no. 6, 2016, doi: 10.1515/auto-2016-0017.

[7] R. Keilig, *Entwurf von schnellschaltenden (hochdynamischen) neutralen Elektromagnetsystemen*, 2007. [Online]. Available: https://www.db-thueringen.de/receive/dbt_mods_00010741

[8] E. Kallenbach, R. Eick, T. Ströhla, K. Feindt, M. Kallenbach, and O. Radler, *Elektromagnete: Grundlagen, Berechnung, Entwurf und Anwendung*, 5th ed. Wiesbaden: Vieweg, 2018. [Online]. Available: https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5214990

[9] R. Boll, *Weichmagnetische Werkstoffe: Einführung in den Magnetismus, VAC-Werkstoffe und ihre Anwendungen*, 4th ed. Berlin, München: Siemens-Aktiengesellschaft, 1990.