Simulation-Based Data Sampling for Condition Monitoring of Fluid Power Drives

Faried Makansi and Katharina Schmitz
1Institute for Fluid Power Drives and Systems (ifas), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany

Faried.Makansi@ifas.rwth-aachen.de

Abstract. Machine learning techniques are continuously gaining attention and importance in several technical domains. In the field of engineering, they can potentially provide manifold advantages for condition monitoring. However, availability of extensive operation data is a limiting factor. In this contribution, a simulation-based approach is presented, which allows an efficient generation of training data. Based on a lumped parameter simulation, a database of time-series data is generated for a hydraulic reference system. In order to incorporate states of faulty machine operation in the database, means to model component faults in the simulation are assessed. Further, a procedure for an automated training data generation is presented.

1. Introduction

Reliability and operational safety are crucial requirements for production machines and automation plants. Mostly, a high machine utilization is demanded to fulfill a required throughput and guarantee a firm return of investment. Monitoring the health status of such machines allows detecting looming failures at an early stage and therefore helps to prevent unplanned downtimes and optimize maintenance strategies. Data-driven techniques, often summarized by the term machine learning (ML), potentially provide manifold advantages in the area of condition monitoring (CM) compared to established approaches. However, the practical implementation of ML-based CM systems can often be time-consuming and costly. Major challenges are the availability of extensive data on regular and faulty machine operation as well as the selection of suitable data features [1].

One approach to circumvent the issue of costly and safety critical data acquisition in real world applications, is to generate training data by means of a lumped parameter simulation. Such simulations are often implemented during the engineering process anyways and are used to investigate a system’s behavior in a few specific design points. However, training data for CM has to represent irregular, faulty system behavior as well, which leads to a vastly increased number of variations of the simulation model compared to the conventional use of the simulation. This can be managed by means of a variational study, where the simulation model is run with multiple sets of design parameters in an automated manner. Hence, if parametric fault models are included into the system simulation, this can be leveraged to automate the data generation procedure.

In this contribution, the process of simulation-based data generation is investigated on a hydraulic reference system. Fault related model parameters are identified and reasonable upper and lower boundaries for their variation are defined. Moreover, a procedure for an automated data generation is presented.
The following section reveals the state of the art of CM in fluid power applications and simulation-based data sampling. Subsequently, the considered hydraulic system and its simulation model are presented in section 3. In section 4, options for the simulation of faulty system behavior are investigated while in section 5 the procedure for data generation is described. Finally, the results are summarized and a conclusion is given.

2. State of the Art

2.1. Condition Monitoring in Fluid Power Systems

In the field of fluid power systems, numerous approaches have been investigated and applied for CM. Solutions range from conventional limit checking, to the comparison of observed and expected system behavior based on a mathematical process-model [2], [3], [4]. However, these approaches either require high manual implementation effort and process knowledge, or have low applicability in cases of complex multi-signal non-linear systems. Concurrently, data-driven methods have also frequently been subject of investigation for CM in fluid power applications. As in [5], [6] and [7], numerous publications in this field focus on the basic proof of applicability of ML-methods for CM of fluid power systems and mainly consider situations of single or few concurrent faults occurring at the same time. Accordingly, only small data sets are used in these studies and questions on the sampling strategy and the distribution of the sampled data are not focused. In [8], datasets in the scale of 3000 measurements are conducted on a test bench for data-based CM of an axial piston pump. Six different isolated faults are investigated which are either varied on binary levels or equally spaced within a defined design region. However, the high number of conducted measurements is a result of repetitions of individual fault conditions. A similar setup is found in [9]. Furthermore, Helwig studies data-based CM techniques on a system level, where sensors are distributed over different locations in the system. If sensor data is combined, higher level information can be provided from which component faults can be inferred as well [10]. Similar to other referenced publications, the sampling strategy for the variation of fault levels and the consideration of concurrent faults is defined heuristically and is practically limited as a consequence of the data collection on a physical test bench.

2.2. Simulation-based data sampling

An alternative to the data collection on a physical test bench is the generation of operation data from computer simulations. Especially for the task of CM this can be beneficial, as safety critical fault settings can be emulated without the risk of damages of the test equipment or harm of the operator. Moreover, it allows to efficiently generate data on a larger scale, in continuous fault settings and in various operating conditions [11]. This links to the statements in [1] and [12], that a major reason for the slow adaptation of ML-based CM techniques in industrial applications is not the lack of in-service data in general, but more the lack of comprehensive data, which covers a diversity of operation scenarios.

In [13], the training of virtual robots is realized with data, which is generated by means of a computer simulation. High fidelity simulation models are used to account for the gap between simulated and real system behavior. Tercan et al. successfully implement a performance estimation model for an injection molding machine, by pretraining an estimation model with data from a simulation and retraining the obtained model with real world data [14]. In the field of condition monitoring, simulation-based data generation is used for the estimation of pump leakage in [15]. Although only a single fault is considered in this study, it is stated that a relatively large number of training points is required, in order to achieve a high accuracy with the simulation-based approach.

By conducting computer simulations instead of measurements on a physical test bench, also different approaches for the experimental design can be considered. A group of experimental designs that are suited for computer experiments are space filling designs (SFDs). In general, these designs aim to spread the sampling points evenly over the experimental region with minimal gaps between them [11]. One type of SFDs are Quasi-Monte-Carlo Designs, which uniformly distribute the
sampling points across the experimental region, based on a randomization scheme derived from pseudo-random sequences. Other commonly used SFDs are Latin Hypercube Designs (LHDs). The basic idea of a LHD is to divide each design factor into an equal number of levels and then distribute each factor level uniformly among the samples, such that each factor level is only sampled once. As by this rule, good space filling properties are not guaranteed, several extensions of this design can be found in the literature. Most of these extensions aim to improve the space filling properties by considering distance measures between samples or by ensuring orthogonality of the samples [16]. Furthermore, an advantage of LHD designs is that they allow to incorporate arbitrary distributions for different factors by transforming the initially uniform distribution to the desired one.

3. Simulation of the Hydraulic Reference System
In this publication, a hydraulic system consisting of a cylinder drive is considered, as it is commonly used in various applications in stationary hydraulics. The simulation model is built according to a physical demonstrator machine, which is a test bench of a hydraulic press. The main components are a hydraulic power supply, a proportional directional control valve and a single acting hydraulic cylinder. The system is modelled in the Modelica-based simulation tool SimulationX, which allows the 1D-simulation of hydraulic circuits. The simulation consists of coupled component models, each containing information on their input and output variables, as well as equations describing their physical behavior. In figure 1, the simulation model of the considered system is depicted, while figure 2 shows the demonstrator.

![Simulation model of the considered hydraulic system](image1)

Figure 1. Simulation model of the considered hydraulic system. The function of the press is to move a ram to a defined position. The ram is attached to the cylinder rod and acts against a load force, which represents the reaction force of a workpiece. In the simulation,
the working cycle is modelled accordingly. To achieve a closed-loop position control of the cylinder, its actual measured position is fed back, compared to the target position and directed to the controller to finally manipulate the position of the valve spool. Figure 3 shows the progressions of the cylinder’s position, velocity and load force over one working cycle. The working cycle starts in rapid traverse where the ram of the press moves towards the workpiece without external load and limitation of the velocity. As the ram reaches a defined distance to the workpiece, the velocity is restricted through limiting the control signal and the volumetric flow to the cylinder, respectively. After the ram is in contact with the workpiece, the load force increases linearly with the stroke of the cylinder. The cycle is defined such that the controller is given a fixed period to reach the set value, before the backstroke is initiated in open-loop control.

Figure 2. Demonstrator of the simulated hydraulic press.

Figure 3. Working cycle of the hydraulic press.

Basic parameters of the component models, such as nominal dimensions, are set according to catalogue data of the components from the demonstrator. The hydraulic power supply provides around 18 l/min of volumetric flow at a pressure of up to 220 bar. Then, the fluid is controlled by a proportional 4/3-way valve with a nominal volumetric flow of 12 l/min. The piston diameter of the hydraulic cylinder is 80 mm with a piston rod diameter of 45 mm and a maximal stroke of 400 mm.

4. Fault Modelling and Fault Effects
In order to detect and diagnose a system’s condition as part of a data-driven CM, the database used, has to include regular and faulty conditions of the system likewise. Therefore, faulty operations of the hydraulic system have to be emulated in the computer simulation as well. This can be done by varying parameters of the component models, which correspond to specific faults. The resulting system behavior of the hydraulic circuit then reflects the effects of these faults. Opposed to conventional process-model-based CM, a high fidelity model of the system behavior under fault conditions is not targeted. Hence, it should reveal general effects of component faults and fault combinations represented in a database to qualitatively identify root-cause dependencies.
4.1. Fault Modelling

In the following, considered faults are leakages and increased friction in the hydraulic cylinder, a faulty measurement of the cylinder position, worn control edges of the control valve, as well as external leakage at a connector in the hydraulic power supply.

In the hydraulic cylinder, leakages mainly occur at sealing spots between elements with relative motion. This is prevalent between piston and cylinder housing as well as between piston rod and cylinder housing, as depicted in figure 4. The clearance between piston and cylinder housing can cause a leakage flow between the two chambers of the cylinder and is therefore called internal leakage. Whereas a leakage flow between piston rod and housing is considered an external leakage, as it is emitted to the environment. In the simulation software used in this study, the internal leakage can be modelled as fluid flow through a gap. It is obtained from the superposition of a fluid flow driven by pressure difference across the gap and a kinematically induced transmission of fluid between the two sides of the gap. For a circular gap with concentric positioning of piston and cylinder, the internal leakage flow $Q_{ci}$ can be described by equation (1).

$$Q_{ci} = \frac{1}{2} \cdot \dot{x}_p \cdot h_1 \cdot D_p \cdot \pi + \frac{h_1^3 \cdot D_p \cdot \pi}{12 \cdot \eta \cdot l_1} \cdot \Delta p$$  \hspace{1cm} (1)

Where $\dot{x}_p$ denotes the relative velocity between piston and cylinder, $h_1$ the gap height, $D_p$ the piston diameter, $\eta$ the dynamic viscosity of the fluid, $l_1$ the length of the gap and $\Delta p$ the pressure difference over the gap. In practice, an increase of the internal leakage can occur as a result of a worn or damaged piston seal which in turn is linked to a change in the gap geometry. Thus, the gap height $h_1$ is considered the main geometric parameter impacting the internal leakage and therefore is used to emulate different levels of internal leakage in the model of the hydraulic cylinder.

Similarly, the external leakage at the hydraulic cylinder can be modelled as fluid flow through a circular gap. In contrast to the sealing system at the piston, the leakage flow is assumed to only be directed from the rod chamber to the environment as depicted in figure 4. Using the relation from equation (1) would therefore require a case separation, depending on the direction of the cylinder movement. In order to maintain model simplicity, and as the major part of the leakage is mostly induced by the pressure difference, the contribution of the relative motion is neglected here. Thus, the external leakage flow $Q_{ce}$ at the hydraulic cylinder is modelled as described in equation (2) with the gap height $h_e$ as fault related parameter.

$$Q_{ce} = \frac{h_e^3 \cdot D_p \cdot \pi}{12 \cdot \eta \cdot l_e} \cdot \Delta p$$  \hspace{1cm} (2)

![Figure 4. Sealing systems in a hydraulic cylinder.](image-url)
Another fault that can result from damaged sealing elements or a bad alignment of piston rod and cylinder housing, is an increased friction between moving components of the hydraulic cylinder. Generally, friction is modelled by means of a Stribeck Curve, which can provide valid estimates of the friction force across different operating conditions and friction regimes. However, the parametrization of a Stribeck Curve requires detailed measurements of the considered system which raises issues of practicability. Alternatively, the friction in the hydraulic cylinder is modelled with constant static friction force and dynamic friction force, respectively. In order to reduce the number of fault combinations in the following data generation process, the static friction force is solely used for the indication and variation of the fault state while the ratio between the two friction forces is kept constant. Based on data from [2] and [18], the ratio is set to 2.5 as expressed in equation (3).

\[
F_{Fr} = \begin{cases}
F_{static}, & \text{for } \dot{x}_P = 0 \\
2.5 \cdot F_{static}, & \text{else}
\end{cases}
\quad (3)
\]

Furthermore, worn control edges of the spool of a control valve are considered for the simulation of a faulty machine state. As a result of abrasion and erosion, the typically precisely manufactured control edges of a valve spool and sleeve can change their shape, especially over long-time usage. This can lead to a change in the relation between the spool position and flow characteristics and consequently the overall control behavior. However, without empirical investigation, this shift of characteristics is unknown and has to be modelled by other means. Consequently, the overlap of the valve is selected as fault defining parameter, such that a growing negative overlap correlates to an increased level of wear.

An external leakage in a connector in the hydraulic power supply is considered as well. It is modelled as a linear function of the system pressure and is varied through the proportionality factor of the leakage flow and system pressure, here denoted as leakage factor \(C_L\) in equation (4).

\[
Q_L = C_L \cdot P_S 
\quad (4)
\]

Finally, a faulty position sensor at the cylinder is modelled by adding a static offset to the sensed cylinder position.

4.2. Fault levels and ranges
The considered faults, their corresponding parameters in the component models and their value ranges for regular and faulty operation are summarized in table 1.

| Fault                              | Model Parameter | Nominal Value | Range            |
|------------------------------------|-----------------|---------------|------------------|
| Cylinder, Internal Leakage         | Gap height \(h_l\) | 1 \(\mu m\)   | 1 \(\mu m\) – 100 \(\mu m\) |
| Cylinder, External Leakage         | Gap height \(h_e\) | 0.5 \(\mu m\) | 0.5 \(\mu m\) – 75 \(\mu m\) |
| Cylinder, Friction                 | Static Friction Force \(F_{static}\) | 500 N          | 500 N – 2500 N   |
| Cylinder, Position Sensor Offset   | Offset          | 0 mm          | 0 mm – 410 mm    |
| Proportional Valve, Worn Control Edges | Spool Overlap \(y_0\) | -2 %          | -2 % – -4 %      |
| Hydraulic Power Supply, External Leakage | Leakage Factor \(C_L\) | 0 l/(min bar) | 0 – 0.1 l/(min bar) |

The nominal gap heights for the cylinder Leakages are selected according to data from [19] and [20]. The range for the leakage factor \(C_L\) is defined, such that states of no leakage, a leakage of a few drops per minute up to a broken connector can be emulated. The absolute leakage flows that result from the
selected ranges are listed in table 2 and refer to a peak differential pressure of 150 bar for the cylinder leakages and 200 bar for the leakage in the power supply.

Table 2. Resulting leakage flow rates for considered fault ranges

| Fault                              | Parameter Range | Leakage Flow |
|------------------------------------|-----------------|--------------|
| Cylinder, Internal Leakage         | 1 µm – 100 µm   | 0 – 23 l/min |
| Cylinder, External Leakage         | 0.5 µm – 75 µm  | 0 – 5.5 l/min|
| Hydraulic Power Supply, External Leakage | 0 – 0.1 l/(min·bar) | 0 – 20 l/min|

It is apparent, that the maximal leakage flows are not probable to be commonly observed in practice. However, these extreme cases are included in the database, to assure that the mapping of the system behavior to fault effects is valid for a broad range of operating conditions. This also applies to the other faults, where maximal values are heuristically defined, as no empirical base is available for estimation.

4.3. Fault Effects

A change of the previously derived fault parameters impacts the input-output-behavior of the hydraulic system. Hence, the fault effect is the change of operational characteristics which is caused by the change of a fault parameter. In order to outline options for an algorithmic mapping of faults and fault effects, exemplary cases are briefly illustrated in the following. For this, the pressure difference at the cylinder ports and the volumetric flow at the valve inlet are each compared at faulty and fault free settings.

Examples for the effect of internal leakage at the hydraulic cylinder are depicted in figure 5. The diagram on the left shows that with changing gap height $h_i$, the pressure difference between cylinder ports clearly changes at the beginning and end of the working cycle. Moreover, the internal leakage effects the volumetric flow throughout the working cycle, as shown in the diagram on the right. Especially in segments of the cycle where the volumetric flow is small, the relative change is apparent.

Figure 5. Effect of different settings of gap height $h_i$ on the pressure difference over cylinder ports (left) and the volumetric flow at the valve inlet (right) over one working cycle.

Similarly, the severity of the effect of an increased friction varies over a working cycle. As the left diagram in figure 6 shows, an effect is mainly observable around the time stamp of 10 seconds, where the cylinder velocity is low and the position is finely controlled. Additionally, the diagram on the right
hand side of figure 6 shows the effect of a change in valve spool overlap on the volumetric flow through the inlet port of the valve. In contrast to the fault effect of an internal leakage, the peak value is not affected by the change in overlap, nor is the volumetric flow at the beginning and end of the cycle.

As a conclusion, fault effects are reflected in system outputs that are commonly observable through sensor data. Providing a large number of training examples and describing signal characteristics with quantitative measures, moreover facilitates the use of ML methods for the mapping of system outputs to fault states.

![figure 6](image)

**Figure 6.** Effect of different settings of static friction force on the pressure difference over cylinder ports (left) and effect of different settings of valve spool overlap on the volumetric flow at the valve inlet (right) over one working cycle.

5. Automated Data Generation

In this section, the procedure of generating data from the simulation model is described. The general scheme is shown in figure 7.

![figure 7](image)

**Figure 7.** Procedure of automated simulation-based database generation.

The first step of this procedure is the setup of the experimental design, which is implemented in a Python-program based on the code-library *pyDOE2*. Here, a Latin Hypercube Design is selected as experimental design. Fault levels are uniformly distributed according the *maximin*-criterion and the number of samples is set to 10000, meaning that the same number of parameter sets is obtained, with which the simulation model is to be run. Provided the parameter identifiers and value ranges, the test matrix is generated and written to a simple text file. In addition to the fault related parameters
discussed in section 4.1, the load force acting on the system and the target position of the hydraulic cylinder are included as variable parameters. Subsequently, the simulation model is exported as an executable file (exe-file) via the code export function of the simulation program. In this step, state variables to be output by the model are also selected. Sensor signals for different pressures, the volumetric flow entering the control valve, as well as cylinder position, valve spool position and control signals are chosen here. The signals are sampled with a rate of 1000 Hz. Using the SRA-add-in of SimulationX, the exported exe-file can additionally be linked to the previously defined experimental design. Thereby, the simulation is automatically run with all parameter sets listed in the test matrix once the exe-file is executed. Moreover, running the simulations from an exported exe-file allows parallel execution of simulation runs on multiple processor cores, which additionally increases the efficiency of the procedure.

The procedure is finally run on four processor cores of an Intel Xeon Gold 6130 processor. Using the CVODE solver, the 10000 simulation runs are completed in approximately six minutes. This includes the writing and saving of the generated data to roughly 50 Gigabytes of text files.

6. Conclusion and Outlook
In this paper, a method for simulation-based data generation was presented, which allows to efficiently generate training data for a data-driven condition monitoring. The data was generated by means of a variational study in a lumped parameter simulation program. Options for parametric fault modelling were described for a hydraulic reference system. Furthermore, a procedure has been derived, which allows to generate large datasets of time-series data with a high variation of fault levels and operating conditions.

As a next step, the derived fault models are to be empirically assessed on the physical demonstrator of the reference system. Furthermore, the preliminary experimental design used in this study can be extended. The consideration of fault specific distributions in the Latin Hypercube Design, is expected to help focus the samples in design regions, which are of higher relevance regarding the separability of fault states and their probability of occurrence. Finally, the applicability of the obtained database is to be evaluated in the frame of a data-driven CM-setup.

Acknowledgements
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.

References
[1] Döbel I, Leis M, Molina Vogelsang M, Neustroev D, Petzka H, Rüping S, Voss A, Wegele M and Welz J 2018 Maschinelles Lernen – Kompetenzen, Anwendungen und Forschungsbedarf: Projektbericht
[2] Stammen C 2005 Condition Monitoring für intelligente hydraulische Linearantriebe (Aachen: Shaker)
[3] Münchhof M 2007 at – Automatisierungstechnik 55
[4] Isermann R 2011 Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems (Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg)
[5] Ramdén T, Krus P and Palmberg J-O 1996 Proceedings of the JFPS International Symposium on Fluid Power 1996 567–72
[6] Watton J and Freebody N 2000 International Journal of Fluid Power 1 59–67
[7] Zachrison A 2008 Fluid Power Applications Using Self-Organising Maps in Condition Monitoring
[8] Fischer A 2013 Prognosekonzepte zur Reduktion von Betriebskosten im Lufttransport (ProReB): Abschlussbericht

[9] Adams S, Beling P, Farinholt K, Brown N, Polter S and Dong Q 2016 Condition Based Monitoring for a Hydraulic Actuator

[10] Helwig NJ 2019 Zustandsbewertung industrieller Prozesse mittels multivariater Sensordatenanalyse am Beispiel hydraulischer und elektromechanischer Antriebssysteme (Aktuelle Berichte aus der Mikrosystemtechnik vol 30) (Aachen: Shaker)

[11] Siebertz K, van Bebber D and Hochkirchen T 2017 Statistische Versuchsplanung: Design of Experiments (DoE) (VDI-Buch) 2nd edn (Berlin, Heidelberg: Vieweg)

[12] Paulmann G and Mkadara G 2018 Condition monitoring of hydraulic pumps – lessons learnt Fluid power networks 11th International Fluid Power Conference (Aachen, Germany, 19th - 21st March 2018) ed H Murrenhoff (Aachen)

[13] El-Shamouty M, Kleeberger K, Lämml A and Huber M 2019 tm - Technisches Messen 86 673–84

[14] Tercan H, Guajardo A, Heinisch J, Thiele T, Hopmann C and Meisen T 2018 Procédia CIRP 72 185–90

[15] Wiens T and Fernandes J 2020 Application of Data Reduction Techniques to Dynamic Condition Monitoring of an Axial Piston Pump Proceedings of the ASME/BATH Symposium on Fluid Power and Motion Control -2019- ASME/BATH Symposium on Fluid Power and Motion Control (Longboat Key, Florida, USA, October 7-9, 2019) (New York, N.Y.: The American Society of Mechanical Engineers)

[16] Gramacy R B 2020 Surrogates: Gaussian process modeling, design, and optimization for the applied sciences (Chapman & Hall/CRC texts in statistical science series) (Boca Raton, FL: CRC Press, Taylor & Francis Group)

[17] Schmitz K 2018 Grundlagen der Fluidtechnik: Teil 1: Hydraulik (Reihe Fluidtechnik. U) (Aachen: Shaker)

[18] Abbas Kazemi-Moghaddam 2000 Fehlerfrühidentifikation und -diagnose eines elektrohydraulischen Linearantriebssystems (Darmstadt: Technische Universität Darmstadt)

[19] Mielke T 2020 Wassereinzug über ölhydraulische Stangendichtsysteme (Reihe Fluidtechnik D vol 98) (Düren: Shaker)

[20] Müller H K and Nau B 2020 Hydraulidichtungen http://www.fachwissen-dichtungstechnik.de/PDFs_01_2020_mit_PW/fw-dt_Kapitel05_01_20.pdf (accessed 30 Nov 2020)