Data-based condition monitoring of a fluid power system with varying oil parameters

Dipl.-Ing. Nikolai Helwig
Centre for Mechatronics and Automation Technology (ZeMA), Eschberger Weg 46, 66121 Saarbruecken, Germany, E-mail: n.helwig@zema.de

Professor Dr. rer. nat. Andreas Schütze
Lab for Measurement Technology, Dept. of Mechatronics Engineering, Saarland University, 66123 Saarbruecken, Germany, E-Mail: schuetze@lmt.uni-saarland.de

Abstract
In this work, an automated statistical approach for the condition monitoring of a fluid power system based on a process sensor network is presented. In a multistep process, raw sensor data are processed by feature extraction, selection and dimensional reduction and finally mapped to discriminant functions which allow the detection and quantification of fault conditions. Experimentally obtained training data are used to evaluate the impact of temperature and different aeration levels of the hydraulic fluid on the detection of pump leakage and a degraded directional valve switching behavior. Furthermore, a robust detection of the loading state of the installed filter element and an estimation of the particle contamination level is proposed based on the same analysis concept.

KEYWORDS: Condition monitoring, multivariate statistics, aeration, contamination

1. Introduction
In recent decades, the condition based maintenance strategy has widely arrived in industrial and mobile applications to realize a resource- and time-efficient operation with significantly reduced machine downtimes. Consequently, the demand for specialized condition monitoring systems has increased steadily /1/. On the other hand, a variety of process sensors is usually installed to control and monitor industrial machine processes anyway. In previous publications, we have shown that a quantification of typical wear and fault mechanisms of hydraulic components can be realized based on the operation-specific statistical analysis of process sensor data thereby reducing or even avoiding the need for specialized condition monitoring /2, 3/. However, in practical applications, oil temperature changes as disturbance variable
drastically impede the condition classification as described in /4/ for the example of the hydraulic circuit of a wind turbine. Besides temperature, another important parameter is the level of undissolved air in oil with an increased value tending to induce vibrations, and, in the long-term, an increased wear of components due to cavitation erosion and an accelerated aging of the fluid /5/. The most common cause for damages and sudden failures of hydraulic systems, however, is solid particle contamination, which can lead to a self-reinforced damage mechanism in case of inadequate filtration /5/. In the following paper, the influence of these oil parameters on the detection of component fault conditions is studied as well as the detection of the oil parameters itself based on the multivariate analysis of the process sensors installed in a hydraulic system.

2. Data analysis concept and experimental setup

2.1. Sensor data analysis

The signal processing scheme is based on the data of a process sensor net, i.e. process-synchronized sensor signals which are used for supervised offline learning with known target conditions (figure 1).

![Figure 1: Sensor data analysis concept](image)

The automated analysis can be divided into the steps feature extraction and selection, dimensional reduction, and classification. Furthermore, the generated statistical model has to be evaluated thoroughly to ensure reliable classification and to rule out undesired effects such as overfitting /6/.
For feature extraction, the raw signals collected for each cycle are split into several time intervals within the cycle corresponding to the characteristic phases of the working cycle, e.g. constant and transient load phases or directional valve operations. These raw sensor signal time intervals are then used for the computation of secondary features. The feature functions used here are the median and the three statistical moments variance, skewness, and kurtosis as well as features that describe the signal shape (e.g. linear fit of load ramp and rise time). To obtain a large feature pool, these features are generated from all process sensors available in the hydraulic system. The physical sensors are complemented with virtual sensors based on a mathematical model measuring the system efficiency (SE, ratio of hydraulic power to electrical power), cooling efficiency (CE, ratio of actual oil temperature decrease passing the cooler to maximum possible temperature decrease concerning the ambient temperature), and the cooling power (CP) of the system cooler based on heat transfer rate. Some system conditions do not show their symptoms in a short time scale using cycle-wise features but instead require analysis over an enlarged time window. For this purpose, time-series features are extracted over 2 up to 60 cycles again using statistical parameters (median, variance, skewness, kurtosis) to describe the long-term behavior of the previously calculated cycle-wise features.

To select the most significant features, all features are sorted according to the absolute value of pairwise Spearman's rank correlation coefficient $\rho$ calculated for each feature value and the targeted system conditions, i.e. oil parameter, over time.

After selection of the 20 up to 50 highest correlated features for each target condition, the dimension of the feature vector is further reduced by linear discriminant analysis (LDA) which optimizes the class separation by maximizing the between class variance and minimizing the within class variance /6, 7/. This is realized by a linear projection of the multidimensional feature vector along the directions with maximum class separability called discriminant functions (DF). The final classification is then based on calibration or training data for the 1st and 2nd DF using the Mahalanobis distance classifier /7/. For the evaluation of the statistical model, k-fold cross validation is used by partitioning the available data samples into k groups where k-1 groups are used for training and the remaining group for testing. This is repeated for all k sample groups being tested using the average classification rate as benchmark parameter. In order to allow a comparison of statistical models with classification rates of 100 %, we introduce a further criterion called scatter ratio which includes the mean ratio of the double standard deviation sum of the two involved classes to the Euclidian distance of their centroids analyzed for all class combinations:
\[ \text{scatter ratio} = \frac{1}{N} \sum_i N \frac{2\sigma_{a,i} + 2\sigma_{b,i}}{\| c_{a,i} - c_{b,i} \|_2} \]  

(1)

Here, \( N \) is the number of pairwise class combinations with the classes a and b of combination i, \( c_{a,i} \) and \( c_{b,i} \) are the centroids and \( \sigma_{a,i} \) and \( \sigma_{b,i} \) the standard deviations of classes a and b of combination i. The class separation is better for a lower value of the scatter ratio with a theoretical optimum of 0.

### 2.2. Experimental setup

To generate suitable training data for the sensor data processing concept described above, a generic hydraulic test bench operating under defined conditions in the lab was used. It consists of a load (figure 2c) and a combined cooling-filtration circuit (figure 2d) which are connected via the oil tank.

**Figure 2: Hydraulic test bench**

The system is equipped with several process sensors, such as pressure (PS), flow (FS), motor power (EPS), temperature (TS) and vibration (VS) sensors, and sensors for particle contamination of the oil (CS, MCS). The sampling rates of the sensor signals range from 100 Hz (pressures, power) to 1/cycle for contamination sensors. In order to simulate a typical industrial application, the hydraulic system performs a constant working cycle with changing load levels generated by the proportional pressure relief valve V6 as well as valve operations of V5. Wear of the internal gear pump MP1 (const. flow 7.5 l/min, 3.3 kW motor) leading to internal leakage can be simulated by two bypass orifices (three cascaded 0.2 mm and 0.25 mm diameter orifices generating 3.3 and 4.6 % leakage rate) activated by V4. The switching behavior of 2/2 directional spool valve V5 is another variable parameter where degradation can be simulated by a reduced control current (100, 85, 73 % of nominal value).
Concerning oil parameters, the temperature of the hydraulic oil (Meguin HLP 32) was increased stepwise by reducing the cooling efficiency of cooler C1 realized by a pulsed operation of the fan with variable duty cycle. A variable injection of pressurized air with up to 0.3 l/min into the suction line of MP1 allows aeration generating mixtures of free air and oil with a maximum ratio of 4 Vol.-% (figure 2a). Finally, the cooling-filtration circuit can be combined with an apparatus for adding defined test dust mixtures at variable concentrations to increase the particle content as shown in figure 2b. During the characterization measurement, different types of oil and component conditions and their grades of severity are combined with each other by nesting the variable parameters to complex profiles (see figure 3a). This ensures that the mutual influences of all variables and conditions are represented in the collected training data.

3. Impact of oil temperature and aeration

Three levels each for the oil cooling and aeration are superimposed with three nested component degradation levels each of MP1 and V5 leading to 81 conditions overall. The mean oil temperatures are increased by 7°C and 17°C, respectively, as a result of the reduced cooling power compared to normal operation with a mean oil temperature of 42°C and a variance of 1°C. The aeration levels are 0, 1.3, and 4.0 Vol.-% air in oil at ambient pressure. During the measurement, a constant working cycle (duration: 60 secs) divided into 13 intervals for feature extraction is repeated to allow comparable load and component characteristics for the statistical analysis (figure 3b).

The feature extraction is performed for all 18 process sensors and 13 intervals of the working cycle obtaining a total number of 1323 features which are ranked according to their Spearman correlation to the target conditions for pump leakage and directional valve degradation. The training dataset is adapted by successively including different
temperature and aeration levels to evaluate the impact of these oil parameters. 20 features are used as input for the LDA ensuring a minimum feature to data sample ratio of 1:10 in order to avoid overfitting effects. In case of pump leakage detection (figure 4a) enlarging the temperature and aeration range by itself leads only to a slight decrease of class separation but with still 100 % classification rate while a combination of a wide oil temperature and aeration range results in an overlap of the pump leakage classes (figure 4b) and, thus, a considerable decrease of the classification rate to 87.7%.

Figure 4: Pump leakage detection vs. oil aeration and oil temperature range showing (a) the scatter ratio for the 1st DF and (b) the corresponding 2-D LDA projections of the training data

A qualitatively similar behavior is also observed for the valve operation LDA projection (figure 5) while the class separation in general is higher compared to pump leakage detection. Here, the impact of high temperatures and air injection is also noticeable but relatively low (decline factor ≈ 3) in contrast to pump leakage detection (decline factor ≈
12). In order to further analyze the strong impact of injected air in oil at high temperatures on pump leakage detection (figure 4), the data set is split into 27 classes describing all combinations of oil temperature, aeration and pump leakage using the highest ranked 100 features extracted from a working cycle for further LDA (figure 6).

Figure 6: LDA plot with 27 classes divided by oil temperatures, aeration levels (air 1: 1.3 Vol.-%, air 2: 4.0 Vol.-%) and pump leakage levels (leak 1: 3 x 0.2 mm ⊗, leak 2: 3 x 0.25 mm ⊗ orifice) with 100 features, n=1134.

The classes exhibit an increased variance with rising oil temperature (positive DF1 direction), especially with changing aeration levels. While this variance is largely undirected for low oil temperature, the scattering direction caused by aeration is pointing mainly along DF1 for medium oil temperature and along DF2 for high temperature. Especially for the highest temperature, both aeration and pump leakage show very similar shift directions and the resulting class overlap makes discrimination difficult. In addition to its negative effect on the condition monitoring classification rate, free air in oil also has several undesired effects as described before. Unlike the oil temperature, measuring the aeration level of oil is complex and costly. Therefore, we also studied the potential for estimating the aeration level from the process sensor data using the same dataset as before. Here, the air injection set-points (0, 1.3, 4.0 Vol.-%) were defined as target value with the other effects (oil temperature and component conditions) as disturbances. Figure 7 shows (a) the classification rate and (b) the scatter ratio of the 1st DF while increasing the number of features and the number of cycles involved in a time-series feature extraction based on a large feature pool and subsequent selection by Spearman correlation ranking. Determination of the aeration...
level is not possible based on a single cycle with sufficient reliability, but a classification rate of up to 100 % can be achieved using a large number of cycles. Furthermore, the scatter ratio of the 1st DF further improves when increasing the number of features and involved cycles even when the classification rate has already reached 100 %. Hence, more than 30 included cycles and more than 40 features are required to allow an estimation of the aeration level with high reliability.

**Figure 7:** (a) Classification performance for the aeration level depending on number of features and number of included cycles for feature extraction (10-fold CV with Mahalanobis classifier), (b) corresponding scatter ratio of the 1st DF.

The distribution of the first one hundred features selected by the automated ranking is shown in **figure 8** in the categories sensors (see figure 2), working cycle intervals (see figure 3b), and feature functions. Here, the percentage of the cumulated absolute value of Spearman correlation of a specific sensor, interval, or function in relation to the overall cumulated correlation of all features can be used to evaluate the respective significance.

**Figure 8:** Ranking of (a) sensors, (b) intervals, and (c) features for estimation of the aeration level.
Regarding the sensor selection, the system efficiency factor of the load circuit (cf. 2.2), the vibration sensor (VS1) installed at the main pump MP1, the volume flow (FS1), as well as the housing temperature of MP1 (TS5) are ranked as most significant. The data-points selected for feature extraction within the working cycle are mainly the intervals 1 (complete cycle) as well as 13, 11, and 9 (stationary load intervals). Furthermore, the majority of features use the median function, followed by the higher statistical moments.

4. Detection of particle contamination

For the following measurement, the cooling-filtration circuit of the hydraulic test bench was coupled with an apparatus for defined injection of oil contaminated with particles (figure 2b). In a second small tank, a mixture of test dust (ISO MTD A3, ISO 12103-1) and oil with a concentration of 30 g/l is kept in suspension by a magnetic stirrer. This highly contaminated mixture is volumetrically diluted by at least a factor 375 with oil from the tank of the test bench (const. flow of 300 ml/min) and subsequently diluted further (1:33) by adding the mixture to the suction line of SP1 (const. flow of 10 l/min).

Figure 9: Characterization measurement with (a) particle contamination levels measured by Hydac CS1000 with particles > 14 μm below detection limit, (b) differential pressure of filter element F2, (c) box plot of measured ISO 4406 classes after filter.

In a characterization measurement over approx. 40 hours (figure 9), the variable flow of the peristaltic pump was changed every 7 hours with setpoints 75, 50, and 100 %
(generating a test dust mass flow of 16, 11, and 22 mg/min) alternated with periods at 0 % . This contamination is superimposed on the inherent background contamination of the oil in the hydraulic test bench which was determined as 14/9/7 according to ISO 4406 by the reference sensor Hydac CS1000 installed behind the filter element F2 (filtration rating 5 μm). Furthermore, the oil temperature was varied randomly (ΔT ≈ 10°C) and thereby also the oil viscosity to simulate realistic operating conditions resulting in the signal ripples of the pressure and CS sensor signals (figure 9a,b). The measurement was aborted after 2400 cycles (approx. 20.5 g test dust added) when the differential pressure over the filter element reached 5 bar. For the further analysis, 60 cycles after each setpoint change are excluded due to the slow transients before the signals are stable again. The classification of the filter loading is shown in figure 10. Sensor data of the three stationary filter loading states (no test dust addition) are extracted for each individual cycle and discriminated using LDA. Figure 10a is based only on the mean differential pressure over the filter as the state-of-the-art indicator for filter change, while in figure 10b the 50 highest ranked features extracted from different process sensors are used. While temperature changes lead to a high variance (up to 3.42) for the mean differential pressure feature and, thus, to an overlap of the classes, the statistical model in (b) shows a significantly reduced cross-sensitivity and allows perfect class separation with greatly reduced within-class variance (max. 1.11). To verify the statistical model, the remaining non-stationary phases during test dust addition are projected in the LDA plot (figure 10c) which shows that they are projected as transients at the correct positions.

Figure 10: Filter loading monitoring with (a) histogram using only the mean of Δp, (b) histogram using 50 selected features with highest correlation, and (c) corresponding LDA plot to (b)
To examine if the particle contamination injection flow can be quantified, the three phases with no dust addition are now combined in a common class and compared with the injection flow speeds 11 mg/min, 16 mg/min, and 22 mg/min. Using 50 time-series features for the LDA, mainly pressure characteristics measured in the filtration circuit, the class centroids are projected in the correct order along the 1\textsuperscript{st} DF (figure 11). However, high within-class variances lead to considerable overlap especially for the lower contamination levels resulting in a classification performance of 88\%. It should be noted that the oil temperature variations interfere considerably with the classification. Still, this approach is useful for identification of features which are robust vs. oil temperature changes. However, experimentally generated contamination levels were significantly higher than typical contamination levels in the field; thus, further improvement is required for practical application.

![Figure 11](image1.png)

**Figure 11:** Estimation of the particle contamination level using 50 time-series features extracted over 60 cycles: (a) LDA projection of four contamination classes and (b) histogram for the 1\textsuperscript{st} DF.

5. Conclusion and outlook
We have shown that the performance of a novel condition monitoring approach based on automated multivariate statistical analysis of process sensor data is impaired significantly especially if both oil temperature and oil aeration levels are varied during the experiments. On the other hand, different aeration levels could be detected and quantified correctly over a wide oil temperature range using time-series features selected based on their correlation. Furthermore, the method was successfully tested for monitoring of the filter loading state with reduced temperature cross-sensitivity compared to just relying on the differential pressure across the filter. The particle contamination level of the fluid, however, could not be estimated from the process
sensor data with high accuracy. The current data analysis procedure is based on supervised learning techniques and has to be expanded to unsupervised learning to be able to apply this method also to systems where different conditions cannot be tested experimentally.

6. Acknowledgement
This work was performed in collaboration with HYDAC Filter Systems GmbH, Sulzbach, Germany, and partially funded by the EFI program (Förderung von Entwicklung, Forschung und Innovation im Saarland). We would like to thank Jörg Kleber from HYDAC for technical support and valuable discussions.

7. References


