Data-Driven Serial Testing of Axial Piston Units: Integrating Signal Processing and Machine Learning for Roller Bearing Fault Detection

Maximilian Romeser[1,2] and Reinhold von Schwerin[1]

[1]Faculty for Informatics, Ulm University of Applied Sciences, Ulm, Germany
[2]Development of Axial Piston Units, Bosch Rexroth AG, Elchingen, Germany
E-Mail: maximilian.romeser@boschrexroth.de, reinhold.vonschwerin@thu.de

Abstract

Data-driven fault detection has emerged as an effective method for early anomaly detection and failure prevention, particularly through vibroacoustic analysis and the use of Machine Learning (ML) or Artificial Intelligence (AI). Given the central role of hydrostatic drives and axial piston units (APUs) in both conventional and electrified powertrains, their reliability is pivotal to prevent costly maintenance and machine downtime. Integral to the functionality of APUs are roller bearings, whose reliability and performance significantly influence the overall efficiency and longevity of the machinery. Given this context, the detection of faults in roller bearings within APUs is crucial to ensure operational safety and extend the lifespan of the systems. This is particularly important in the realms of condition monitoring and predictive maintenance, as well as serial testing, which is essential for comprehensive quality control, especially when considering alternative bearings. Therefore, advanced methods for detecting faulty roller bearings are needed, but challenges include hydraulically induced noise and manufacturing-related serial dispersion in serial testing scenarios. This paper presents an experimental investigation into roller bearing fault detection through vibroacoustic analysis. The study emphasizes the integration of advanced signal processing techniques and semi-supervised anomaly detection methods, utilizing various machine learning algorithms. Additionally, the research explores optimal sensor selection and operating conditions. The findings demonstrate that, under appropriate operating conditions and with the support of empirical knowledge, robust fault diagnosis can be achieved in the presented serial testing scenario.

Keywords: Axial Piston Units, Fault Detection, Anomaly Detection, Vibroacoustic Analysis, Machine Learning, Condition Monitoring, Serial Testing, Roller Bearings

1 Introduction

The robustness and high power density of hydrostatic drivetrains have long established them as a preferred choice in off-road applications. These systems are particularly valued for their precise control accuracy, high dynamics, and the ability to provide stepless transmission. Additionally, hydrostatic drives enable linear movements for various work functions, making them indispensable in demanding environments [1]. Given these advantages, hydrostatic drives, and APUs in particular, play a central role in the powertrain, whether it is a conventional drivetrain with an internal combustion engine or an electrified drivetrain, as shown in [2]. Therefore, the reliability of APUs is vital to prevent costly maintenance and machine downtime. As a result, methods for fault detection and diagnosis are the focus of various research efforts, primarily to enable condition monitoring (CM) and predictive maintenance, which are already well-established for gearboxes and turbines [3]. Additionally, robust and automated fault detection can serve as a valuable quality control measure during serial testing, particularly when products are in the ramp-up phase of production. Monitoring changes in suppliers is also crucial in this context.

Regarding the reliability of rotating machinery in general, and especially APUs [4], roller bearings are essential components [5]. Consequently, the detection of faulty bearings constitutes a significant area of research. In this context, the analysis of vibroacoustic behavior has emerged as an effective tool, enabling the detection of fault-related geometric deviations through anomalous excitations during operation [5]. This allows the detection of most failures in rotating machinery at an earlier stage than by monitoring temperatures or other signals [3, 6] However, in the case of axial piston units, this is particularly challenging compared to other rotating machinery. The abrupt impulses induced by the reciprocating pistons are extremely dominant, making it difficult to detect fault-induced anomalies in the emitted structure-borne or airborne noise [4]. Furthermore, when considering scenarios for fleet-based CM [7] or serial testing processes, it is important to account for serial dispersion. Conventional CM approaches typically assume a nominal reference state for the machine under consideration, where the signal behavior is initially stable, and faults manifest as anomalies relative to this reference. However, in the aforementioned scenarios, such an approach is not feasible due to the short testing duration, varying operating conditions and the potential presence of initial faults. Instead, anomalies in the data of new test units must be evaluated in relation to the historical data of previous, healthy units. Due to serial dispersion, the vibroacoustic behavior also varies, and fault-induced anomalies must be sufficiently pronounced to stand out against this variability [8].

Given the complexity of the vibroacoustic and structural dynamic behavior of APUs, this study aims to develop a data-driven approach based on measurements from an ensemble of units. The approach follows the concept of semi-supervised anomaly detection and falls within the field of ML and AI. An ensemble of tested, healthy units will be used to characterize production variability and train the model, while various faulty units will be evaluated during model testing. Since it is not straightforward to determine how fault patterns manifest in vibroacoustic behavior, different signal processing and feature generation methods will be investigated. The goal is to extract fault-relevant information from the data to ensure robust, data-based fault detection. Since vibroacoustic behavior can be highly dependent on operating conditions and sensor application, these factors will be comprehensively evaluated to provide a holistic assessment of fault detection and to determine the optimal conditions for serial testing processes. To the best of the author's knowledge, no study has yet explored the data-driven detection of faulty roller bearings in an ensemble of APUs. Therefore, the contributions of this study are as follows:

- Proposal and evaluation of a semi-supervised approach for detecting APUs with faulty roller bearings in a serial testing scenario using ML
- Evaluation of the combination of various approaches for signal processing and feature generation under various operation conditions
- Comparison of different sensor applications for analysis of vibroacoustic behavior of APUs

2 State of the Art

Fault detection in axial piston units has become a prominent field of research in recent years, driven by advancements in ML and signal processing, as shown in [9]. Typically, this research is conducted within the framework of CM or in a more general context. Only Gaugel et al. [10] consider a serial testing scenario of APUs, without delving into the specific faults and signals considered and instead focusing on time series segmentation. In the following, the authors of this paper limit themselves to approaches that address vibroacoustic analysis. In particular, data-driven approaches including signal processing and ML techniques are frequently discussed, as model-based approaches are very challenging due to the system's complex vibroacoustic behavior [11]. A comprehensive study on the use of conventional ML in terms of supervised classification for CM scenarios is provided by Torrika [12], who particularly highlights the steps of feature extraction and reduction. An analogous approach to feature generation in a supervised classification setting can be found in [13]. Further intensive research on the condition monitoring (CM) of axial piston pumps using supervised methods such as Support Vector Machines (SVMs) and gradient boosting algorithms, utilizing a feature engineering approach based on tsfresh, was conducted by Horn et al. [14]. In contrast, Gnepper et al. [15] and Siyuan et al. [16] explore a semi-supervised anomaly detection scenario and employ principal component analysis to characterize the nominal state of the APU within the context of CM.

To better highlight fault information, signal processing techniques such as Fast Fourier Transformation (FFT) [15, 17, 18], Wavelet Packet Transformation (WPT) [9, 16, 17], Empirical Mode Decomposition (EMD) [18], Cepstrum Analysis [11], or Time Synchronous Averaging [13] are frequently employed. Specifically, in the detection of faulty bearings in APUs, various techniques are commonly employed to generate sensitive filter bands and enhanced signal envelopes [4,19-22]. For the fault detection of worn slippers, the principle of cyclostationarity plays a particularly important role here, as investigated by Casoli et al. [23]. By considering the results in [17], it can be observed that commonly studied fault patterns, such as wear on slippers, port plates or cylinders, can have immediate effects on the dominant harmonics of the piston frequency. However, this is not the case for roller bearings, which imposes different requirements on signal processing. Advanced signal processing techniques also support the application of Deep Learning, which has been extensively studied in the context of fault detection for APUs in recent years [24-31]. As such, model training typically involves feature extraction based on scalograms or spectrograms, rather than directly utilizing the raw time series data as in [31]. However, being supervised classification settings the foundation of these approaches relies on the identification of known, specific fault states resulting from wear and damages, as well as on sufficient data volumes without considering serial dispersion due to manufacturing tolerances. As Keller et al. [32] emphasize, this must be ensured for adequate validation. Nevertheless, in the case of CM, these factors are frequently irrelevant, since the goal is to identify faults relative to the normal condition of the same unit, as mentioned before. This contrasts with the present use-case.

Although the developed methods are often tested at various operating conditions, only Horn et al. [14] and Gnepper et al. [15] investigate a wide operating range and determines the quality of fault detection with respect to slipper clearance or also slipper wear and valve plate cavitation [14] depending on this range. Additionally, while the experimental setups of some of the mentioned studies include multiple structure-borne or airborne sound sensors, such as in [4] or [9], this aspect is not always analyzed in detail. Casoli et al. [13, 17, 23] and Du et al. [11], for instance, only consider structure-borne sound in axial and radial directions, whereas Gnepper et al. [15] analyze different sensor concepts and positions. Furthermore, the investigations in [32] indicate that positioning the sensor close to the bearing is effective, even when the valve plate is damaged. Other studies address the fusion of multi-sensor data, including acoustic signals [31] or pressure signals [24].

3 Experimental Design

To replicate a scenario of serial testing with an ensemble of both healthy and faulty units, an experimental series is conducted on a laboratory test bench. The experimental design, including the materials and the test sequence, as well as the data acquisition setup is described below.

3.1 Material and Fault Types

In this study, APUs in bent axis design were tested using a laboratory hydraulic test bench, as shown in Figure 1. Each APU was initially tested in its original delivery condition. Subsequently, to characterize the range of vibroacoustic variation attributable to component interchangeability, the drive shafts were systematically interchanged among the housings. For example, in the first round of swaps, the drive shaft from the first APU was placed into the second housing, the second unit's drive shaft into the third housing, and so on. While a full exploration of all 49 possible unique pairwise combinations was desirable, time constraints limited this cross-assembly testing, constituting the first phase of the experimental series, to 30 unique drive shaft-housing combinations (achieved through 23 swaps). Throughout this phase, the drive assemblies remained equipped with their original, undamaged bearings, which were not removed. This approach enabled assessment of vibroacoustic variability while isolating the effects of drive shaft and housing variations.

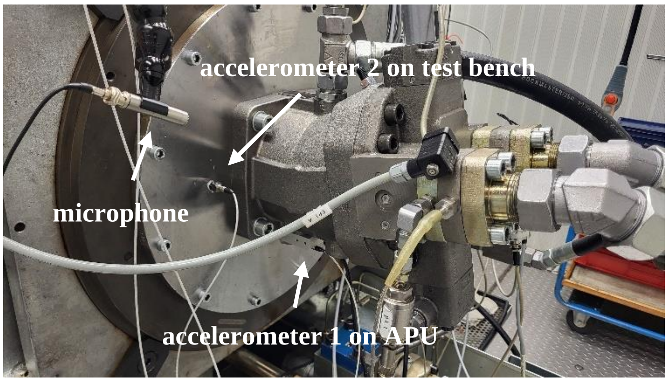

Figure 1: Experimental setup at the laboratory test bench

In the second phase of the experimental series, 14 drive assemblies with faulty roller bearings are installed. In 12 cases, this relates to the large roller bearing and in two cases the small roller bearing. Faults in other components of the drive assembly are neglected for the purposes of this investigation, meaning that these components remain in a healthy condition. The types of faults in the tapered roller bearings are

Figure 2: Manipulated roller elements (left) and outer ring (right) with abrasively generated spalls

intentionally chosen to be heterogeneous to generate a broad portfolio of potential anomalies in the data and to ensure comprehensive validity. The bearings are manipulated, with different sections of the raceways or end faces of the rolling elements abrasively processed, as exemplarily depicted in Figure 2. The defects vary in depth and width, as well as in number and location of the defects. Regarding further investigation, it is crucial to generate both easily detectable faults (e.g., signals with significantly higher signal energy) and subtle faults, identifiable only by the presence of small peaks in the frequency spectrum, to evaluate the robustness of the fault detection method. Due to modulation effects resulting from tolerances in the APU, various effects of impact excitations on the structure-borne sound emissions are observed. It should be noted that due to the cage, subsequent processing of the inner ring is not possible. Each faulty drive assembly is mounted once in one of the housings, and the assembled unit is then tested on the test bench. The assignment of faulty drive assemblies to housings is done randomly.

3.2 Test Sequence

Each unit or drive assembly-housing combination, regardless of the presence of a fault, undergoes the same automated test sequence on the test bench for high reproducibility. Initially, a 15-minute run-in period is conducted to achieve an oil temperature level of 40 °C. Since the regulation is not perfect, fluctuations can occur during operation, especially when load points are varied. In the end, the overall oil temperature range, across all following load points, was 42 °C to 56 °C. For individual load points, the experimentally determined temperature variation did not exceed 8 °C. It is plausible that these fluctuations, particularly in relation to oil viscosity, may influence the vibroacoustic emissions. However, this can also occur in real serial testing or fleet-based CM scenarios, and thus, it is treated as regular process variation.

The test procedure involves the traversal of a characteristic map within a three-dimensional operating range. An operating point is defined by the shaft speed, the pressure in the high-pressure line, and the swivel angle. The load in the closed circuit is adjusted in four stages between 100 bar and 400 bar using a load valve. The speed is varied in five equally sized stages between 400 and 2000 rpm. The swivel angle is set to either its minimum or maximum position. The swivel angle as well as the high-pressure influence the load on the bearing, which is why these parameters are referenced accordingly in the Table 1 for clarity. Consequently, the resulting section of the test sequence includes 40 load points. Between the load points, the operating conditions are adjusted in a ramp-like manner. Each load point is then held for ten seconds.

Table 1: Indexing of the different bearing loads (0: Min. swivel angle, 1: Max. swivel angle)

Bearing Load	1	2	3	4	5	6	7	8
Pressure [bar]	100	200	300	400	100	200	300	400
Swivel Angle	0	0	0	0	1	1	1	1

3.3 Sensors and Data Acquisition

To capture the vibroacoustic behavior, three sensors were applied: Firstly, a free-field microphone is permanently installed to measure airborne noise, as shown in Figure 1, at approximately 30 cm above the APU. Secondly, a triaxial piezo-electric accelerometer is permanently attached to the test bench using an adhesive adapter to capture structure-borne noise emissions (accelerometer 2). Since the sensor application is performed just once before the test series, high reproducibility of the results is ensured. This permanent application is particularly advantageous for serial testing, as it eliminates the need for manual sensor placement by the operator or automatic positioning systems. Additionally, it is beneficial to measure as close as possible to the relevant source of structure-borne noise, i.e., the roller bearing, to achieve a high signal-to-noise ratio (SNR). For this purpose, another accelerometer of the same type is permanently mounted using an anodized screw adapter, as shown in Figure 3 (accelerometer 1). The adapter is attached to the bolting of the bearing cross-flushing and is therefore reattached with each new unit mounted on the test bench, i. e. for each test. Consequently, this has a negative impact on the dispersion between the conducted tests.

The sampling rate is consistently set to 50 kHz. To avoid aliasing effects, a Butterworth low-pass filter with a cutoff frequency of 25 kHz is employed. Data acquisition is performed in intervals and controlled by an automated test sequence. After a stabilization period of three seconds, interval measurements of five seconds duration are conducted at each of the 40 load points. For further investigations, the signal channel and the time series from the microphone are referenced as Lp, the channels from accelerometer 1 as a1x, a1y and a1z, and those from accelerometer 2 as a2x, a2y and a2z. Consequently, the relevant dataset consists of 40 x 7 time series, e. g. data instances, for each of the 44 tested units, with each time series comprising 250000 samples. Consequently, there are 40 x 7 data subsets including 44 instances per each.

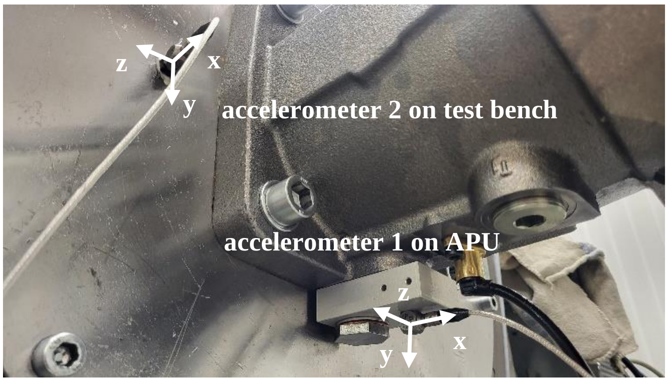

Figure 3: Detailed view of the attached accelerometers

4 Data-Driven Methodology

The methodology employed in this paper for generating a data-driven model for fault detection consists of three main steps: signal processing, feature generation, and anomaly detection, with the latter encompassing the actual ML. As illustrated in Figure 4, various methods for signal processing and different feature sets are utilized. The features thus serve as the input for the semi-supervised anomaly detection model, enabling it to learn the characteristics of vibroacoustic data from healthy units. While the features can be directly applied to the raw time series data $x(t)$, it is often more effective to precede the feature generation process with signal processing methods, as demonstrated by Bienefeld et al. [33] in a comprehensive study on this subject. To compare these methods while considering the dependency on the operating conditions and sensors, an appropriate metric is used to evaluate the anomaly detection and thus the quality of fault detection.

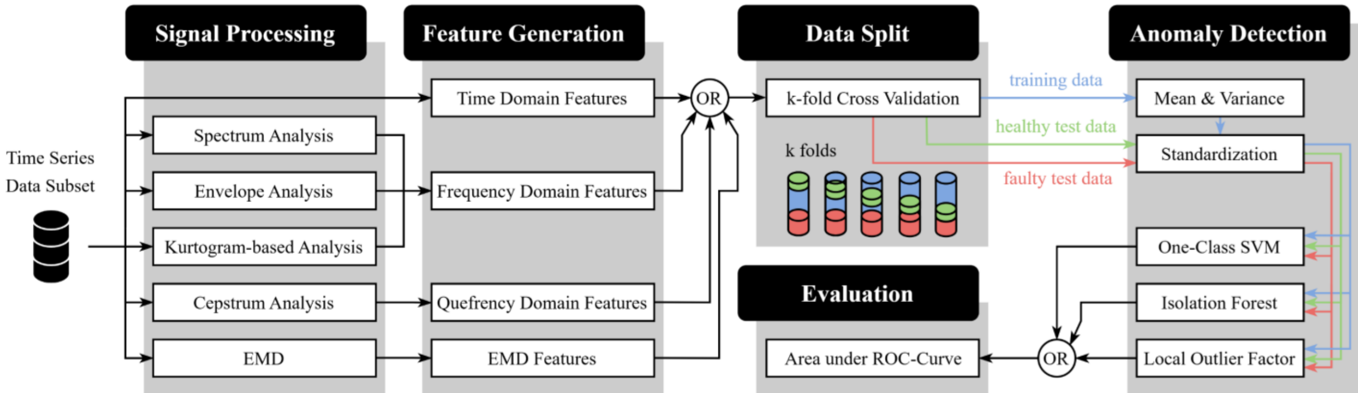

Figure 4: Integration of the data-driven methodology. The data pipeline enables the variation of signal processing techniques and their corresponding features as well as the variation of the anomaly detection model (OR gates). Note, that the pipeline is set up for each individual data subset regarding different operating conditions and sensor signals.

4.1 Signal Processing

In the scope of this work, five signal processing methods are investigated: spectrum analyses in frequency domain (FD) which includes the analysis of the raw spectrum, the envelope spectrum and the kurtogram-based envelope spectrum, cepstrum analysis and EMD. Additionally, the direct generation of features from the time series data is considered, too. A deeper understanding of the methods is provided by Randall et al. [5]. A low-pass filter with a cutoff frequency of 3 kHz was selected based on visual inspection of the raw time series data and consideration of the dominant frequencies observed in preliminary testing. Above 3 kHz, the spectra were dominated by broad-band elevations indicative of test bench and sensor coupling resonances, particularly pronounced for accelerometer 1. These resonances varied significantly between individual units, potentially masking the APU-specific vibroacoustic characteristics of interest. The 0-3 kHz range captured the most prominent features related to APU operation, as indicated by Figure 6.

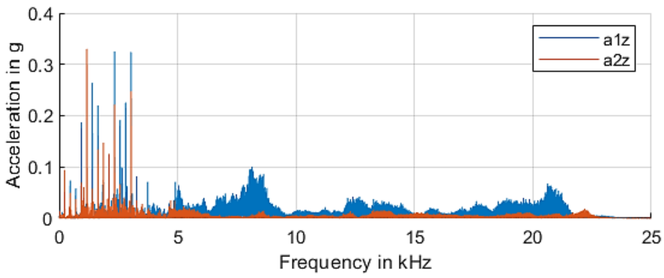

Figure 6: Entire spectra of a healthy unit in z-direction before low-pass filtering

4.1.1 Spectrum Analyses

Bearing faults often generate characteristic vibrations at specific frequencies depending on the type of fault, operating parameters, and bearing dimensions. By applying FFT to the vibroacoustic data, these frequency components can be identified in the frequency domain, i. e. in the raw spectrum $X(f) = \mathcal{F}\{x(t)\}$.

As the raw signal spectrum often provides limited diagnostic information about bearing faults, envelope analysis has become an established method for bearing diagnostics [5]. The envelope-based method exploits the fact that a high-frequency carrier signal, which reflects the structural resonances and is influenced by other excitations, is amplitude modulated by the recurring impulses caused by bearing faults. The relevant information in the raw signal $x(t)$ can thus be identified through demodulation, where the envelope $e(t)$ can be generated using the Hilbert transform $\mathcal{H}\{x(t)\}$ as follows:

$$e(t) = \sqrt{x(t)^2 + \mathcal{H}\{x(t)\}^2} \tag{1}$$

and the FFT of this envelope reveals the envelope spectrum $\mathcal{F}\{e(t)\}$, as exemplarily shown in Figure 5. It should be noted, however, that the envelope can be significantly affected by piston frequencies, as demonstrated by Xiao et al. [4]. As mentioned by Randall et al. [5], it thus makes sense to identify frequency bands that are primarily dominated by bearing resonances before demodulation. In the context of this work, these were determined empirically, and a bandpass filter with a passband of 2 to 3 kHz was chosen for the envelope-based analysis. This contrasts with the purely FFT-based approach, which examines the entire frequency band up to 3 kHz without performing envelope analysis.

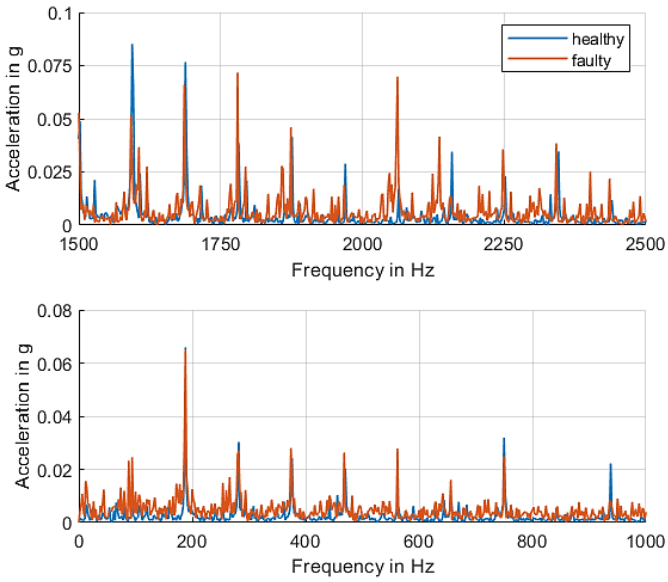

Figure 5: Raw spectrum (top) and envelope spectrum (bottom) of a healthy and faulty unit considering a2y

Nevertheless, the identification of a suitable filter is generally non-trivial. However, spectral kurtosis provides a tool that relates the impulsiveness of the signal to its frequency components. The underlying hypothesis is that frequency bands containing information relevant to bearing fault detection exhibit higher spectral kurtosis. For simplicity, a detailed explanation of the method is omitted here, and instead, explicit reference is made to [5]. However, it is important to understand that a window is slid over the time series and a FFT is performed for each window, resulting in the Short-Time Fourier Transformation (STFT) and resulting in the spectrogram $S(t, f)$. This process transforms the signal into the time-frequency domain. The spectral kurtosis $K(f)$ is calculated as

$$K(f) = \langle \frac{|S(t,f)|^4}{|S(t,f)|^2} \rangle - 2 \tag{2}$$

with $\langle\cdot\rangle$ characterizing the time-averaging operator. Frequency ranges with high spectral kurtosis are then suitable filter bands. Determining these ranges depends on the ratio of the window length to the unknown spacing between the periodically occurring impulses. Therefore, the kurtogram is constructed, which further represents $K(f)$ as a function of the window length. Due to the time-frequency uncertainty principle, the frequency resolution increases with longer window lengths.

4.1.2 Cepstrum Analysis

Another method to reduce the influence of disturbances in relation to fault-induced anomalies is cepstrum analysis. In the cepstrum, families of periodically spaced spectral components appear as prominent peaks, known as rahmonics, which include both harmonics and sidebands in the spectrum. These phenomena are caused by impulsive forces due to roller element bearing faults combined with amplitude modulations resulting from phase shifts between the load and the fault positions [34]. According to Du et al. [11], this approach has been successfully applied in the context of bearing fault detection in APUs.

The cepstrum $C(T)$ is generated by first transforming the signal into the frequency domain using the FFT. After logarithmizing the spectral amplitudes, the signal is then transformed into the quefrency domain (QD) with $T = 1/f$, which can be expressed as

$$C(T) = |\mathcal{F}^{-1}\{\log|\mathcal{F}\{x(t)\}|^2\}|^2. \tag{3}$$

In simple terms, the cepstrum represents the inverse transformation of the logarithmized power spectrum $\log|\mathcal{F}\{x(t)\}|^2$, thereby presenting the cepstrum in relation to the quefrency components in time domain (TD) once again. The reciprocal quefrency T provides insights into dominant families of frequency components.

4.1.3 Empirical Mode Decomposition

A purely data-driven and heuristic approach to signal processing is EMD, which enables the decomposition of $x(t)$ into N so-called Intrinsic Mode Functions (IMFs) and a residual $r_N(t)$:

$$x(t) = \sum_{i=1}^{N} IMF_i(t) + r_N(t) \tag{4}$$

The IMFs characterize different frequency bands, which is particularly advantageous for diagnosing faulty roller bearings when there is no prior knowledge of relevant bands or the signal structure. To determine the individual IMFs, the algorithm iteratively decomposes $x(t)$ in TD through a process called sifting, which proceeds as follows:

1. Find the local minima and maxima of $x(t)$
2. Construct lower and upper envelopes $v_1^-(t)$ and $v_1^+(t)$ by using the local extrema
3. Calculate the mean of envelopes $m_1(t)$
4. Subtract $m_1(t)$ from $x(t)$ to obtain residual $r_1(t)$

This process is repeated N times with $x(t) := IMF_i(t) = r_i(t)$ till a specific condition is met: In this study, the process is completed after $N = 5$ repetitions, resulting in four IMFs and the final residual $r_5(t)$.

4.2 Feature Generation

Given the diverse nature of fault cases, it is challenging to generate fault-related features in advance, as these may not be suitable for extracting the corresponding anomalies from the signal in every fault scenario. It is worth mentioning that, even for roller bearings alone, a multitude of fault cases is possible, which cannot all be considered within the scope of this study. Therefore, it is even more important to develop appropriate features. If the features are to be derived directly from the time series, e. g. one data instance, statistical metrics are used for this purpose. This includes the first four statistical moments: mean $\mu_x = F_{TD_1}$, variance $\sigma_x^2 = F_{TD_2}$, skewness F_{TD_3}, and kurtosis F_{TD_4}, with the latter two defined as

$$F_{TD_i} = \mathrm{E}\left[\left(\frac{x-\mu}{\sigma}\right)^i\right] \;\forall\, i \in \{3{,}4\}. \tag{5}$$

Depending on the amplitude, outliers receive a higher weight. Additionally, the Root-Mean-Square (RMS) and the median are used as features F_{TD_5} and F_{TD_6}. Analogous to [32], it makes sense to relate features in FD (due to preprocessing with the methods from sections 4.1.1 to 4.1.4) to specific frequency bands. This approach allows, for example, the isolation of bands where bearing faults excite the corresponding bearing resonances [35] from bands where piston frequencies are predominantly dominant. Therefore, the spectrum is divided into five frequency bands, according to Table 2, and the mean and kurtosis are calculated for each band. The same approach for feature generation can also be applied to the cepstrum in QD. In the case of the kurtogram-based method, it is also useful to use the boundaries of the selected frequency band as an additional feature. Since EMD is performed in TD, the RMS is calculated instead of the mean to characterize signal power. An overview of the features is given in Table 3. Consequently, the dataset is divided into 6 x 40 x 7 subsets in the feature domain. Each subset is standardized to ensure that all features have equal influence in the ML process.

Table 2: Frequency bands for feature generation in frequency domain. The corresponding quefrency bands have reciprocal start and end values.

Band No.	1	2	3	4	5
Start Freq. [Hz]	0	120	360	840	1800
End Freq. [Hz]	120	360	840	1800	3200

Table 3: Overview of different feature sets

Set	Size	Feature Description
TD	6	Mean, variance, skewness, kurtosis, RMS, median
FD	10 (+ 2)	Mean and kurtosis per band (+ passband frequencies defined by the kurtogram)
QD	10	Mean and kurtosis per band
EMD	10	RMS and kurtosis per IMF or residuum

4.3 Anomaly Detection

In this study, it is assumed that faulty units generate anomalous data. Given that anomalous data are typically rare, it is often practical for anomaly detection to use normal data to construct a model that can reflect the characteristics of normal data (semi-supervised anomaly detection or novelty detection). These models are then capable of identifying anomalous data based on their deviating characteristics and thus a higher anomaly score. In this study, three ML techniques are employed for anomaly detection (see Figure 4): the One-Class Support Vector Machine (OC-SVM), the Isolation Forest (IF), and the Local Outlier Factor (LOF). The reason for using multiple techniques lies in the fact that it is not trivial to determine which method will yield the best results for a given dataset. By applying these diverse approaches, we aim to comprehensively evaluate their performance and robustness in detecting anomalies.

To ensure comparable influence of features during model training, all feature sets are standardized to a similar scale. As test data must remain unseen, standardization is performed exclusively based on the training data, as illustrated in Figure 4. Specifically, the mean and variance are computed from the training set and subsequently applied to both training and test data (z-standardization).

4.3.1 One-Class Support Vector Machine

OC-SVM is a variant of the SVM designed for anomaly detection [35]. Unlike a binary SVM, which separates two classes with a hyperplane and is trained on normal and anormal data, OC-SVM is trained only on normal data. Hence, the objective is to find a hypersphere that maximizes the margin from the origin, where normal data lies, ensuring that the anomalous data lies in the positive half-space. A data instance p can be expressed as a vector with n_F Features. When slack variables are introduced to allow for some misclassifications, the optimization objective can be expressed as:

$$\min \frac{1}{2}\left||w|\right|^2 + \frac{1}{\nu n_{Train}} \sum_{i}^{n_{Train}} \xi_i - b \qquad (6)$$
$$\forall \langle w, \Phi(p_i)\rangle \geq b - \xi_i \ \wedge \ \xi i \geq 0$$

where ξ is a vector of slack variables, Φ can be a nonlinear mapping function to the kernel space, b is the bias term, ν is a regularizing parameter for ensuring most training instances are within the hypersphere while minimizing the weights w, and n_{Train} is the number of data instances for training. Once the optimization is solved, an anomalous data instances o can be detected if the following condition is fulfilled:

$$\mathrm{sgn}(\langle \mathrm{w}, \Phi(\mathrm{o})\rangle) - b = 1 \qquad (7)$$

In this work, the radial basis function is used as the kernel function. The parameter ν, which is a hyperparameter not adjusted during training, is pre-set to a value of 0.5. The closer ν is to 1, the more flexible and curvier the decision boundary becomes.

4.3.2 Isolation Forest

Instead of profiling normal data, the IF algorithm [37] isolates anomalies by recursively partitioning the data. Starting with n instances, the algorithm randomly selects an attribute a, e. g. a feature, and a split value a_{split} to divide the data. An internal node is created with the condition $a < a_{split}$, splitting the data into two subsets, where $a < a_{split}$ and where $a > a_{split}$. This process is recursively applied until a termination condition is met: The tree reaches a height limit, a subset contains only one instance, or all data points in a subset have identical values. The resulting so-called iTree is a binary tree and assuming that all instances are distinct, each instance is located as an external node when the iTree is fully grown. This tree structure ensures that anomalies, being fewer and more distinct, are located closer to the root, resulting in shorter path lengths. The path length $h(p)$ of an instance p is defined as the number of edges traversed from the root to an external node in an iTree. Anomalies tend to have shorter path lengths. The anomaly score $s(p)$ is then calculated as

$$s(p) = 2^{-\frac{E[h(p)]}{c(n)}}. \qquad (8)$$

where $\mathrm{E}[h(p)]$ is the average path length from several iTrees, in our case 100, and $c(n)$ is the average path length for unsuccessful searches in a binary search tree, where explicit reference is made to [37].

4.3.3 Local Outlier Factor

The LOF algorithm, first introduced in [38] is a density-based method where the process of calculating the LOF includes defining the degree of outlying. By introducing the concept of locality, LOF performs well in scenarios with a dataset with fluctuating density [39]. The LOF of an instance p can be calculated through

$$LOF_k(p) = \frac{\sum_{q \in N_k(p)} \frac{Lrd_k(q)}{Lrd_k(p)}}{|N_k(p)|}. \qquad (9)$$

which is the average of the local reachability density (Lrd) of the instance and the k-nearest neighbors, where $N_k(p)$ is the number of neighbors of p. Note that $N_k(p) > k$ is possible for multiple instances q with the same distance. Lrd is defined as the inversion of the average reachability distance of the k-nearest neighbors of p:

$$Lrd_k(p) = 1/\frac{\sum_{q \in N_k(p)} Rd_k(p,q)}{|N_k(p)|} \qquad (10)$$

The reachability distance $Rd_k(p,q)$ is defined as the distance $d(p,q)$ in the feature space but is at least the k-distance of q whereas the k-distance of q, denoted as $kd(q)$, is the distance of q to the k-nearest neighbor. This can be expressed as follows:

$$Rd_k(p,q) = \max\{d(p,q), kd(q)\} \qquad (11)$$

In the present use-case, Euclidean distance is used and k is the number of training samples. In general, higher values for LOF indicate anomalous data instances whereas values around or below 1 indicate normal data instances.

4.4 Data Split and Evaluation

For the evaluation of anomaly detection quality, the concept of fold cross-validation is employed, as the dataset is relatively small and the significance of the results with randomly held-out test data (hold-out validation) would otherwise be low. As shown in Figure 4, five train-test splits of the dataset are generated, and a model is trained with the five resulting different training datasets, ensuring that the same instance does not appear in multiple splits. The exception to this rule is the faulty data, which is completely included in each split, so only the healthy data instances vary. The determination of model performance or the quality of anomaly detection is then based on the arithmetic mean of the evaluation metrics considered across the different splits. The metric considered is the AUC, i.e., the area under the receiver-operating-characteristic-curve (ROC curve), which varies the threshold with respect to the anomaly scores of the aforementioned anomaly detection techniques. Therefore, the ROC curve plots the true positive rate (TPR) over the false positive rate (FPR).

In this study, data from faulty units are labeled as positive, anomalous instances and data from healthy units as negative, normal instances. According to this terminology, the model can incorrectly predict positive instances as negative or vice versa. There is an inherent trade-off, as a model that very confidently predicts positive instances may also frequently predict negative instances as positive, resulting in ambivalent effects regarding TPR and FPR. Therefore, a model is considered robust only if it provides correct predictions overall. Accordingly, a high AUC (close to 1) indicates robust fault detection, whereas a value close to 0.5 indicates random fault detection. In addition to the AUC, the TPR for a FPR of 0, as well as the FPR for a TPR of 0, are also analyzed to evaluate the quality of fault detection under specific requirements. In the former case, the TPR ideally takes a value of 1, while in the latter case, the FPR should be 0.

5 Results

To keep the evaluation of the experimental results clear and comprehensible, the analysis of the signal processing methods and signal channels is initially conducted using the averaged AUC values. Therefore, the arithmetic means of the AUC values using the OC-SVM, the IF and the LOF is calculated. To characterize fault detection performance, we consider results from numerous model training runs. Each training process is performed independently for each load point (as explained in Figure 4) and for each combination of measurement channel and signal processing method. The results are treated as individual data points within a distribution. These distributions are visualized as box plots in Figure 7. According to the box plots in Figure 7, it can be observed that an accelerometer is preferable to a microphone based on fault detection quality, provided no signal processing methods are applied. However, when choosing the envelope-based method, results can be achieved that are approximately equivalent to those obtained with the application of accelerometers. Here too, considering the median and the load point-specific dispersion in the representation, it is evident that the envelope-based method yields better results. The highest median AUC, around 0.954, and the smallest dispersion ranges are achievable using the envelope-based method and the accelerometer fixed to the test stand in the planar directions. However, methods based on Kurtogram, Cepstrum, and EMD generally tend to deliver poorer results, even when compared to the FFT-based method. For a more detailed examination of the load points and a deeper understanding of the sensors, only the envelope-based and FFT-based methods will be considered henceforth, as the latter does not require domain knowledge in the form of a suitable filter band in contrast to the envelope-based method.

Figure 9 clearly confirms that the envelope-based method, when considering accelerometer 2, delivers higher AUC values across the entire operating range. It further shows that the rotational speed can have a significant impact on the quality of fault detection. While with this combination, the AUC only drops to a worst-case value of 0.843 at 400 rpm, values of 0.630 can be achieved using the envelope-based method applied to accelerometer 1 in the y-direction, and values of 0.571 can be achieved using the microphone and the FFT-based method. When using accelerometer 1 and the FFT, there are partially opposing effects, as higher rotational speeds result in lower AUC values, especially at large swivel angles (bearing loads 5 to 8). However, the effect of bearing load on fault detection is not as apparent and necessitates a more detailed examination using the FPR assuming a TPR of 0, and the TPR assuming a FPR of 0. Figure 8 shows that the envelope-based method tends to yield higher TPR values and lower FPR values at smaller swivel angles, although there are outliers, such as at 1600 rpm and bearing load 1. In this context, it is important to emphasize that the FPR can increase rapidly when the TPR is 0, especially if a single faulty unit does not exhibit truly anomalous behavior. If the goal is to avoid falsely detecting good units (see TPR), higher rotational speeds remain essential for robust fault detection. In particular, in these scenarios, the superiority of the envelope-based method becomes evident. A summary of the maximum AUC values and TPR, as well as the minimum FPR values for the

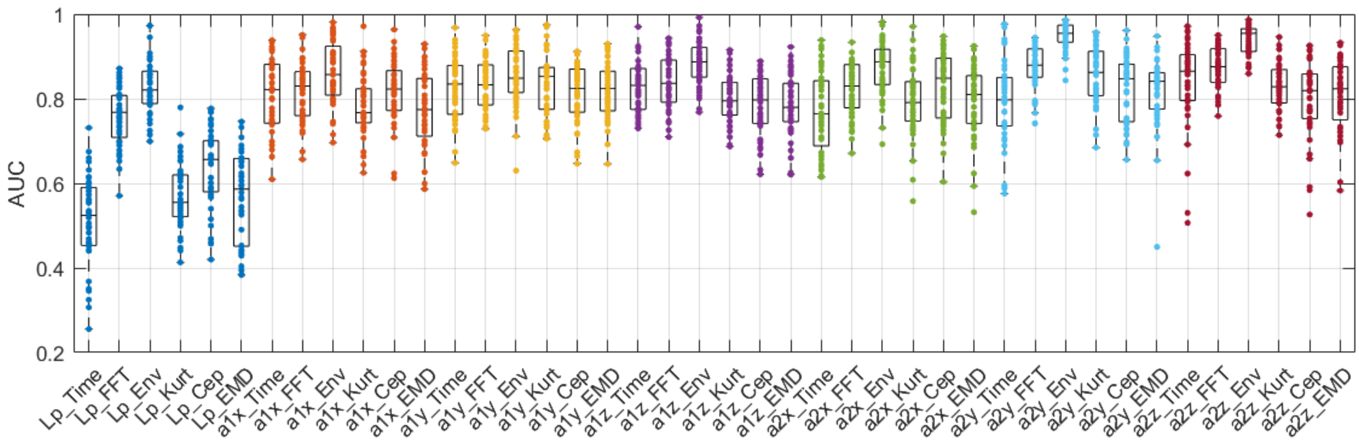

Figure 7: Distribution of AUC values obtained by averaging the models across the operating conditions. Each box represents the AUCs for a specific signal and signal processing method (Env: Envelope, Kurt: Kurtogram, Cep: Cepstrum).

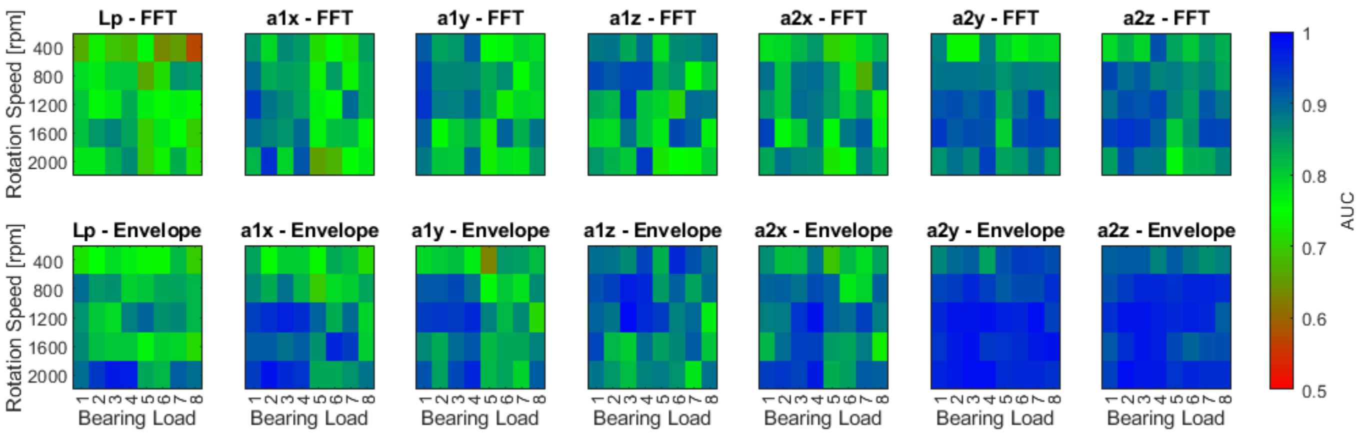

Figure 9: AUC values obtained by averaging the models depending on the operating condition

different sensors, is provided in Table 4 with corresponding signals, signal processing methods and operating conditions. The envelope method consistently delivers the best values for the metrics, and the z-direction is usually the most suitable.

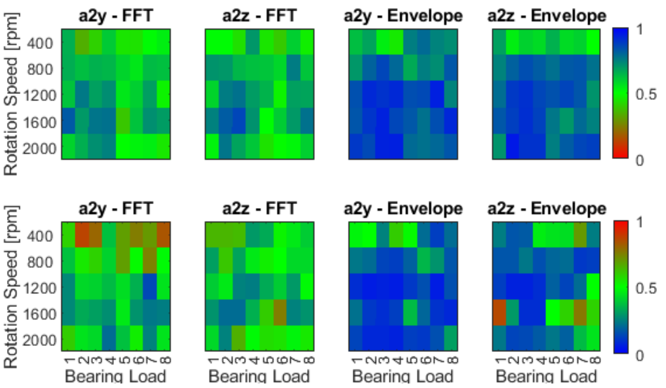

Figure 8: TPR (top) and FPR (bottom) values obtained by averaging the models depending on the operating condition

Table 4: Best values of the metrics for different sensors obtained by averaging the models (Acc: Accelerometer, Mic: Microphone)

Sensor	Metric	Value	Operating Condition	Signal	Signal Processing
Mic.	AUC	0.973	2000 rpm, BL 3	Lp	Env.
	TPR	0.906	2000 rpm, BL 3	Lp	Env.
	FPR	0.133	2000 rpm, BL 4	Lp	Env.
Acc. 1	AUC	0.992	1200rpm, BL3	a1z	Env.
	TPR	0.952	1200rpm, BL3	a1z	Env.
	FPR	0.056	1200rpm, BL3	a1z	Env.
Acc. 2	AUC	0.987	2000 rpm, BL 2	a2z	Env.
	TPR	0.952	2000 rpm, BL 2	a2z	Env.
	FPR	0.056	1200rpm, BL 2	a2y	Env.

Particularly when considering accelerometer 1, medium rotational speeds are well-suited, whereas for the microphone and accelerometer 2, the maximum rotational speed typically yields the highest results at low swivel angles. The FPR values, due to the averaging of the values for the three models, should be interpreted as follows: one of the models for anomaly detection has falsely detected a good unit as defective when it is required that all defective units must be correctly identified.

To gain a deeper insight into the performance of individual models for anomaly detection, it is crucial to compare their metrics rather than relying on the mean values as previously considered. The distributions, which arise from different operating conditions, reveal that no definitive general statement regarding performance can be made when varying scenarios for data collection and processing are taken into account, as illustrated in Figure 10. Using the IF (red), the median values for the metrics obtained from applying the FFT-based method to the signals from accelerometer 2 are inferior compared to those obtained using the OC-SVM (blue) and the LOF (yellow). However, an analysis of the metrics derived from the envelope-based method presents a different picture, with the OC-SVM yielding poorer median values and exhibiting significantly more variability depending on the operating conditions. Particularly in the lower plot of Figure 10, it becomes evident that the variability for signal a2y and the envelope-based analysis is considerably smaller, although the FPR can generally vary widely with respect to variable operating conditions. Given that the different models incorporate distinct concepts for anomaly detection, it is prudent to examine the correlations between the individual models in greater detail. Figure 11 depicts the AUC values of the models for the previously analyzed signals and signal processing methods. Compared to a perfect correlation (dashed line), the discrepancies between the IF and the other two models are apparent. A depiction with a linear regression (red line) is inadequate here, resulting in a lower Pearson Correlation of 0.52 and 0.65, respectively. In contrast, a value of 0.91 when comparing the OC-SVM and the LOF indicates a significantly more similar categorization of units and detection of faults.

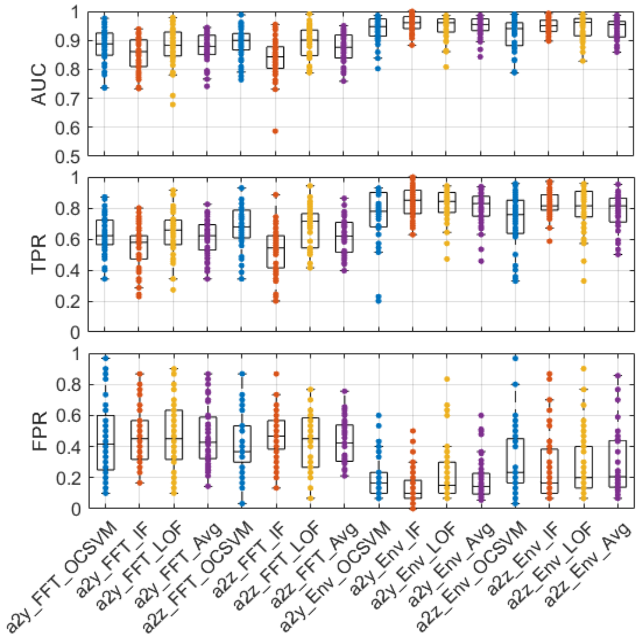

Figure 10: Distribution of AUC, TPR and FPR values for every single model across the operating conditions

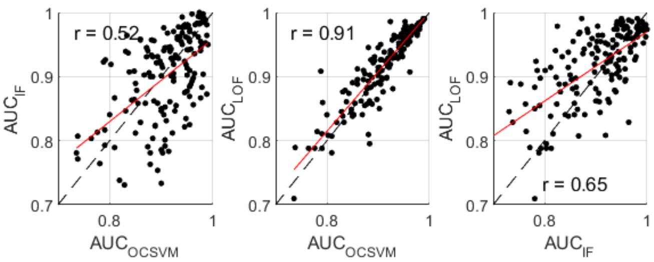

Figure 11: Correlation between the AUC values of the different models considering a2y and a2z (dashed line: perfect correlation, red line: linear regression)

6 Discussion and Future Work

The results of the experiments indicate that detecting faulty roller bearings in a serial testing scenario is a non-trivial task, as different methods of data acquisition and processing, as well as operating conditions, can have a significant impact. The choice of sensors is critical, especially in the absence of domain knowledge and when only limited signal processing methods, such as FFT, are available. In the time domain, robust fault detection using a microphone is not feasible, as the generated features are subject to significant fluctuations. On one hand, the noise emissions from the test bench infrastructure interfere with the measurements. On the other hand, the transmission path through the air attenuates the noise emissions of the test object compared to direct structure-borne sound measurements. However, the superimposed noise emissions of the test bench predominantly affect lower frequency bands, as good results can still be achieved using a filter band within the envelope analysis. Nonetheless, it must be considered that additional acoustic disturbances, such as music, conversations, or other nearby machinery, could induce interference, which was not addressed in this study. It is plausible that the impact on structure-borne sound measurements is lower due to the attenuation over the air transmission path, but this needs to be validated for the application, as disturbances can also be introduced through the structure, directly affecting structure-borne sound measurements. This also applies to other sources of interference, such as fluctuating oil temperatures or system changes due to maintenance work. Thus, further research is crucial to determine the influence of these fluctuations, especially oil temperature, on model performance. This research should investigate whether the impact of temperature variation is negligible compared to other variations (e.g., manufacturing tolerances), allowing models to leverage inherent fluctuations for robustness. If not, the need for temperature-specific training, feature integration, or test bench modifications should be evaluated.

In contrast to systematic disturbances, the repeatability of the measurements with the available data is well verified. It is confirmed that the repeatability of accelerometer 2 on the test bench is higher, as no reapplication is necessary when testing a new unit. The sensor coupling and thus the transmission path of the structure-borne sound does not vary, unlike for accelerometer 1 on the unit, where the variation in sensor application is compounded by the general serial dispersion. Particularly in the planar direction of the test bench flange, robust fault detection is possible across a wider operational range, as hydraulic pulsations have a lower impact here. Overall, it is evident that the advantage of the sensor being consistently applied outweighs the disadvantage of being further from the relevant noise source. This can be attributed to the rigid connection of the unit to the test bench. The dependency of fault detection on operating conditions can also be attributed to hydraulic pulsations: Lower swivel angles result in reduced noise and hydraulic pulsations due to the lower inertia of the displaced oil volume, leading to fewer noise disturbances that superimpose the structure-borne or airborne noise generated by the rolling bearing fault in the whole spectrum. The effects of rotational speed become particularly evident when high reproducibility, as seen with accelerometer 2, is achieved and noise disturbances are filtered out across a wide frequency range, as in the case of envelope analysis. Lower rotational speeds result in reduced excitation of the rolling bearing and the transmission path to the sensor. Consequently, there is a risk that minor faults may be masked by the background noise and the residual disturbances within this filter band. Conversely, high rotational speeds may have a more disruptive impact on accelerometer 1 if the adapter is induced to vibrate, which is application-dependent. Particularly in the z-direction, fault detection may be compromised at high rotational speeds.

The superiority of the envelope-based method, which is reflected in a high quality of fault detection with an AUC exceeding 0.9 across a wide range of operating conditions when considering accelerometer 2, and even with suboptimal sensors such as the microphone, is based on the empirically determined filter band. This filter band characterizes a range where hydraulic pulsations are attenuated by the structure, and bearing resonances are pronounced. However, it is critical to question whether this remains consistent in a serial testing scenario where various variants and sizes of axial piston units are tested. The FFT-based method does not require such prior knowledge and can be generically applied to structure-borne or airborne sound signals. To ensure robust fault detection, it would be prudent to determine the filter band automatically, for example, through ML. A data-driven approach could involve determining anomaly detection models in individual frequency bands and isolating anomalies on a band-specific basis. Furthermore, within the scope of signal processing, there is the potential to create enhanced envelopes (see chapter 2) to calculate bands where hydraulic pulsations are minimized. The kurtogram is not a viable solution, as filter bands selected solely based on spectral kurtosis may be significantly influenced by the frequency peaks of hydraulic pulsations. Heuristic signal processing methods such as EMD offer no added value due to the dominance of hydraulic pulsations, and the cepstrum is also significantly influenced by the harmonics of the piston frequency and thus by hydraulic pulsations.

Outliers or faults in a2y and the envelope-based method, for example, are more effectively detected by the IF compared to the FFT method. For the latter, the location of faulty data instances in the feature space favors better detection by the OC-SVM and the LOF, with a notable strong correlation between the latter two. In this context, it makes sense to identify the faults that are or are not detected by individual models and to refer to the configuration in the feature space, which, however, is not within the scope of this paper. Nevertheless, it should thus be verified whether robust frameworks for fault detection can be constructed using an ensemble of models, a common approach in ML. Whether majority voting (hard voting) or soft voting is more suitable needs to be verified. The use of relatively simple models is justified due to the limited data available, ensuring the development of generalizable models. If more data can be generated, this opens the possibility of employing more complex models based on deep learning, such as autoencoders or convolutional neural networks for anomaly detection. Particularly, the application of the latter in conjunction with the transformation of time series data into image representations holds significant potential. The underlying scenario of semi-supervised anomaly detection facilitates the learning of the nominal behavior of relevant units. In serial testing processes with low variant diversity and high production volumes, model training can be effectively conducted using historical data. However, if the nominal behavior changes over time, this poses a challenge in a semi-supervised scenario, as the robustness of the solution concerning other variants must be reassessed. In such cases, unsupervised methods for outlier detection without model training (see [7]), as well as transfer or cross-domain learning approaches such as siamese neural networks, can provide viable solutions.

7 Conclusion

In this paper, we propose the integration of signal processing methods and ML in a semi-supervised anomaly detection scenario, which can be utilized for fault detection in serial testing processes. The proposal specifically focuses on the analysis of acoustic measurements but can be adapted for other signals as well. Additionally, the pipeline can be used for CM applications. The results indicate that detecting faulty bearings against the background of serial variation is not a trivial task. Operating conditions can have a significant impact on fault detection, with higher rotational speeds and lower swivel angles generally being preferable. When prior knowledge of relevant frequency bands exists, envelope analysis can yield comparatively favorable results even with microphones, demonstrating an improvement in the operating-point-specific median AUC of approximately 0.31 compared to time-based features. However, to obtain robust results in scenarios with a large variety of variants without such knowledge, it is crucial to generate repeatable results with a fixed installed accelerometer on the test bench. Using the envelope-based method, AUC values of at least 0.9 were achieved even at moderate speeds around 800 rpm, with a maximum observed value of 0.987. While the APU-mounted accelerometer achieved a best-case AUC of 0.992, its performance exhibited significantly higher variability depending on the operating point. Consequently, the potential exists to achieve more robust results with the investigated ML techniques through appropriate signal processing using a practicable test bench sensor. In our studies, it was possible to ensure nearly perfect fault detection in the best case considering the mentioned values for the AUC. This corresponds to a high TPR of 95,2% at a FPR of 0%. In Contrast to that, other signal processing methods commonly used in general bearing fault diagnosis like the cepstrum and kurtogram-based analysis as well as EMD are counterproductive for APUs.

References

[1] S. Haug, Ganzheitliche Optimierung einer Axialkolbenpumpe durch bedarfsangepasste Entlastung tribologischer Kontakte, Bd. 95, Karlsruhe: KIT Scientific Publishing, 2022.

[2] P. C. Palavicino, B. Sarlioglu, D. Bobba, W. Lee und T. Minav, „Electrification of hydraulic systems using high.efficiency permanent magnet motors,“ in *12th International Fluid Power Conference*, Dresden, 2020.

[3] M. Tiboni, C. Remino, R. Bussola und C. Amici, „A Review on Vibration-Based Condition Monitoring of Rotating Machinery,“ *applied sciences,* Bd. 12, Nr. 2, 2022.

[4] C. Xiao, H. Tang, Y. Ren, J. Xiang und A. Kumar, „A fault frequency bands location method based on improved fast spectral correlation to extract fault features in axial piston pump bearings,“ *Measurement,* Bd. 171, 2021.

[5] R. B. Randall und J. Antoni, „Rolling element bearing diagnostics - A tutorial,“ *Mechanical Systems and Signal Processing,* Bd. 25, pp. 485-520, 2011.

[6] M. Bengtsson und G. Lundström, „On the importance of combining “the new” with “the old” – One important prerequisite for maintenance in Industry 4.0,“ *Procedia Manufacturing,* Bd. 25, pp. 118-125, 2018.

[7] K. Hendrickx, W. Meert und Y. Mollet, „A general anomaly detection framework for fleet-based condition monitoring of machines," *Mechanical Systems and Signal Processing,* Bd. 139, 2020.

[8] J. Bonart und W.-G. Drossel, Entwicklung einer End-of-Line Körperschallprüfmethodik für hochintegrierte elektrifizierte Antriebstopologien, Wissenschaftliche Scripten, 2024.

[9] J. G. M. Lázaro und C. B. Pinilla, „A methodology for detection of wear in hydraulic axial piston pumps," *International Journal on Interactive Design and Manufacturing,* Bd. 14, pp. 1103-1119, 2020.

[10] S. Gaugel, B. Wu, A. Anand und B. Wu, „Supervised Time Series Segmentation as Enabler of Multi-Phased Time Series Classification: A Study on Hydraulic End-of-Line Testing," in *IEEE 21st International Conference on Industrial Informatics (INDIN),* 2023.

[11] J. Du, S. Wang und H. Zhang, „Layered clustering multi-fault diagnosis for hydraulic piston pump," *Mechanical Systems and Signal Processing,* Bd. 36, pp. 487-504, 2013.

[12] T. Torrika, Bewertung von Analyseverfahren zur Zustandsüberwachung einer Axialkolbenpumpe, Aachen: Rheinisch-Westfälische Technische Hochschule Aachen, 2011.

[13] P. Casoli, M. Pastori, F. Scolari und M. Rundo, „A Vibration Signal-Based Method for Fault Identification and Classification in Hydraulic Axial Piston Pumps," *energies,* Bd. 12, Nr. 5, 2019.

[14] S. Horn, A. Shorbagy, M. Lenz, S. Knoll, M. Winter und J. Weber, „Causal Condition Monitoring for Axial Piston Pumps: Prediction of Different Faults under Variable Operating Conditions," in *2024 International Maha Fluid Power Conference*, West Lafayette, 2024.

[15] O. Gnepper, H. Hitzer und O. Enge-Rosenblatt, „Predictive Diagnosis in Axial Piston Pumps: A Study for High Frequency Condition Indicators Under Variable Operating Conditions," *International Journal of Prognostics and Health Management,* Bd. 14, Nr. 1, 2023.

[16] L. Siyuan, D. Linlin und J. Wanlu, „Study on Application of Principal Component Analysis to Fault Detection in Hydraulic Pump," in *Proceedings of 2011 International Conference on Fluid Power and Mechatronics*, Beijing, 2011.

[17] P. Casoli, M. Pastori und F. Scolari, „A multi-fault diagnostic method based on acceleration signal for a hydraulic axial piston pump," in *AIP Conference Proceedings* , Modena, 2019.

[18] S. Xia, Y. Xia und J. Xiang, „Piston Wear Detection and Feature Selection Based on Vibration Signals Using the Improved Spare Support Vector Machine for Axial Piston Pumps," *materials,* Bd. 15, Nr. 23, 2022.

[19] Y. Zhou, A. Kumar, C. Parkash, G. Vashishtha, H. Tang und J. Xiang, „A novel entropy-based sparsity measure for prognosis of bearing defects and development of a sparsogram to select sensitive filtering band of an axial piston pump," *Measurement,* Bd. 203, 2022.

[20] A. Kumar, C. P. Gandhi, H. Tang, G. Vashishtha, R. Kumar, Y. Zhou und J. Xiang, „Adaptive sensitive frequency band selection for VMD to identify defective components of an axial piston pump," *Chinese Journal of Aeronautics,* Bd. 35, Nr. 1, pp. 250-265, 2022.

[21] Y. Zhou, A. Kumar, C. Parkash, G. Vashishtha, H. Tang, A. Glowacz, A. Dong und J. Xiang, „Development of entropy measure for selecting highly sensitive WPT band to identify defective components of an axial piston pump," *Applied Acoustics,* Bd. 2023, 2023.

[22] A. Kumar, H. Tang, G. Vashishtha und J. Xiang, „Noise subtraction and marginal enhanced square envelope spectrum (MESES) for the identification of bearing defects in centrifugal and axial pump," *Mechanical Systems and Signal Processing* , Bd. 165, 2022.

[23] P. Casoli, A. Bedotti, F. Campanini und M. Pastori, „A Methodology Based on Cyclostationary Analysis for Fault Detection of Hydraulic Axial Piston Pumps," *energies,* Bd. 11, Nr. 7, 2018.

[24] Q. Chao, H. Gao, J. Tao, C. Liu, Y. Wang und J. Zhou, „Fault diagnosis of axial piston pumps with multi-sensor data and convolutional neural network," *Frontiers of Mechanical Engineering,* Bd. 17, Nr. 3, 2022.

[25] Y. Zhu, G. Li, S. Tang, R. Wang, H. Su und C. Wang, „Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN," *Applied Acoustics,* Bd. 192, 2022.

[26] S. Tang, Y. Zhu und S. Yuan, „An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnois of hydraulic piston pump," *Advanced Engineering Informatics,* Bd. 50, 2021.

[27] S. Tang, Y. Zhu und S. Yuan, „Intelligent fault diagnosis of hydraulic piston pump based on deep learning and Bayesian optimization," *ISA Transactions,* Bd. 129, Nr. Part A, pp. 555-563, 2022.

[28] Y. Zhu, L. Guangpeng, R. Wang, S. Tang, H. Su und K. Cao, „Intelligent Fault Diagnosis of Hydraulic Piston Pump Based on Wavelet Analysis and Improved AlexNet," *sensors,* Bd. 21, Nr. 2, 2021.

[29] S. Wang und J. Xiang, „A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps," *Soft Computing,* Bd. 24, pp. 2983-2997, 2020.

[30] S. Tang, Y. Zhu und S. Yuan, „A novel adaptive convolutional neural network for fault diagnosis of hydraulic piston pump with acoustic images," *Advanced Engineering Informatics,* Bd. 52, 2022.

[31] Y. He, H. Tang, Y. Ren und A. Kumar, „A deep multi-signal fusion adversarial model based transfer learning and residual network for axial piston pump fault diagnosis," *Measurement,* Bd. 192, 2022.

[32] N. Keller, A. Sciancalepore und A. Vacca, „Demonstrating a Condition Monitoring Process for Axial Piston Pumps with Damaged Valve Plates," *International Journal of Fluid Power,* Bd. 23, Nr. 2, 2022.

[33] C. Bienefeld, F. M. Becker-Dombrowsky, E. Shatri und E. Kirchner, „Investigation of Feature Engineering Methods for Domain-Knowledge-Assisted Bearing Fault Diagnosis," *entropy,* Bd. 25, Nr. 9, 2023.

[34] R. B. Randall, „A history of cepstrum analysis and its application to mechanical problems," *Mechanical Systems and Signal Processing,* Bd. 3, Nr. 19, 2017.

[35] C. Scheffer und P. Girdhar, Practical Machinery Vibration Analysis and Predictive Maintenance, Elsevier, 2004.

[36] P. Perera, P. Oza und V. M. Patel, „One-Class Classification: A Survey," arXiv, 2021.

[37] F. T. Liu, K. M. Ting und Z.-H. Zhou, „Isolation Forest," in *2008 Eighth IEEE International Conference on Data Mining*, Pisa, Italy, 2009.

[38] M. M. Breunig, H.-P. Kriegel, T. N. Raymond und J. Sander, „LOF: Identifying Density-Based Local Outliers," *ACM SIMOD Record,* Bd. 29, Nr. 2, 2000.

[39] O. Alghushairy, R. Alsini, T. Soule und X. Ma, „A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams," *big data and cognitive computing,* Bd. 5, Nr. 1, 2021.