Article citation info:
Chazy B, Pawlik P. The use of a naive Bayes classifier in continuous monitoring systems for rotating machinery. Diagnostyka. 2025;26(3):2015313. https://doi.org/10.29354/diag/209804.

DIAGNOSTYKA, 2025, Vol. 26, No. 3

e-ISSN 2449-5220
DOI: 10.29354/diag/209804

THE USE OF A NAIVE BAYES CLASSIFIER IN CONTINUOUS MONITORING SYSTEMS FOR ROTATING MACHINERY

Bogumił CHAZY *, **Paweł PAWLIK**

Department of Mechanics and Vibroacoustics, Faculty of Mechanical Engineering and Robotics,
AGH University of Krakow, al. A. Mickiewicza 30, 30-059 Kraków, Poland
* Corresponding author, e-mail: chazy@agh.edu.pl

Abstract

This paper presents a method for diagnosing rotating machines operating under varying conditions based on order analysis, Gaussian mixture models (GMMs) and Bayesian inference. A classifier was constructed based on the values of the amplitudes of the order spectrum and the value of the rotational speed changing due to the variable load of the machine.

An analysis was conducted on the functionality of the method in diagnosing misalignment and unbalance of a drive system consisting of a drive motor and planetary gearbox. The drive train was subjected to a variable load of the main gear of a bucket wheel excavator at a variable oil temperature. In the diagnostic experiment, the method was shown to be highly effective in diagnosing preset system faults. The implementability of the method in embedded systems was also investigated. The method was implemented in a system with a real-time operating system and an FPGA. The method was tested in continuous monitoring mode on a laboratory bench.

Keywords: vibroacoustic diagnostics, machine condition monitoring, real-time operating system, variable load, machine learning

List of Symbols/Acronyms

FFT – Fast Fourier Transform
xo1 – Order No. 1 axis X
yo72 – Order No. 72 axis Y
zo8 – Order No. 8 axis Z

1. INTRODUCTION

Condition monitoring of rotating machinery operating in industrial plants is increasingly important [1–5]. Failure of essential machinery causes production downtime and generates the need for costly unplanned repairs or the replacement of damaged assets. In extreme cases, faulty machinery can lead to loss of health or life of personnel. Early detection of equipment failures is an essential aspect of industrial plants.

In the machinery industry, more and more companies are choosing to implement advanced diagnostic systems that cover entire plants or groups of plants. These solutions often monitor multiple machines simultaneously and use advanced analytical methods to assess their condition. Wireless diagnostic devices or specialised diagnostic software can also be found on the market, which serve as central units for large assemblies of diagnostic equipment. In most cases, these systems are based on the analysis of RMS, peak value or basic statistical parameters. These parameters work well for rotating machines operating under constant conditions. For machines in fluctuating operating conditions, it is necessary to use more sophisticated diagnosis methods based on synchronous methods, which additionally should not be sensitive to load changes [6,7]. Such methods are being developed in the scientific literature, e.g. for diagnosing gearboxes [8–12], rolling bearings [13–15], or entire systems [16–18].

For the method to be implemented in industrial continuous monitoring systems, it must be relatively simple and not have too many parameters depending on the object to be diagnosed. Commercially available systems operate as stand-alone devices performing measurement and simple condition assessment based on alarm thresholds, while in-depth data analysis is performed by cooperating computer programs [19–23]. No universal devices offer analysis of machines operating under variable load conditions among these systems. There are dedicated devices for machines operating in a limited load range (wind turbines, hydroelectric power plants) based on the ISO 20816-3:2022 standard [24]. Furthermore, these devices require

Received 2025-03-14; Accepted 2025-08-22; Available online 2025-09-10

manual input of alarm threshold values, which could be automated using machine learning methods. What is lacking is a solution that is designed for machines operating under variable loads, that allows automatic calibration to work with any rotating machine using machine learning methods, and one in which all the analysis is done automatically within a single device - the analyser - and the software is only used to receive the analysis results.

This paper presents a machine diagnosis method that performs fault identification in real-time (during measurement). Unlike existing solutions dedicated to machines operating within a certain load range, the proposed method is dedicated to machines operating under variable conditions. The method is based on synchronous analysis, Gaussian mixture models (GMM) and the naive Bayes classifier algorithm. The method is implemented in a measurement-diagnostic system, which performs the measurement, analyses the data and presents the analysis results to the user. The system also performs an automatic evaluation of the machine's condition.

This article in Chapter 2 describes the proposed method, whose algorithm is adapted for implementation in continuous diagnostic monitoring systems. Chapter 3 describes the analysis of the functionality of the proposed method on a laboratory bench. Chapter 4 presents the measurement data analysis: algorithmic selection of diagnostically significant symptoms, and a naive Bayes classifier training as part of automatic machine condition assessment. Chapter 5 proposes the implementation of the method in a device with a real-time operating system and an FPGA. The final chapter provides a summary of the method discussed and the results of the analysis of the diagnostic experiment data.

2. THE PROPOSED DIAGNOSTIC METHOD

The proposed method was developed for machines operating under variable load conditions.

The method assumes initial measurements to adjust the algorithm to specific machines and defects. Measurements should be taken for the normal operating condition of the machine and defective conditions.

Measurement of the vibration-acoustic signals takes place synchronously. Due to the occurrence of variable loads, an order analysis is performed - resampling the measurement signals based on the tachometer signal so as to obtain signals synchronous to the shaft rotation [25]. The resampled vibroacoustic signals are used to determine the one-sided order power spectra [26] according to the formula:

$$G_{AA}(Order) = \begin{cases} S_{AA}(Order), Order = 0 \\ 2S_{AA}(Order), Order \neq 0 \end{cases} \quad (1)$$

where:

$$S_{AA}(Order) = \frac{FFT(A) \cdot \overline{FFT(A)}}{N^2} \quad (2)$$

$\overline{FFT(A)}$ denotes the complex conjugate of the discrete Fourier spectrum of the synchronous signal A with N samples. The signal length N is the length of the frame into which the resampled measurement signal is segmented in the rotation angle domain. The number of N samples is chosen in such a way as to obtain an appropriate spectrum resolution.

The spectral components of the order power spectra for the subsequent frames constitute the features used further for analysis and, ultimately, for classification.

The selection of those orders – symptoms that are diagnostically significant is done algorithmically. For this purpose, a recursive feature elimination (RFE) algorithm was used using a random forest classifier [27,28].

Among the features from the reduced set, those features for which the Spearman correlation coefficient between them and the other features indicates a strong correlation are removed.

A naive Bayes classifier makes a diagnostic decision based on selected symptoms, whose distributions are described by Gaussian mixture models. It is a classifier that is simple to implement and takes up little of the device's operating memory. It makes a classification decision based on the Bayes theorem, which states the probability of condition X_i occurring with a measured symptom S [29]. The S-symptom can be a parameter obtained by measurement that provides information on the technical condition of the machine.

$$P(X_i|S) = \frac{P(X_i) \cdot P(S|X_i)}{P(S)} \quad (3)$$

$$P(X_i|S) = \frac{P(X_i) \cdot P(S|X_i)}{\sum_{k=1}^{K} P(S|X_k) \cdot P(X_k)} \quad (4)$$

where:

- $P(X_i|S)$ is the conditional probability of state X_i occurring after measuring symptom S,
- $P(S|X_i)$ is the conditional probability that the observed symptom value S corresponds to the state X_i,
- $P(X_i)$ is the probability of occurrence of state X_i,
- $P(S)$ is the probability of a symptom S,
- K is the number of states X_i.

The diagnostic decision falls on that state X_i, belonging to the vector of states $\boldsymbol{X}$, for which the probability $P(X_i|S)$ is highest. Given that the denominator of Equation 4 for each state is equal, and assuming that the occurrence of each state is equally likely, in practice, the diagnostic decision falls on the state for which the probability $P(S|X_i)$ is the highest. We assume that the a priori probability of each state is the same due to the absence of information indicating otherwise. However, Bayesian inference allows this probability to be taken into account if it is available.

If there is more than one symptom, then $\boldsymbol{S}$ is a vector of symptoms. Assume that the components of the vector $\boldsymbol{S}$ are independent. The probability $P(\boldsymbol{S}|X_i)$ is then equal to [30]:

$$P(\boldsymbol{S}|X_i) = \prod_{j=1}^{J} P(S_j|X_i) \quad (5)$$

where J is the dimension of the vector **S** (vectors are shown in bold).

In practice, however, because the product of a high number of small numbers can be numerically unstable and quickly reach 0, the natural logarithm of the probability (log-likelihood) is used [31]. This allows the product operation to be replaced by a sum operation.

$$ln\left(P(\boldsymbol{S}|X_i)\right) = \sum_{j=1}^{J} ln\left(P\left(S_j|X_i\right)\right) \quad (6)$$

The logarithm is a monotonic function, so instead of comparing probabilities to make a classification, log-likelihood values can be compared.

The probability $P(S_j|X_i)$ can be obtained using a simple probability distribution model obtained from the measurement data. A Gaussian mixture model can be used as such a model [28, 32]. Because the symptom distributions cannot be described by a normal distribution (Chapter 4), they are approximated by the sum of three normal distributions. The non-normality of the distributions of the order spectrum components is due to the varying load, which affects the amplitude of the vibrations. For the sum of the distributions – a Gaussian mixture, the log-likelihood $ln(P(S|Xi))$ equals:

$$ln\left(P(\boldsymbol{S}|X_i)\right) = \sum_{j=1}^{J} ln\left(\sum_{n=1}^{N} w_n f_{\mu_{n,j}\sigma^2_{n,j}}\left(S_j\right)\right) \quad (7)$$

where:

- $f_{\mu n,j,\sigma^2 n,j}(x)$ is the probability density function of a normal distribution with mean $\mu_{n,j}$ and variance $\sigma^2{}_{n,j}$ for the n-th probability density function in the GMM and the j-th feature,
- w_n is the weight of the n-th probability density function in the Gaussian mixture model,
- N is the number of components of the Gaussian mixture model.

In summary, the proposed machine diagnostic method consists of the following steps:

1. Synchronous measurement of acceleration and rotational speed - for proper and faulty states of the machine.
2. Resampling of acceleration signals to obtain signals synchronous to shaft rotation.
3. Calculation of order power spectra.
4. Algorithmic selection of diagnostically significant features.
5. Elimination of features that are highly correlated with other features.
6. Approximation of feature distributions using Gaussian mixture models.
7. Training of a naive Bayes classifier using previously derived statistical models.
8. Using a trained model to classify the condition of the machine.

3. DIAGNOSTIC EXPERIMENT

3.1. Laboratory bench

The laboratory bench shown in Figure 1 is a model of a loaded industrial machine. It consists of a drive motor followed by a planetary reduction gear with a ratio of 4. The gear consists of a ring gear with 72 teeth, a sun gear with 24 teeth and three planetary gears with 24 teeth each. A claw coupling connects the output shaft of the gearbox to the load motor. The motor speeds are controlled by three-phase inverters.

The machine load is set by reducing the speed of the load motor. When the motors are working in unison, the speed of the planetary gearbox output shaft and the set speed of the load motor are equal, and the machine runs without load. When, on the other hand, the set speed of the load motor is less than a quarter of the speed set for the driving motor, the driving motor will have to overcome the load. The load will be greater the greater the difference between these speeds.

Fig. 1. Photography of the laboratory station: 1 - control inverters, 2 - drive motor, 3 - planetary gearbox, 4 - claw coupling, 5 - brake motor

The motors are controlled by an application that uses the analogue outputs of the NI USB 6002 measurement card to generate voltage signals, with which it applies the appropriate voltage to the control inputs of the inverters. The application controls the motors so that the driving motor operates under varying load conditions. In the first phase, the motors are ramped up linearly so that no load is present. A load profile is then applied to the input of the inverter responsible for the load. To replicate the load variation to that which would occur during operation of an industrial machine, this signal is replayed from the recorded speed waveform of the planetary gearbox input shaft of the KWK 1500s excavator during its operation. Variations in rotational speed were caused by varying loads on the bucket wheel mounted on the gearbox output. This signal was scaled to the plenary gearbox under test on the laboratory bench. By inflicting this type of load, we can investigate how the damage inflicted on the laboratory bench will affect the diagnostic signals of the planetary gearbox. Investigating this under industrial conditions is impossible due to the impossibility of introducing damage into an industrial machine.

An example of a machine load profile in the form of voltages applied to the control inputs of the inverters is shown in Figure 2.

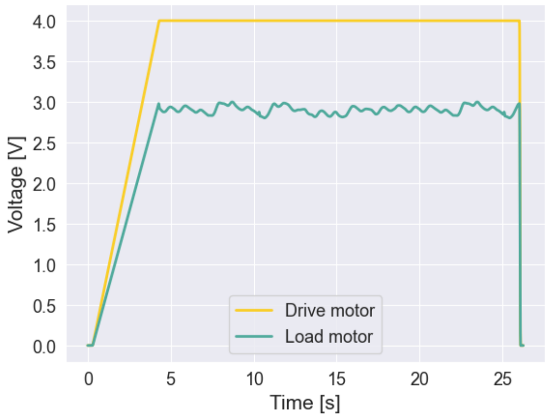

Fig. 2. Set load profile

Because of the variable load, the shaft speed on the laboratory bench is also variable. It oscillates around 710 rpm +/- 10 rpm. An example of the shaft rotational speed during one load cycle is shown in Figure 3.

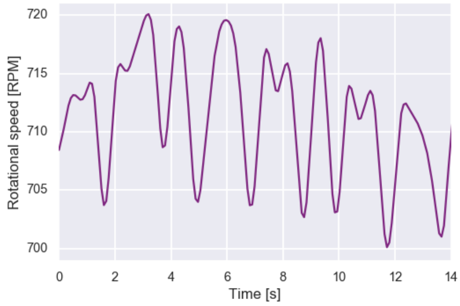

Fig. 3. The shaft rotational speed

The sensor locations on the test bench are shown in Figure 4.

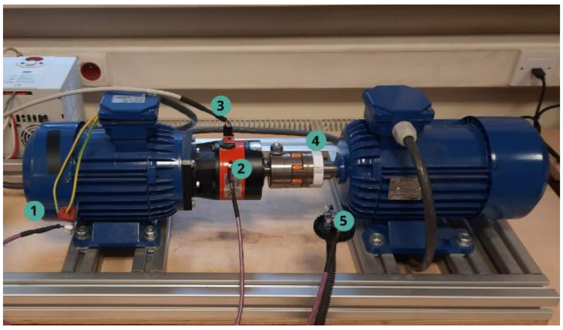

Fig. 4. Sensor locations on the test bench:
1 - current sensor, 2 - thermometer,
3 - accelerometer, 4 - rotation marker,
5 - tachometer

The sensors used were a MEMS ADXL335 acceleration sensor, ARD 7662 tachometer, TA12-200 AC current sensor, LM35 temperature sensor.

The sensors were connected with the analogue inputs of the NI myRIO-1900 device.

3.2. Experiment

The measurement system described above was first used to take measurements. Measurements were needed to collect data for three different machine states: an operational state and two defects - to then analyse them and train the classifier.

The measurement was made for each of the three states of the machine:

- undamaged condition - the machine was working without defects,
- unbalanced state - with a 0.68 g unbalanced mass added to the coupling of 0.05 m diameter,
- misalignment condition - with 2 mm thick washers placed under the front feet of the driving motor, causing shaft misalignment.

Due to the varying operating conditions (i.e. loads and temperatures), signals of 15 min were recorded. This allowed the entire heating time of the machine to be recorded. The acceleration and rotation marker signals were recorded at a sampling rate of 5kHz to record the entire bandwidth of the acceleration sensor (up to 1.6kHz) in accordance with the Nyquist-Shannon sampling theorem and with a sufficient margin. The temperature and current drawn by the drive varied over the same range for each state. Order power spectra, temperature, and current were used to analyse the data as symptoms.

4. DATA ANALYSIS

As part of the data analysis, symptoms that carry diagnostic information were algorithmically selected and used to train a classifier model. The order power spectrum for each of the three directions contains orders from 0 to 96, which, together with the RMS value of the current and the temperature of the gearbox, gives 293 features.

To assess whether the distributions of features within the groups are normal, a Shapiro-Wilk test was performed for each feature and each group [33,34]. The following hypotheses were posed:

- null hypothesis - the distribution of the feature can be described by a normal distribution,
- alternative hypothesis - the distribution of the feature cannot be described by a normal distribution.

A normality test of the distribution was performed for a significance level of $\alpha = 0.05$. Table 1 shows the p-values for each group. Due to the large number of characteristics, only the five characteristics with the highest p-value values among the three states were included.

Table 1. P-value values for the features

Feature	operational	imbalance	misalignment
zo1	4,74e-06	2,89e-02	1,70e-09
yo2	2,83e-15	2,70e-02	3,09e-16
zo9	2,07e-03	8,58e-10	1,98e-34
zo6	9,71e-04	1,53e-25	4,50e-39
zo11	1,43e-10	9,09e-04	1,05e-35

Out of the 293 features and three groups, in all cases, p-value < α, so the alternative hypothesis must be accepted - symptom distributions cannot be described by a normal distribution.

4.1. Feature selection

A recursive feature elimination (RFE) algorithm with a random forest classifier was used to initially reduce the number of features [27,28]. The target feature count was 100. The following algorithm was then performed for the remaining features: the dataset was divided into 10 parts, and then, similarly to the cross-validation method, the random forest classifier was trained 10 times using nine of the 10 parts of the dataset, each time excluding a different part of the data from training. After each iteration of this algorithm, the feature significance returned by the model was saved and averaged for each feature after completion. The average significance for each feature was thus obtained. Figure 5 shows these significances with the standard deviation and cumulative significance shown. Based on this data, the 10 most significant features were selected.

Features related to temperature and current appeared to be of little relevance in terms of diagnosis because these parameters varied in the same way for each case and were not affected by damage.

Among the features derived from the vibration acceleration signals, one would expect to find fault information in the 1st, 2nd and 4th harmonics of the shaft, while the significance analysis of the features did not show these harmonics as significant. This may be due to the variable loading of the system. However, in this case, the most significant features appeared to be the amplitudes of the order spectrum from the gear meshing band. According to [35] the meshing frequency of a planetary gearbox is equal to:

$$f_z = f_1 \frac{z_3 z_1}{z_3 + z_1} \tag{8}$$

where f_1 is the rotational frequency of the input shaft, z_3 is the number of teeth in the ring gear, and z_1 is the number of teeth in the sun gear. The gear meshing frequency on the laboratory bench is 18 times the rotational frequency of the input shaft, which corresponds to 72 order after considering the gear ratio.

The next step was to exclude features that are highly correlated with other features. Correlated features carry the same information, leading to feature redundancy and causing overfitting [36]. Using the naive Bayes classifier, one of the assumptions is that the features are independent. To avoid overfitting and meet the model assumptions, it is necessary to exclude correlated features.

The Spearmann correlation coefficient for each pair of features was calculated from a set limited to ten features. The matrix of Spearman's correlation coefficients is shown in Figure 6. Some features were then excluded so as to eliminate strong correlations - with a coefficient above 0.7 [37].

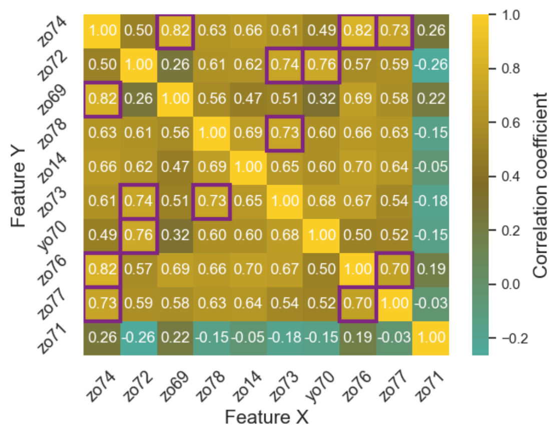

Fig. 6. Spearman correlation coefficient matrix

Based on the above analysis, the set was reduced to six symptoms: zo72, zo69, zo78, zo14, zo77, zo71. An example of the distribution of a pair of features derived from the reduced set is shown in Figure 7. The figure shows that, based on the selected features, it is possible to differentiate the individual groups - the technical states of the machine.

4.2. Classifier training

To train the classifier, each feature of the reduced set was first normalised to an interval between 0 and 1. The minimum and maximum values for each feature were saved so that the new data could later undergo the same transformation. Next, the data was split into a training set and a test set in a 3:1 ratio - the first set was used to train the model, and the second set was used to evaluate it.

Training consisted of determining a Gaussian mixture model for each state and for each feature based on the training dataset. The number of components of the mixture was set equal to 3, and the covariance matrix was chosen to be diagonal.

For the number of states K, the number of symptoms J and the number of components of the Gaussian mixture model N, the following were obtained:

- K vectors of weights $\boldsymbol{w}$ of dimension N,
- K average matrices M of dimension $N \times J$,
- K variance matrices $\boldsymbol{\Sigma}$ of dimension $N \times J$.

The classifier was then evaluated using data from the test set. Each sample from the dataset was classified by assigning it to the class Xi with the highest log-likelihood value according to the formula x. The confusion matrix of the classifier is shown in Figure 8, and its metrics are presented in Table 2.

Table 2. Classifier metrics

Machine condition	Precision	Sensitivity	Support
Imbalance	0.99	0.96	187
Misalignment	1.00	1.00	188
Good condition	0.96	0.99	188
Accuracy			0.98

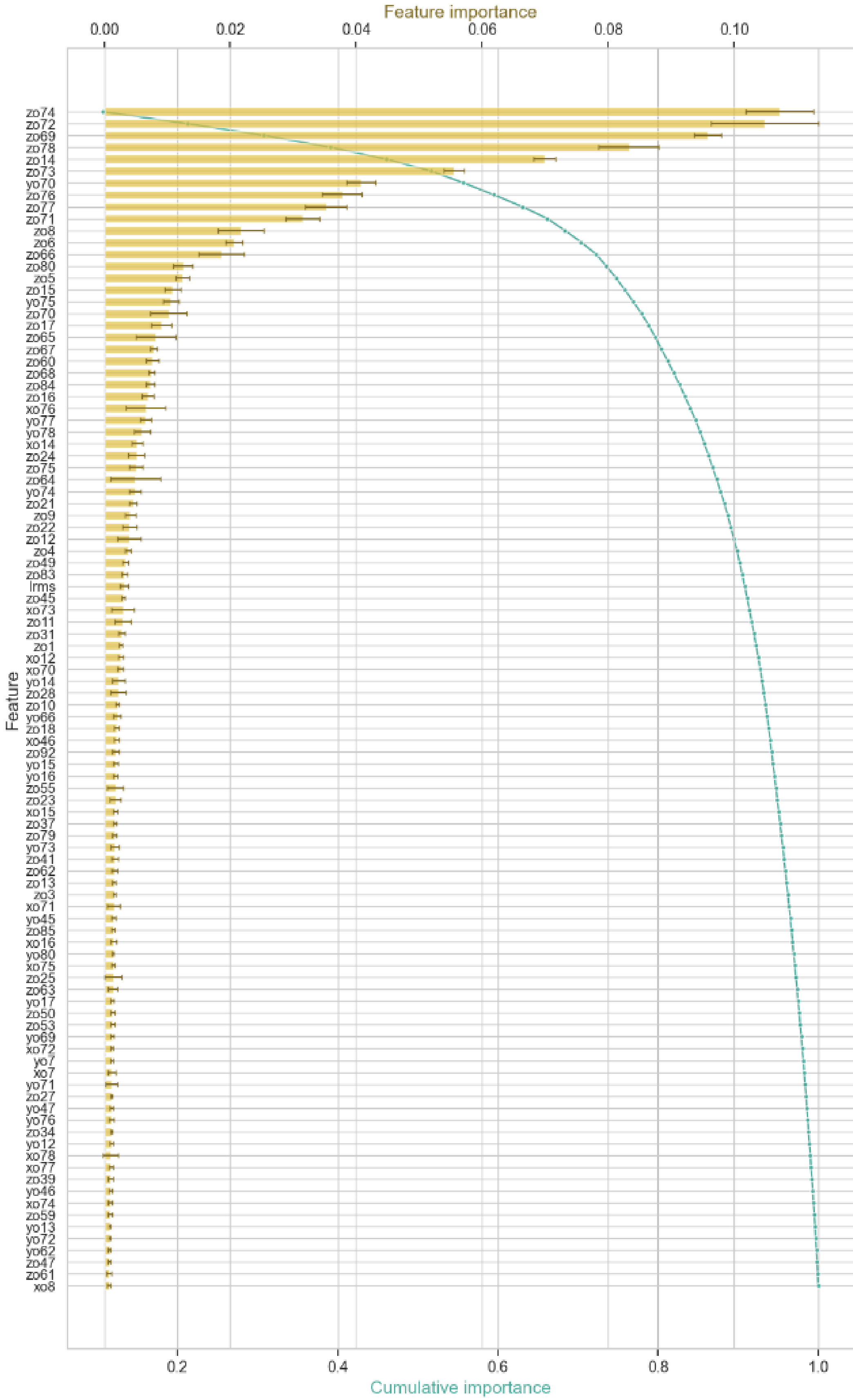

Fig. 5. Feature importance ranking

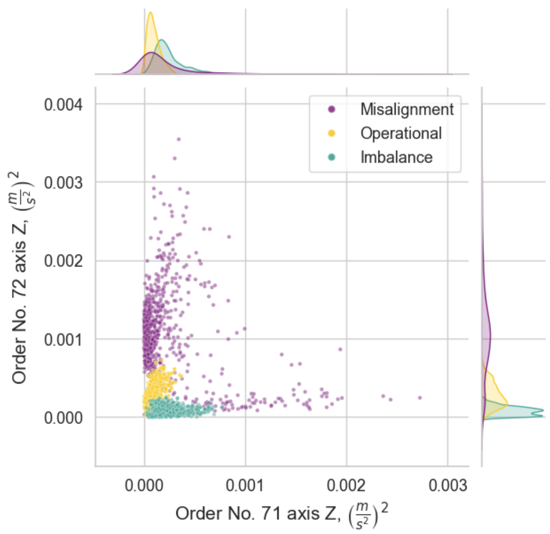

Fig. 7. Example of the distribution of a pair of features from the reduced set

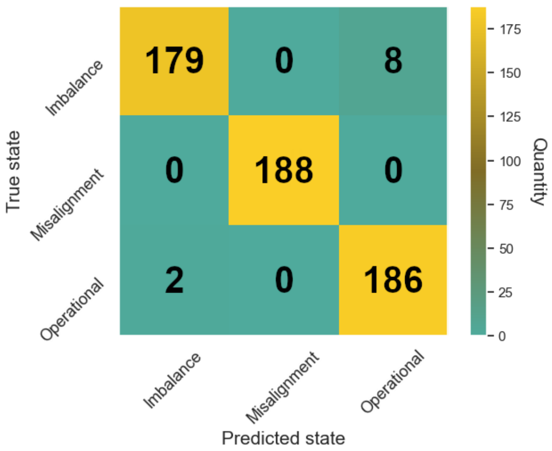

Fig. 8. Confusion matrix

The resulting classifier most accurately identifies the misalignment state. As for the other two states: imbalance and operational state, the model achieved lower metrics. Eight samples from the imbalance state were classified as operational state and three samples from the operational state were assigned to the imbalance class. Despite this, the classifier achieves very high metrics - the accuracy of the classifier was 0.98, which means that of all classifications, 98% were correct.

5. IMPLEMENTATION OF THE METHOD IN THE RT SYSTEM PROPOSITION

The classifier prepared in this way was implemented on an NI MyRIO device. This device has a controller with a real-time operating system (RTOS) and an FPGA circuit, which has its industrial equivalent [38]. The RT operating system ensures reliable system operation and is a commonly used system in diagnostic equipment. The measurement and diagnostic system created is shown in diagram form in Figure 9.

The software that comprises the diagnostic system: the program running on the FPGA, the program running on the RT OS and the application running on the Windows PC were implemented using the LabVIEW development system.

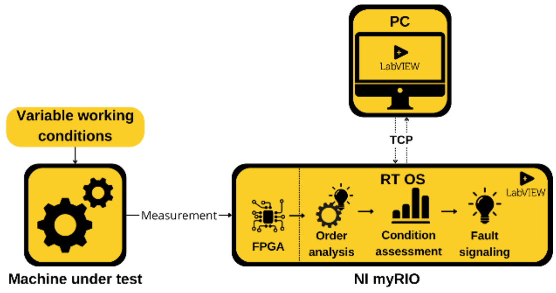

Fig. 9. System diagram

The basic assumption of the developed diagnostic system is that it is intended to operate under varying load conditions during the operation of a machine. Measurement of the vibroacoustic signals, processing of these signals, and evaluation the machine's technical condition are carried out by the NI MyRIO device.

The program implemented on the FPGA is responsible for reading the values of the signals from the A/D converter. Once per time period which is the inverse of the specified sampling frequency, it reads the instantaneous value of the measured signals and sends them to the RT via a DMA FIFO queue.

The measured values are:

- vibration acceleration in three directions,
- shaft rotational speed.

The program running on the real-time operating system is based on a four-thread architecture. Its structure is shown in Figure 10. Each thread is a state machine and communication between them is carried out through queues.

The thread used to receive control data from the PC manages the other threads. The PC application communicates with this thread using TCP/IP protocol.

The thread responsible for recording the measurement signals receives the data from the DMA FIFO queue that was sent from the FPGA and then sends it to the analysis thread.

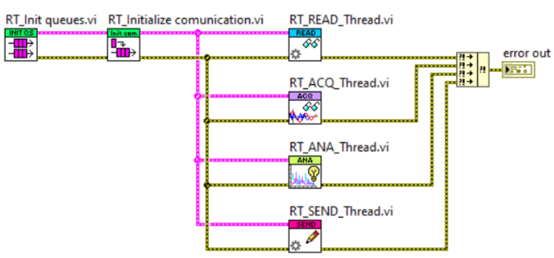

Fig. 10. RT program architecture

In the analysis thread, as a part of the processing of the measurement signals, due to the occurrence of variable loading, an order analysis is performed -

according to the method described in Chapter 2. In order to assess the technical condition of the machine, those amplitudes of the order spectrum are used - symptoms whose diagnostic relevance has been indicated by the analysis described in Chapter 3.

Based on the selected symptoms, the Naive Bayes Classifier, whose training is described in Chapter 2, makes a diagnostic decision.

The automatic evaluation of the machine's condition takes place ~~even~~ before the data is sent from the program running on the RT to the PC-based application.

The measurement data, the results of the order analysis and the diagnostic decision are sent to the next thread, which sends them to the application running on the PC via TCP.

Using the Windows PC software, it is possible to communicate with the NI MyRIO device to set the measurement configuration and the processing parameters and to receive the raw and processed measurement data and the diagnostic decision. The classification result is displayed on the program's user interface.

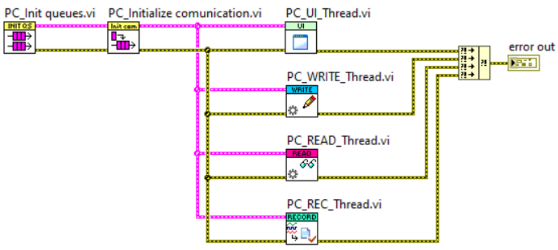

Fig. 11. PC application architecture

The application running on the PC is based on four threads: a user interface thread, which is used to control the system and display the data and analysis results, and state machine-based threads for sending data to the RT, receiving data from the RT and writing data to a file. The architecture of the application is shown in Figure 11.

The device was tested on the laboratory bench described in section 3.1 regarding the continuous operation of the device in an application for continuous diagnostic monitoring. The myRIO device recorded measurement signals while simultaneously analysing and assessing the technical condition of the object under test.

6. SUMMARY

This paper presents a new method for diagnosing machines operating under varying conditions. The method is based on the order power spectrum of the vibration acceleration signal and a naive Bayes classifier. A functionality analysis of the proposed method was carried out on a laboratory test bench for planetary gearboxes. The method showed 98% effectiveness in classifying misalignment, unbalance and correct machine operation. A prototype of the diagnosis system with the myRIO-1900 computing unit was built. The use of the prototype in continuous diagnostic monitoring systems was tested.

Traditional diagnostic systems are based on the ISO 20816-3:2022 standard [24], according to this standard, rotating machines should be operated under nominal operating conditions and at a fixed operating temperature during diagnosis [30].

The developed system is adapted to work during the operation of a machine operating in variable conditions. It measures the diagnostic signals, processes them - among other things, it carries out an order analysis and performs an automatic evaluation of the technical condition of the machine and presents the measurement data and the classification result to the user (using TCP/IP protocol).

Dedicated systems for variable-speed machines based on synchronous methods already exist in industry [39], but in most cases, they are inefficient for machines operating under variable load and temperature [6]. Dedicated methods for machines operating under variable conditions are being developed [40] but are not yet implemented in industrial diagnostic monitoring systems.

Automatic evaluation of the machine's condition is achieved by analysing the measurement data to create a 'naive Bayes classifier', which uses a Gaussian mixture model to assign the machine to one of three states: operable, imbalance and misalignment. The fault base can be extended to include other faults that occur in industry.

The analyser is a prototype for an industrially deployable diagnostic system.

The analyser is universal - the classification algorithm can be calibrated to work with any rotating machine. However, this requires measurements with artificially introduced faults. It must be taken into account that sometimes this is not possible due to the need to stop the machine.

The proposed method requires data from a faulty machine, which can be difficult to obtain in an industrial environment. However, it is possible to define the most common damages such as misalignment, imbalance and train the model with this data. To detect other defects, the method can be adapted to return information about an unidentified machine condition. The classifier, trained for artificially introduced states, will indicate the one of the states for which the probability is highest, even if the damage is not included in the training data. Then the probability that the machine belongs to each state will be low. It is possible to improve the method by introducing a probability threshold below which the classifier indicates an unidentified machine state.

Reliable and reproducible measurements are necessary for the analyser to function correctly. Otherwise, a significance analysis of the symptoms may indicate those that do not contribute diagnostic information, thus misleading the classifier, causing disturbances and unstable condition assessment.

Further work is planned with a more complex analysis that considers signals describing the machine's load and operating conditions.

Source of funding: *The project was financed by the Polish Ministry of Science and Higher Education [project No. 16.16.130.942].*

Author contributions: *Research concept and design, B.C., P.P; Collection and/or assembly of data, B.C.; Data analysis and interpretation, B.C.; Writing the article, B.C., P.P.; Critical revision of the article, P.P.; Final approval of the article, B.C, P.P.*

Declaration of competing interest: *The authors declare no conflict of interest.*

REFERENCES

1. Wang T, Han Q, Chu F, Feng Z. Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox : A review. Mech Syst Signal Process. 2019;126:662–85. https://doi.org/10.1016/j.ymssp.2019.02.051.
2. Lei Y, Lin J, Zuo MJ, He Z. Condition monitoring and fault diagnosis of planetary gearboxes: A review. Measurement. 2014; 48:292–305. https://doi.org/10.1016/j.measurement.2013.11.012.
3. Klekot G, Dąbrowski Z, Dziurdź J. 50 years of scientific thought on machine diagnostics in Poland and its influence on maritime applications. Polish Maritime Research. 2024; 31(4):161–73. https://doi.org/10.2478/POMR-2024-0060.
4. Luo X, Ning J, Wu M. Design and analysis of wind turbine fault diagnosis system based on convolutional neural network. Diagnostyka. 2025;26(1):2025113. https://doi.org/10.29354/diag/201334.
5. Goolak S, Gubarevych O, Yurchenko V, Kyrychenko M. A review of diagnostic information processing methods in the construction of systems for operating diagnostics of rotor eccentricity of induction motors. Diagnostyka. 2025;26(1): 2025104. https://doi.org/10.29354/diag/202757.
6. Antoni J, Randall RB. Differential diagnosis of gear and bearing faults. J Vib Acoust. 2002; https://doi.org/10.1115/1.1456906.
7. Pawlik P. Single-number statistical parameters in the assessment of the technical condition of machines operating under variable load. Eksploatacja i Niezawodnosc – Maintenance and Reliability. 2019;21(1):164–9. https://doi.org/10.17531/ein.2019.1.19.
8. Li Y, Feng K, Liang X, Zuo MJ. A fault diagnosis method for planetary gearboxes under non-stationary working conditions using improved Vold-Kalman filter and multi-scale sample entropy. J Sound Vib. 2019;439:271–86. https://doi.org/10.1016/j.jsv.2018.09.054.
9. Guan Y, Liang M, Necsulescu DS. Velocity synchronous bilinear distribution for planetary gearbox fault diagnosis under non-stationary conditions. J Sound Vib. 2019;443:212–229. https://doi.org/10.1016/J.JSV.2018.11.039.
10. Burdzik R, Konieczny Ł, Warczek J, Cioch W. Adapted linear decimation procedures for TFR analysis of non-stationary vibration signals of vehicle suspensions. Mech Res Commun. 2017;82:29–35. https://doi.org/10.1016/j.mechrescom.2016.11.002
11. Bartelmus W, Zimroz R. A new feature for monitoring the condition of gearboxes in non-stationary operating conditions. Mech Syst Signal Process. 2009;23(5):1528–34. https://doi.org/10.1016/j.ymssp.2009.01.014.
12. Liu J, Wang T. Application of vibration signal analysis method in the fault diagnosis of mechanical gearboxes. Diagnostyka. 2025;26(1):2025107. https://doi.org/10.29354/diag/199976.
13. Wang X, Zheng J, Ni Q, Pan H, Zhang J. Traversal index enhanced-gram (TIEgram): A novel optimal demodulation frequency band selection method for rolling bearing fault diagnosis under non-stationary operating conditions. Mech Syst Signal Process. 2022; 172:109017. https://doi.org/10.1016/J.YMSSP.2022.109017.
14. Bote-Garcia JL, Mokhtari N, Gühmann C. Wear monitoring of journal bearings with acoustic emission under different operating conditions. PHM Society European Conference. 2020;5(1):8-8. https://doi.org/10.36001/PHME.2020.V5I1.1202.
15. Jiang L. Construction and application of a bearing fault diagnosis model based on improved GWO algorithm. Diagnostyka. 2024;25(3):1–10. https://doi.org/10.29354/diag/189904.
16. Antoniadou I, Manson G, Staszewski WJ, Barszcz T, Worden K. A time–frequency analysis approach for condition monitoring of a wind turbine gearbox under varying load conditions. Mech Syst Signal Process. 2015;64–65:188–216. https://doi.org/10.1016/j.ymssp.2015.03.003.
17. Mauricio A, Helm D, Timusk M, Antoni J, Gryllias K. Novel cyclo-nonstationary indicators for monitoring of rotating machinery operating under speed and load varying conditions. J Eng Gas Turbine Power. 2022;144(4). https://doi.org/10.1115/1.4049778/1096343.
18. Komorska I, Puchalski A. Rotating machinery diagnosing in non-stationary conditions with empirical mode decomposition-based wavelet leaders multifractal spectra. Sensors 2021;21:677. https://doi.org/10.3390/S21227677.
19. Emerson. AMS wireless vibration monitor. Data Sheet. 2024.
20. Emerson. AMS Asset Monitor Asset Source Interface Product Data Sheet. 2024.
21. Emerson. AMS Machine Works Product Data Sheet. 2023.
22. Siemens Energy. VIB3000 Machinery Protection Highest safety and availability for your plant assets. 2021.
23. Siemens Energy. Condition Monitoring Software Vibroexpert CM500 Technical Description. 2021.
24. ISO/TC 108/SC 2. ISO 20816-3:2022 - Mechanical vibration — Measurement and evaluation of machine vibration — Part 3: Industrial machinery with a power rating above 15 kW and operating speeds between 120 r/min and 30 000 r/min. 2022.
25. National Instruments Corporation. LabVIEWTM Order Analysis Toolkit User Manual. 2005.
26. Zieliński TP. Od teorii do cyfrowego przetwarzania sygnałów. Wydział EAIiE AGH; 2002.
27. Scikit-learn Developers. sklearn. Feature Selection. RFE. 2024.

28. Buitinck L, Louppe G, Blondel M, Pedregosa F, Müller AC, Grisel O, et al. API design for machine learning software: experiences from the scikit-learn project. European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases. 2023. https://doi.org/10.48550/arXiv.1309.0238.
29. Cempel C. Diagnostyka wibroakustyczna maszyn. Wydawnictwo Politechniki Poznańskiej; 1985.
30. Jaramillo VH, Ottewill JR, Dudek R, Lepiarczyk D, Pawlik P. Condition monitoring of distributed systems using two-stage Bayesian inference data fusion. Mech Syst Signal Process. 2017;87:91–110. https://doi.org/10.1016/j.ymssp.2016.10.004.
31. Robinson E. Introduction to likelihood statistics. 2016.
32. Scikit-learn Developers. GaussianMixture. 2024.
33. The SciPy Community. scipy.stats.shapiro. 2024.
34. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. https://doi.org/10.1038/s41592-019-0686-2.
35. Bartelmus W, Zimroz R. Vibration spectra characteristic frequencies for condition monitoring of mining machinery compound and complex gearboxes. Mining Science. 2011;133(40):17–34.
36. Mamdouh Farghaly H, Abd El-Hafeez T. A high-quality feature selection method based on frequent and correlated items for text classification. Soft comput. 2023;27(16): 11259–74. https://doi.org/10.1007/S00500-023-08587-X/FIGURES/10.
37. Akoglu H. User's guide to correlation coefficients. Turk J Emerg Med. 2018;18(3):91–3. https://doi.org/10.1016/j.tjem.2018.08.001.
38. National Instruments Corporation. User guide and specifications NI myRIO-1900. 2018.
39. Jaramillo VH, Ottewill JR, Dudek R, Lepiarczyk D, Pawlik P. Condition monitoring of distributed systems using two-stage Bayesian inference data fusion. Mech Syst Signal Process. 2017;87:91–110. https://doi.org/10.1016/J.YMSSP.2016.10.004.
40. Braun S. The synchronous (time domain) average revisited. Mech Syst Signal Process. 2011;25(4):1087–102. https://doi.org/10.1016/J.YMSSP.2010.07.016.
41. Randall RB. Vibration-based diagnostics of gearboxes under variable speed and load conditions. Meccanica. 2016;51(12):3227–39. https://doi.org/10.1007/S11012-016-0583-Z

Bogumił CHAZY, recieved Master's degree in Faculty of Mechanical Engineering and Robotics from AGH University of Cracow, Cracow, Poland, in 2024. He is a doctoral student, and as part of his PhD, he is working on developing machine learning methods for the diagnosis of rotating machinery. He is a Certified LabVIEW Developer.
e-mail: chazy@agh.edu.pl

Paweł PAWLIK, is an Associate Professor at the Department of Mechanics and Vibroacoustics, AGH University of Krakow, Poland. His research focuses on the development of measurement systems and advanced methods for signal analysis. He is a member of the Polish Society of Technical Diagnostics (PTDT) and a certified LabVIEW Architect.
e-mail: pawlik@agh.edu.pl