DIAGNOSTYKA, 2026, Vol. 27, No. 2

e-ISSN 2449-5220
DOI: 10.29354/diag/225192

MACHINE LEARNING-BASED DIAGNOSIS OF SLIDING BEARINGS IN ROLLING MILL DRIVE SYSTEMS UNDER SEVERE OPERATING CONDITIONS

Volodymyr KLITNOI [1], Serhii HUBSKYI [2], Viktor KIRIK [1], Andrii YAVTUSHENKO [2], Dmytro BORODIN [1]

[1] Department Machine Components and Hydropneumatic Systems,
Institute of Education and Science in Mechanical Engineering and Transport,
National Technical University «Kharkiv Polytechnic Institute», Ukraine
[2] Department Computer Modeling and Integrated Forming Technologies,
Institute of Education and Science in Mechanical Engineering and Transport,
National Technical University «Kharkiv Polytechnic Institute», Ukraine
[*] Corresponding author, e-mail: Volodymyr.Klitnoi@khpi.edu.ua

Abstract

Rolling mill drive bearings operate under extremely severe conditions, including high loads, elevated temperatures, and intense vibration, which leads to accelerated wear and premature failure. In industrial practice, the technical condition of such bearings is typically monitored indirectly using temperature and vibration measurements. However, these approaches do not ensure reliable prevention of emergency failures, as evidenced by multiple incidents in metallurgical plants caused by bearing overheating and damage. To address these limitations, this paper proposes a diagnostic approach based on vibration analysis combined with machine learning techniques. A diagnostic model using the Random Forest algorithm is developed to classify the technical condition of sliding bearings into three states: normal, warning, and critical. The model utilizes a feature vector composed of vibration and operational parameters, including RMS vibration, peak vibration, standard deviation, dominant frequency, temperature, load, rotational speed, and operating time since last maintenance. The model was preliminarily evaluated using a numerically generated dataset based on industrial measurements, rolling mill technical specifications, and established diagnostic threshold values. The obtained results should be considered as a preliminary assessment of the proposed approach and require further validation using extended real industrial datasets.

Keywords: sliding bearings, rolling mill, condition monitoring, Random Forest, machine learning-based diagnostics

List of Symbols/Acronyms

RMS – root mean square;
PJSC – public joint-stock company;
TCI – technical condition index;
V_{RMS} – root mean square value of vibration [mm/s];
V_{max} – peak vibration value [mm/s];
V_{std} – standard deviation of vibration [mm/s];
f_{dom} – dominant frequency of the vibration spectrum [Hz];
T – bearing temperature [°C];
T_q – transmitted torque [kNm];
n – rotational speed [rpm];
t_{op} – operating time since the last maintenance [hours];
k – class label [normal, warning, or critical condition];
M – total number of decision trees in the ensemble;
$h_m(X)$ – prediction of the m-th decision tree for the input feature vector X;
δ – indicator function;
V_v – vibration velocities measured in vertical direction;
V_h – vibration velocities measured in horizontal direction;
V_a – vibration velocities measured in axial direction.

1. INTRODUCTION

Rolling mills are critical components of metallurgical production, where the reliability of drive systems directly affects productivity, product quality, and operational safety [1]. Bearings used in electric drives of rolling mill stands operate under extremely severe conditions, including high cyclic loads, elevated temperatures, shock impacts, and intensive vibration. These factors significantly accelerate wear pro-cesses and increase the risk of premature failure [2].

In industrial practice, the technical condition of sliding bearings in rolling mill drive systems is typically assessed using indirect monitoring methods, primarily based on temperature and

Received 2026-05-19; Accepted 2026-06-29; Available online 2026-06-26

vibration measurements. Although such approaches are widely implemented, they do not provide sufficient reliability in detecting early-stage defects or preventing emergencies. In particular, temperature monitoring often reflects already advanced degradation, while conventional vibration indicators may not fully capture complex defect interactions in heavily loaded friction units.

Real operational conditions reveal that bearing damage rarely occurs in a single isolated form. Instead, multiple defect mechanisms develop simultaneously. These include wear under contaminated lubrication conditions, scuffing caused by over-load, and fatigue failure due to dynamic edge loading and misalignment [2]. Each of these defects negatively affects lubrication regimes, reduces load-carrying capacity, and promotes further degradation. Fatigue-related damage is especially critical, as it can lead to fragmentation of the antifriction layer and catastrophic bearing failure.

Industrial evidence confirms the limitations of existing diagnostic approaches. For example, several emergency events in metallurgical production were associated with overheating and damage of rolling mill drive bearings, leading to unplanned downtime and significant economic losses [2]. These cases demonstrate that conventional monitoring techniques are insufficient for ensuring reliable operation under real production conditions.

Therefore, there is a need to develop more advanced diagnostic methods capable of identifying complex defect patterns at early stages and supporting timely maintenance decisions. In this context, vibration analysis remains one of the most informative sources of data on dynamic processes in friction units. However, its effective-ness can be significantly enhanced through the application of modern data-driven approaches.

Machine learning techniques provide new opportunities for improving the accuracy and robustness of bearing diagnostics. By analysing multidimensional datasets that include both vibration and operational parameters, such methods enable auto-mated classification of technical condition and identification of hidden patterns that are difficult to detect using traditional approaches.

This paper proposes a diagnostic model for sliding bearings in rolling mill drive systems based on the Random Forest algorithm. The scientific and practical contribution of the study consists in adapting an interpretable integrated approach to the diagnostics of sliding bearings operating under severe industrial conditions, taking into account real damage mechanisms, combined defects, and operational parameters. The proposed model combines vibration and operating features and transforms probabilistic classifier outputs into a technical condition index suitable for maintenance decision support in industrial environments.

2. LITERATURE REVIEW

A wide range of approaches has been developed for the condition monitoring and fault diagnosis of bearings in rotating machinery. Among the most commonly used methods are temperature monitoring, vibration analysis, oil analysis, acoustic emission, and other non-destructive testing techniques [3]. Temperature-based methods are widely applied in industrial practice due to their simplicity and ease of implementation; however, they typically reflect already advanced stages of degradation and are not sensitive to early fault development. Oil analysis techniques, including wear debris monitoring, provide valuable information about lubrication conditions and material wear, but they often require laboratory analysis and are not always suitable for real-time diagnostics in harsh industrial environments. Acoustic emission methods have demonstrated high sensitivity to incipient defects and microscopic damage processes, as they are capable of capturing high-frequency elastic waves generated by surface interactions. Nevertheless, their practical application is limited by high sensitivity to noise, complex signal interpretation, and the need for specialized equipment.

In addition to the above approaches, various hybrid and alternative monitoring techniques have been proposed, including ultrasonic methods, electrical signal analysis, and surface condition evaluation [4]. Despite their advantages, these methods often face challenges related to implementation complexity, cost, or limited applicability under variable operating conditions [3]. As highlighted in several review studies, vibration-based monitoring remains one of the most widely used and effective techniques for bearing diagnostics due to its ability to capture dynamic responses of mechanical systems and detect changes in operating conditions at relatively early stages [3, 5]. For this reason, vibration-based methods are extensively applied in industrial practice, particularly in rolling mill equipment, where they serve as a primary tool for monitoring the condition of bearings operating under heavy loads and harsh environments.

A number of recent studies have focused specifically on rolling element bearings operating in rolling mill equipment under severe industrial conditions. For instance, in [6], experimental investigations demonstrated the effectiveness of combining vibration analysis with acoustic emission techniques for condition monitoring and fault detection in rolling mill roller bearings. While vibration signals provide useful information about bearing health and operating conditions, they are often affected by mechanical noise and do not allow accurate quantification of defect size. In contrast, acoustic emission signals exhibit higher sensitivity to defect evolution and enable better characterization of fault severity under varying load and speed conditions, although their practical application is associated with increased complexity and signal processing requirements.

In another study [7], intelligent fault diagnosis methods based on deep learning have been proposed for rolling mill bearings operating under complex industrial conditions and limited data availability. In particular, an improved deep belief net-work combined with spectral kurtosis-based feature extraction, principal component analysis, and swarm intelligence optimization algorithms was shown to enhance diagnostic accuracy and generalization capability under small-sample conditions. The results also indicate that appropriate sensor placement significantly influences diagnostic performance, while future improvements are expected through the use of multi-sensor data fusion to further increase the robustness and reliability of fault detection in rolling mill systems.

In vibration-based diagnostics of rolling element bearings, machine learning techniques, including advanced deep learning methods, are widely applied. In study [8], blind source separation using Independent Component Analysis was employed to decompose vibration signals into statistically independent components, improving the identification of fault-related frequency features under noisy operating conditions and enabling effective integration into machine learning pipelines. In [9], hybrid diagnostic frameworks combining handcrafted features with deep learning models were proposed, where feature fusion from time–frequency representations significantly enhanced classification accuracy, although at the cost of increased model complexity and computational requirements.

Classical machine learning approaches also remain relevant, as demonstrated in [10], where methods such as decision trees, support vector machines, k-nearest neighbours, and Random Forests, often combined with principal component analysis, were successfully applied for fault classification, offering a balance between interpretability and computational efficiency. At the same time, deep learning-based approaches, such as the gram matrix and multiscale convolutional neural networks model presented in [11], have shown high robustness to noise and strong classification performance due to multi-scale feature extraction capabilities; however, their practical implementation is often constrained by the need for large labelled datasets and high computational resources.

Recent studies have proposed hybrid intelligent approaches combining signal processing, optimization algorithms, and machine learning to improve the accuracy of rolling element bearing's fault diagnosis. In particular, study [12] presents a comparative analysis of knowledge-based and data-driven approaches, showing that frequency-domain vibration analysis and identification of characteristic bearing frequencies are essential for early fault detection. Knowledge-based methods offer faster implementation and higher interpretability, while data-driven approaches provide greater flexibility but require large datasets and higher computational effort, highlighting the trade-off between interpretability and data dependency.

In study [13], a hybrid method integrating variational mode decomposition, kernel extreme learning machine, and an improved whale optimization algorithm demonstrated high diagnostic accuracy through optimized feature extraction and model parameter tuning. However, the use of multiple processing and optimization stages increases computational complexity, which limits the applicability of such approaches in real-time industrial monitoring systems.

However, the application of vibration-based diagnostics to sliding bearings is significantly more challenging. Unlike rolling element bearings, where defects generate distinct periodic impulses, sliding bearings operate under hydrodynamic lubrication conditions, resulting in more complex and nonlinear dynamic behaviour. Defect development in such systems is often associated with gradual degradation processes, including lubricant film instability, wear, misalignment, and thermal effects, which do not always produce clear spectral signatures.

As shown in study [14], lubrication conditions are strongly coupled with vibration response in journal bearings. Experimental investigations demonstrate that lubricant type significantly affects vibration characteristics, including RMS, variance, and frequency-domain responses, while also influencing system stability under different operating speeds. These findings highlight the critical role of lubrication regimes and operating conditions in the interpretation of vibration signals and the reliability of condition monitoring for sliding bearings.

To address the complexity and nonlinearity of vibration signals in sliding bearings, advanced signal processing techniques have been increasingly applied. In particular, study [15] explores adaptive signal decomposition methods for analysing non-stationary vibration signals in journal bearings. Approaches based on empirical mode decomposition combined with Hilbert spectral analysis enable the extraction of intrinsic mode functions and instantaneous frequency components, providing more detailed identification of fault-related features compared to traditional Fourier- and wavelet-based methods. These results demonstrate that mode decomposition-based techniques are effective for capturing localized time-scale characteristics of vibration signals and improving fault detectability under complex operating conditions.

Recent studies have demonstrated that the integration of vibration analysis with machine learning techniques is an effective approach for fault diagnosis of journal bearings. In particular, methods based on power spectral density combined with classifiers such as k-nearest neighbour and artificial neural networks enable reliable identification of operating conditions, including oil starvation and severe wear, with artificial neural networks generally providing higher classification accuracy [16]. Furthermore, review studies highlight significant progress achieved through the application of hybrid and ensemble models, including deep neural networks, which improve diagnostic accuracy and robustness by up to 15–20%, although challenges related to data availability and model generalization remain [17]. Experimental investigations on hydrodynamic journal bearings confirm that wear processes lead to noticeable changes in vibration spectra, including the appearance of characteristic harmonics, which can be effectively used for fault detection; among various classifiers, Random Forest demonstrates superior performance compared to k-nearest neighbour and support vector machine [18]. Similarly, comparative analyses of multiple machine learning models indicate that ensemble methods, such as Gradient Boosting and Random Forest, consistently achieve the highest classification accuracy (above 95%) and improved reliability in distinguishing different wear conditions [19]. Overall, these results confirm that ensemble learning approaches, particularly Random Forest-based models, represent a promising direction for vibration-based diagnostics of sliding bearings due to their high accuracy and robustness.

It should also be noted that, despite the significant progress achieved in vibration-based diagnostics and machine learning applications, several important research gaps remain. First, relatively few studies focus specifically on sliding bearings operating in rolling mill drive systems under severe industrial conditions. Second, many existing diagnostic models are developed and preliminarily evaluated using laboratory datasets, which may not fully reflect the complexity and variability of real industrial environments. Finally, relatively limited attention has been given to the integration of probabilistic classifier outputs with practical condition indices suitable for maintenance decision support and industrial operation. These limitations motivate the development of interpretable and practically applicable diagnostic approaches adapted to real operating conditions of rolling mill equipment.

3. RESEARCH METHODOLOGY

The aim of this study is to develop and validate a machine learning-based diagnostic approach for condition assessment of sliding bearings in rolling mill drive systems operating under severe conditions, using vibration and operational data for reliable classification of bearing health states.

To achieve this goal, it is necessary to solve the following tasks:

- analyse the operating conditions and post-failure damage characteristics of bearing liners in sliding bearings of rolling mill drive systems;
- analyse existing vibration-based and data-driven diagnostic approaches for sliding bearings;
- develop a diagnostic model based on the Random Forest algorithm for classifying the condition of sliding bearings;
- perform numerical validation of the proposed diagnostic model using experimental or industrial datasets.

3.1. Post-failure analysis of bearing liners in rolling mill sliding bearings

Industrial experience confirms that plain bearings in rolling mill drive systems are prone to failures caused by severe operating conditions, including high loads, elevated temperatures, and dynamic impacts. This is evidenced, in particular, by a number of accidents recorded at PJSC Zaporizhstal, where overheating and damage to plain bearings in rolling mill electric drives led to emergency shutdowns and unplanned downtime.

Figure 1 shows several bearings that have been removed from service after accidents. In particular, bearing № 11 of the lower horizontal roll drive motor supports both the armature and the spindle and is located between them. Another critical component is bearing № 7 of the upper horizontal roll drive motor. These bearings operate under severe load conditions that combine radial, axial and dynamic loads.

Fig. 1. Rolling mill drive motor plain bearings

A detailed analysis of one case from the emergency installations shows that during a routine inspection of the equipment in machine room No. 1, a temperature increase of up to 58 °C was detected at control point № 3 of bearing № 11 of the lower rolling motor. Despite normal oil supply conditions (maximum flow rate according to sensors), a further increase in temperature was observed. Within a short period, the temperature increased to 60 °C, which led to the decision to stop the rolling process. After the stop, a rapid increase in temperature to 82 °C was recorded, indicating the development of a critical failure.

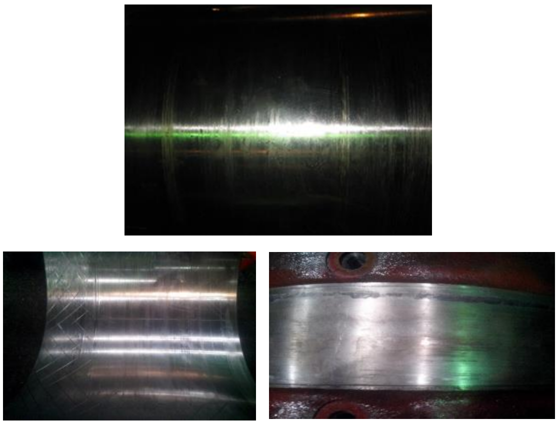

Fig. 2. Damaged bearing liner of № 11 after failure: evidence of babbitt deformation, improper contact pattern, and wear of the thrust flange

Post-accident inspection of bearing liner № 11 (Fig. 2) revealed several characteristic defects, including traces of babbitt smearing on the armature neck, incorrect contact patch configuration, increased wear of the bearing flange, formation of material deposits, and rotation of the thrust element. At the same time, the parameters of the lubrication system (oil level and ring rotation) remained within normal limits. The main cause of failure was determined to be the loss of the hydrodynamic oil film due to displacement of the bearing liner under axial impacts.

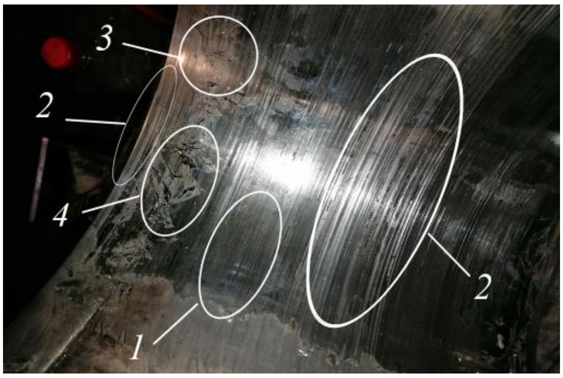

Fig. 3. Typical combined defects of a journal bearing liner

Further analysis of the defective bearing liners (Fig. 3) shows that plain bearing failures rarely occur as a single isolated defect. Instead, a combination of degradation mechanisms is usually observed, including: (1) running-in under clean lubricant conditions; (2) abrasive wear caused by lubricant contamination; (3) burrs during overloading; and (4) fatigue damage due to dynamic edge loading and misalignment. Most of these defects significantly degrade lubrication conditions, reduce load-bearing capacity, and accelerate further wear. Fatigue-related damage in particular is the most critical, as it leads to cracking, spalling, and ultimately delamination of the anti-friction layer, which can lead to catastrophic bearing failure.

According to the established technical criteria, unacceptable conditions for plain bearings are excessive wear of the babbitt layer, cracking, spalling of the antifriction material and damage to the bearing flanges. These defects are often associated with increased vibration levels and excessive clearances, which further destabilize the lubrication regime and contribute to the development of failures.

Thus, the presented analysis demonstrates that failures of plain bearings in rolling mill drive systems are characterized by complex, multifactorial degradation processes that are difficult to detect at early stages using conventional monitoring methods. This emphasizes the need for improved diagnostic approaches that can detect combined defect patterns and assess the condition of bearings under real operating conditions. In this context, the application of vibration diagnostics in combination with machine learning methods, in particular a diagnostic model based on the Random Forest algorithm, is a promising solution for reliable classification of bearing conditions and prevention of catastrophic failures.

3.2. Machine learning-based diagnostic model for sliding bearings using the Random Forest algorithm

Based on the above analysis, a diagnostic model based on the Random Forest algorithm is proposed for assessing the condition of plain bearings in rolling mill drive systems. The model uses vibration and operating parameters as input data and classifies the technical condition into three categories: normal, warning, and critical.

The effectiveness of vibration-based machine learning models strongly depends on the selection and quality of input features used to represent the system state [20]. The input to the model is defined as an 8-dimensional feature vector:

$$X = [V_{RMS}, V_{max}, V_{std}, f_{dom}, T, T_q, n, t_{op}]^T \quad (1)$$

The diagnostic significance of the selected input parameters is summarized in Table 1.

In addition to classifying the condition, the model estimates the probability of belonging to each class:

$$P(\text{class} = k \mid X) = \frac{1}{M}\sum\delta(h_m(X) = k) \quad (2)$$

These probabilities allow evaluating the confidence level of the model's decision and are widely used in ensemble-based classification systems for uncertainty assessment [21].

For practical decision-making, TCI is introduced:

$$TCI = 0 \cdot P(\text{ Normal } \mid X\,) + 1 \cdot P(\text{ Warning } \mid X\,) + 2 \cdot P(\text{ Critical } \mid X\,) \quad (3)$$

The TCI varies in the range from 0 to 2 and provides an integral quantitative assessment of bearing condition. The following thresholds are recommended:

$TCI < 0.4$– normal condition (continued operation);

$0.4 \leq TCI < 1.3$– warning (scheduled maintenance required);
$TCI \geq 1.3$– critical condition (immediate shutdown required).

Table 1. Diagnostic significance of input features

Feature	Diagnostic significance
RMS vibration	Overall level of vibration energy, wear development
Peak vibration	Local impacts, lubricant film instability, scuffing
Standard deviation	Instability of vibration signal, variability of friction regime (Stribeck effect)
Dominant frequency	Changes in dynamic behaviour, harmonic components, misalignment
Temperature	Increased friction, lubrication degradation, overheating
Torque	Severity of operating conditions and accelerated wear factor
Rotational speed	Normalization of vibration behaviour relative to shaft rotation
Time since maintenance	Accumulated operational degradation of the tribological unit

The threshold values were selected empirically based on the probabilistic output distribution of the trained classifier applied to the simulation dataset, ensuring minimal overlap between adjacent condition classes.

The training dataset is formed as a set of pairs (feature vector, condition class). During model development, the dataset is divided into training (80%) and testing (20%) subsets to ensure objective evaluation of classification performance. This data splitting strategy is widely used in machine learning applications as it provides a balanced trade-off between model training efficiency and reliable performance assessment, ensuring sufficient data for learning while maintaining an independent dataset for validation and generalization testing [23].

The proposed model is characterized by ease of implementation, as it does not require complex signal preprocessing. The ensemble nature of the Random Forest algorithm provides noise immunity and reduces the impact of measurement uncertainties. An important advantage is the interpretability of the model, which allows assessing the relative importance of input features in the diagnostic process. In addition, the model is adaptive and can be updated with new data without complete retraining, which makes it promising for industrial application after validation in real time.

4. RESULTS AND DISCUSSION

4.1. Dataset generation assumptions

The validation of the proposed diagnostic model was performed using a numerically generated dataset intended to reproduce realistic operating conditions of sliding bearings in rolling mill drive systems. The generated data should not be considered as a fully experimental industrial dataset; instead, it represents a simulation-based approximation constructed using available industrial measurements, technical specifications, and known diagnostic criteria.

The normal operating ranges of vibration and temperature parameters were defined based on real measurements obtained from bearing units of a reversible rolling mill (stand 1680 PJSC Zaporizhstal). The technical operating parameters, including rotational speed, motor power, and load ranges, were also derived from the actual characteristics of the rolling mill drive system.

The warning and critical condition ranges were formed using industrial vibration diagnostic thresholds adopted in metallurgical practice together with known failure scenarios of sliding bearings, including lubrication degradation, wear development, thermal overload, and misalignment effects.

To improve the realism of the dataset, Gaussian noise was added to the generated samples to account for measurement uncertainty and environmental disturbances. In addition, partial overlap between adjacent condition classes was intentionally introduced to reproduce the gradual nature of bearing degradation and avoid unrealistic class separability.

The purpose of the generated dataset is to provide a preliminary feasibility assessment of the proposed diagnostic approach and to evaluate the applicability of the Random Forest classifier under controlled yet physically justified conditions. The presented validation should therefore be considered as an initial stage preceding further investigation using extended real industrial datasets.

The rolling mill drive operates under the following nominal conditions: rotational frequency of 500 rpm, motor power of 3250 kW, and torque in the range of 40–165 kNm. These parameters were used as a basis for defining the ranges of operational variables in the dataset. Additionally, vibration data (V_v, V_h, and V_a) were collected from four bearing units to establish baseline values corresponding to normal operating conditions. Measurements were performed at PJSC Zaporizhstal following standard industrial monitoring protocols. Vibration sensors were mounted on the bearing housings to capture data in three mutually perpendicular axes (vertical, horizontal, and axial), while temperature sensors were placed on the housing surface adjacent to the bearing insert. The resulting measurements provided the basis for defining the normal operating ranges of the diagnostic parameters.

To form a scalar diagnostic feature, the overall vibration level was calculated as:

$$V_{RMS} = \sqrt{V_v^2 + V_h^2 + V_a^2} \tag{4}$$

The peak vibration value V_{max} and standard deviation V_{std} were either obtained directly from the signal or estimated based on statistical relationships with RMS values. These parameters reflect the amplitude and variability of vibration signals and are sensitive to defect development in sliding bearings.

The f_{dom} was assigned based on the rotational frequency and its harmonics, reflecting typical vibration behavior of sliding bearings under different defect conditions.

According to industrial vibration diagnostics standards, the condition of bearing units can be classified into four categories: normal (up to 1.8 mm/s), acceptable (up to 4.5 mm/s), warning (up to 11.2 mm/s), and dangerous (above 11.2 mm/s). In this study, these categories were generalized into three classes used in the proposed model: normal, warning, and critical. The “acceptable” range was merged with the normal state, while the “dangerous” range corresponds to the critical condition.

To ensure realistic representation of degradation processes, the feature ranges for each class were defined based on physical understanding of sliding bearing behaviours. In particular, degradation is associated with an increase in vibration amplitude, temperature rise, changes in dominant frequency, and higher variability of signals caused by lubrication breakdown, wear, and misalignment.

The input feature vector includes vibration and operational parameters are presented in Table 2.

Table 2. Feature ranges used for simulation-based dataset generation

Parameter	Normal	Warning	Critical
$\boldsymbol{V_{RMS}}$, mm/s	0.1 – 2.0	2.0 – 8.0	8.0 – 15.0
$\boldsymbol{V_{max}}$, mm/s	0.5 – 3.0	3.0 – 12.0	12.0 – 25.0
$\boldsymbol{V_{std}}$, mm/s	0.05 – 0.8	0.8 – 3.0	3.0 – 8.0
$\boldsymbol{f_{dom}}$, Hz	8 – 12	10 – 25	20 – 60
$\boldsymbol{T}$, °C	20 – 50	50 – 75	75 – 90
$\boldsymbol{T_q}$, **kNm**	40 – 100	80 – 140	120 – 165
$\boldsymbol{n}$, rpm	480 – 520	450 – 550	400 – 600
$\boldsymbol{t_{op}}$, h	0 – 500	500 – 2000	2000 – 5000

Figure 4 illustrates a representative simulated degradation scenario of a sliding bearing. The time histories demonstrate physically meaningful evolution of the diagnostic parameters, where increasing vibration severity is accompanied by rising temperature and TCI values. Small stochastic fluctuations were intentionally retained to reproduce industrial measurement uncertainty and operational disturbances. The figure represents only one synthetic realization of the degradation process and is intended solely for illustrative purposes; the complete classification dataset consists of independent samples generated within the predefined parameter ranges and correlations.

The lower bounds of vibration parameters were derived from real measurements of bearing units under normal conditions, where RMS vibration values ranged between 0.1 and 2.5 mm/s and temperatures did not exceed 25°C. The upper bounds for degraded states were defined based on industrial threshold values and typical failure scenarios, including lubrication loss, surface wear, and thermal overload. The degradation scenarios used for dataset generation were additionally informed by the analysis of real bearing failures discussed in Section 3.1. In particular, the observed cases of overheating, babbitt deformation, abrasive wear, and fatigue damage were used to establish physically meaningful relationships between vibration growth, temperature increase, and progression toward critical operating conditions.

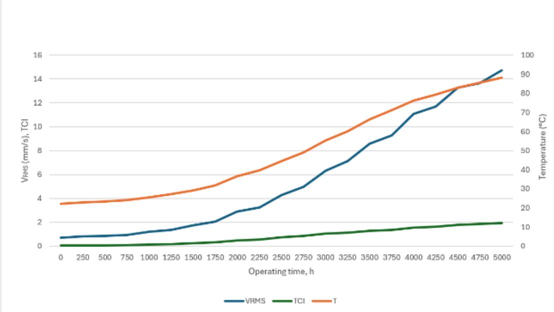

Fig. 4. Generated time-series profiles of vibration (V_{RMS}), temperature, and TCI illustrating the gradual degradation process of a sliding bearing

To improve the physical consistency of the generated data, several correlations between diagnostic parameters were incorporated during dataset generation. In particular, increasing vibration severity was accompanied by higher temperature levels and larger vibration variability (V_{std}), while longer operating times were associated with an increased probability of warning and critical conditions. The dominant frequency ranges were also adjusted according to the expected changes in vibration behaviour under bearing degradation scenarios. Furthermore, partial overlap between adjacent condition classes was intentionally preserved to reproduce the gradual transition between degradation states observed in industrial practice. Although a complete multivariate statistical model based on long-term industrial observations is currently unavailable, the generated dataset preserves the principal physical relationships characteristic of sliding bearing degradation processes.

The final dataset consisted of balanced samples for all three classes and was used for training and testing the diagnostic model. Such an approach allows evaluating the model performance under controlled yet realistic conditions, ensuring the representativeness of the obtained results for industrial applications [24].

4.2. Classification results and model performance

To evaluate the performance of the proposed diagnostic model, the generated dataset was used to train and test the Random Forest classifier.

The model performance was assessed using standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics allow comprehensive evaluation of both overall classification performance and the model's ability to correctly identify each condition class.

The obtained classification results are summarized in Table 3.

Table 3. Classification performance of the proposed diagnostic model

Class	Precision	Recall	F1-score
Normal	0.96	0.95	0.95
Warning	0.94	0.93	0.93
Critical	0.97	0.98	0.97
Overall	–	–	0.95

The overall classification accuracy of the model reached 95.6%, indicating high preliminary classification performance in distinguishing between different technical states of sliding bearings. The highest performance was achieved for the "critical" class, which is particularly important for industrial safety, as it corresponds to conditions requiring immediate shutdown.

To further assess the applicability of the proposed Random Forest-based approach, several commonly used machine learning algorithms were additionally evaluated using the same simulation-based dataset. The comparative analysis included Decision Tree (DT), k-Nearest Neighbours (k-NN), and Support Vector Machine (SVM) classifiers. The obtained results are summarized in Table 4.

Table 4. Comparison of classification performance for different machine learning models

Model	Accuracy	Precision	Recall	F1-score
Decision Tree	0.88	0.87	0.88	0.87
k-NN	0.91	0.90	0.91	0.90
SVM	0.89	0.88	0.89	0.88
Random Forest	0.956	0.95	0.95	0.95

The obtained results demonstrate that the Random Forest classifier significantly outperforms the alternative machine learning models across all evaluation metrics. The ensemble structure of the Random Forest algorithm enables more robust handling of overlapping condition classes and complex nonlinear relationships between vibration criteria and operational parameters.

The single Decision Tree model showed the lowest classification stability (Accuracy of 0.880). This suboptimal performance is explicitly caused by the model's high sensitivity to local variations and its tendency to overfit when dealing with the stochastic nature of the added Gaussian noise. The k-NN classifier demonstrated acceptable performance (Accuracy of 0.910) due to its ability to capture local data density, but it remained inherently sensitive to class overlap zones, where nearest neighbours from adjacent degradation states frequently intermix.

Interestingly, the SVM model (Accuracy of 0.890) underperformed compared to k-NN. While SVM typically handles high-dimensional vectors well, its performance was restricted by the continuous, partially overlapping boundary zones between the "Warning" and "Critical" states, which complicates the construction of an optimal separating hyperplane under noisy conditions.

In contrast, the Random Forest model effectively mitigates these limitations by aggregating the predictions of multiple randomized decision trees. The voting mechanism inherent to the ensemble reduces variance and smooths out the impact of measurement uncertainties. These comparative results firmly validate the selection and suitability of the Random Forest algorithm as the core classifier for vibration-based diagnostics of sliding bearings operating under severe, variable, and noisy industrial conditions of rolling mills.

To further analyze model performance, the confusion matrix for the test dataset is presented in Table 5.

Table 5. Confusion matrix of the diagnostic model

Actual \ Predicted	Normal	Warning	Critical
Normal	38	2	0
Warning	3	35	2
Critical	0	1	39

The confusion matrix shows that most classification errors occur between adjacent states, which is expected due to the gradual nature of bearing degradation. Importantly, misclassification between "normal" and "critical" states is practically absent, confirming the robustness of the proposed model in detecting severe faults.

The results demonstrate that the Random Forest algorithm effectively captures nonlinear relationships between vibration and operational parameters. The ensemble nature of the model contributes to noise resistance and stable performance even under variable operating conditions.

Additionally, the proposed TCI was calculated based on the probabilistic outputs of the Random Forest classifier for the test dataset. The obtained TCI values demonstrated a clear separation between operational states, with low values corresponding to normal conditions, intermediate values indicating warning states, and high values associated with critical bearing conditions. This confirms that the probabilistic interpretation of the ensemble model output can be effectively used for decision support in condition monitoring systems. The TCI-based representation further enhances the practical applicability of the proposed approach by providing a unified scalar indicator for maintenance decision-making in rolling mill drive systems.

4.3. Noise robustness analysis

To evaluate the robustness of the proposed diagnostic model under severe operating conditions, additional tests were performed using different levels of Gaussian noise added to the simulation-based dataset. Three noise levels corresponding to 5%, 10%, and 15% signal perturbation were considered, where the noise amplitude was defined as a percentage of the standard deviation of each corresponding feature in the dataset. These levels were selected to reproduce measurement uncertainty and industrial disturbances typical for rolling mill environments.

To provide a comprehensive robustness assessment, all four machine learning models evaluated in Section 4.2 were tested under identical noise conditions using the same training and testing datasets. The comparative classification accuracy values are summarized in Table 6.

Table 6. Classification accuracy of machine learning models under different noise levels

Model	0% noise	5% noise	10% noise	15% noise
Decision Tree	88.0	84.5	79.2	73.8
k-NN	91.0	87.3	82.4	76.5
SVM	89.0	85.8	81.6	75.1
Random Forest	95.6	94.2	92.1	88.7

The results demonstrate that the Random Forest classifier exhibits the highest robustness to increasing noise levels among all evaluated models. While all methods show a gradual reduction in classification accuracy with increasing noise intensity, the Random Forest algorithm demonstrates an accuracy drop of only 6.9 percentage points at a 15% noise level, compared to 13.9–14.5 percentage points for alternative models.

The superior robustness of the Random Forest classifier is attributed to its ensemble nature, where averaging across multiple decision trees reduces sensitivity to local fluctuations and random disturbances in the input data. In contrast, the Decision Tree classifier exhibited the highest sensitivity to noise due to its overfitting tendencies on perturbed samples. The k-NN model suffered from distorted inter-sample distances in class overlap zones under noisy conditions, while the SVM classifier experienced difficulties in maintaining a stable separating hyperplane between the adjacent Warning and Critical states.

These findings additionally confirm the suitability of the Random Forest algorithm for vibration-based diagnostics of sliding bearings operating under harsh rolling mill conditions.

4.4. Discussion, limitations, and industrial applicability

The obtained results confirm the suitability of the proposed approach for vibration-based condition monitoring of sliding bearings operating under harsh rolling mill conditions.

It should be noted that the presented results are obtained using simulation-based data. While this approach ensures controlled conditions and reproducibility, further validation using real industrial datasets is required to fully confirm the applicability of the model. Nevertheless, the achieved accuracy and consistent classification behavior indicate strong potential of the proposed method for practical implementation in rolling mill drive systems.

Although extensive labelled industrial datasets containing complete histories of bearing degradation are currently unavailable, the proposed approach relies on documented failure cases and post-accident inspections of damaged bearing liners obtained from rolling mill drives at PJSC Zaporizhstal, as discussed in Section 3.1. Consequently, the simulation-based dataset was constructed using physically justified degradation mechanisms and industrial observations rather than purely hypothetical assumptions.

At the same time, several limitations of the present study should be noted. First, the proposed model was preliminarily evaluated using a simulation-based dataset rather than a fully experimental industrial dataset. Second, the condition classes were defined using predefined parameter ranges based on industrial diagnostic thresholds and known failure scenarios, which may not fully capture the complexity of real degradation processes. In addition, the proposed TCI requires further calibration and validation using long-term industrial observations. The present model also does not account for sensor drift, long-term changes in lubricant properties, or certain transient operating effects that may influence vibration behaviour in real production environments. Therefore, further investigation using extended industrial datasets is necessary to improve model robustness and practical applicability.

Although the present study is based on data from a reversible rolling mill (stand 1680 PJSC Zaporizhstal), the proposed methodology is not limited to a specific facility. The selected input features represent general physical indicators of sliding bearing condition and are commonly monitored in rolling mill drive systems. Therefore, the approach can be adapted to other rolling mills equipped with similar bearing units. However, practical implementation at another facility would require model retraining and calibration using local operating and maintenance data to account for differences in equipment design, loading conditions, lubrication regimes, and measurement systems.

As a next step, the proposed diagnostic model is planned to be validated using real industrial data obtained from rolling mill drive systems. In case of successful validation, the approach can be further developed for practical implementation in metallurgical production, which will allow for increased reliability, early detection of faults and reduction of unplanned downtime.

5. CONCLUSIONS

This paper addresses the problem of condition monitoring and fault diagnosis of sliding bearings operating in rolling mill drive systems under severe operating conditions. Based on the conducted analysis and obtained results, the following conclusions can be drawn:

The analysis of real cases in rolling mill drives at PJSC Zaporizhstal, including post-accident inspection of damaged bearing liners, confirms that the degradation of plain bearings is usually caused by a combination of failure mechanisms, such as abrasive wear, burrs and fatigue damage. The analysis of the defective liners revealed complex damage patterns associated with oil film instability, misalignment and dynamic loading, which significantly complicates early fault detection and requires more advanced diagnostic approaches.

Based on the literature review on vibration-based diagnostics and machine learning approaches for bearing condition monitoring, the Random Forest algorithm was selected as the most appropriate method for the proposed study. This choice is due to its ability to effectively work with multidimensional data, noise resistance, high classification accuracy and interpretability of results. In addition, the analysis of current research shows that ensemble methods, in particular Random Forest, demonstrate consistently high results in the diagnostics of bearings, both rolling and sliding.

A diagnostic model based on the Random Forest algorithm has been proposed for assessing the condition of plain bearings in rolling mill drive systems. The model uses an 8-dimensional feature vector of vibration and operating parameters and classifies the technical condition into three states: normal, warning, and critical. It also provides class probabilities for decision confidence assessment and introduces TCI as an integral indicator of bearing condition in the range from 0 to 2. The model is characterized by simplicity of implementation, noise robustness, interpretability, and adaptability, making it suitable for real-time industrial condition monitoring applications.

Numerical validation of the proposed diagnostic model was carried out using a simulation-based dataset constructed on the basis of industrial measurements and technical specifications of a reversible rolling mill (stand 1680, PJSC Zaporizhstal), including vibration levels, temperature conditions, and operational parameters of bearing units. The obtained results demonstrated stable classification performance, reliable identification of critical operating conditions, and robustness under increased noise levels typical for harsh industrial environments.

Although the obtained results are based on simulation-based validation, they indicate the feasibility of the proposed diagnostic approach for sliding bearings operating in rolling mill drive systems rather than providing final industrial verification. The achieved classification performance and stable identification of critical states demonstrate the potential of integrating vibration and operational parameters with an interpretable Random Forest classifier. However, before industrial implementation, the proposed model requires further validation and calibration using extended real industrial monitoring datasets collected under different load, speed, lubrication, and wear conditions.

Source of funding: *The general approach has been partially developed within the research project «Creation of experimental samples of rolling bearings with enhanced performance characteristics based on energy efficiency and durability criteria» (State reg. no. 0125U001616) supported by the Ministry of Education and Science of Ukraine.*

Author contributions: *research concept and design, V.K.; Collection and/or assembly of data, V.Ki.; Data analysis and interpretation, S.H.; Writing the article, V.K.; Critical revision of the article, S.H.; Final approval of the article, V.K.*

Declaration of competing interest: *The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.*

REFERENCES

1. Gaydamaka A, Klitnoi V, Kulik G, Bobrytskyi S, Borodin D. Study of the functioning and wear of the teeth of the intermediate gear of the vertical gearbox rolls of the slabing state – 1150. Diagnostyka. 2025;26(1). https://doi.org/10.29354/diag/200630.
2. Gaydamaka A, Klitnoi V, Kulik G, Bobrytskyi S, Borodin D. Diagnostics of sliding bearings of rolling equipment. Bull Natl Tech Univ KhPI Ser Eng CAD. 2021;(2):22–28.
3. Kannan V, Zhang T, Li H. A review of the intelligent condition monitoring of rolling element bearings. Machines. 2024;12(7):484. https://doi.org/10.3390/machines12070484.
4. Wu G, Yan T, Yang G, Chai H, Cao C. A review on rolling bearing fault signal detection methods based on different sensors. Sensors. 2022;22(21):8330. https://doi.org/10.3390/s22218330.
5. Burda E, Zusman G, Kudryavtseva I, Naumenko A. An overview of vibration analysis techniques for the fault diagnostics of rolling bearings in machinery. Shock Vib. 2022; 2022:6136231. https://doi.org/10.1155/2022/6136231.
6. Nirwan N, Ramani H. Condition monitoring and fault detection in roller bearing used in rolling mill by acoustic emission and vibration analysis. Mater Today Proc. 2021;51. https://doi.org/10.1016/j.matpr.2021.05.447.
7. Peng R, Zhang X, Shi P. Bearing fault diagnosis of hot-rolling mill utilizing intelligent optimized self-adaptive deep belief network with limited samples. Sensors. 2022;22(20):7815. https://doi.org/10.3390/s22207815.
8. Mika D, Józwik J, Ruggiero A. Vibration-based diagnostics of rolling element bearings using the

independent component analysis method. Sensors. 2025;25(23):7371. https://doi.org/10.3390/s25237371.
9. Chennana A, Megherbi AC, Bessous N. Vibration signal analysis for rolling bearings faults diagnosis based on deep-shallow features fusion. Sci Rep. 2025;15:9270. https://doi.org/10.1038/s41598-025-93133-y.
10. Kumar M, M P, B P. Vibration-based condition monitoring of shaft bearing systems using machine learning techniques. Asian J Converg Technol. 2024; 10:1–13. https://doi.org/10.33130/AJCT.2024v10i02.007.
11. Zhang X, Cai S, Cai W, Mo Y, Wei L. A fault diagnosis method for rolling bearing based on gram matrix and multiscale convolutional neural network. Sci Rep. 2024;14(1). https://doi.org/10.1038/s41598-024-83315-5.
12. Iunusova E, Gonzalez M, Szipka K, Archenti A. Early fault diagnosis in rolling element bearings: comparative analysis of a knowledge-based and a data-driven approach. J Intell Manuf. 2023;35:1–21. https://doi.org/10.1007/s10845-023-02151-y.
13. Yuan B, Lu L, Chen S. Research on bearing fault diagnosis based on vibration signals and deep learning models. Electronics. 2025;14(10):2090. https://doi.org/10.3390/electronics14102090.
14. Jebur N, Soud W. Comparative analysis of grease types on bearing performance: time-frequency and statistical investigations. Tribol Finn J Tribol. 2025; 42:24–34. https://doi.org/10.30678/fjt.157405.
15. Babu TN, Devendiran S, Aravind A, Rakesh A, Jahzan M. Fault diagnosis on journal bearing using empirical mode decomposition. Mater Today Proc. 2018;5(2):12993–13002. https://doi.org/10.1016/j.matpr.2018.02.284.
16. Moosavian A, Ahmadi H, Tabatabaeefar A, Khazaee M. Comparison of two classifiers: k-nearest neighbor and artificial neural network for fault diagnosis on a main engine journal bearing. Shock Vib. 2013;20:263–272. https://doi.org/10.1155/2013/360236.
17. Jebur N, Soud W. Recent advances in journal bearings: wear fault diagnostics, condition monitoring and fault diagnosis methodologies. Eng Technol J. 2024;43(1). https://doi.org/10.30684/etj.2024.148997.1737.
18. Bhat H, Yadav R, Bhudhar S, Mandarha A, Phalle V. Vibration analysis of hydrodynamic conical journal bearing and fault prediction using machine learning. 2023. https://doi.org/10.2139/ssrn.4611915
19. Jebur N, Soud W. Wear fault diagnosis in journal bearings using vibration analysis and AI. FME Trans. 2025;53:401–408. https://doi.org/10.5937/fme2503401J.
20. Almutairi K, Wen H, Sinha J. Standardisation of vibration-based parameters for rotor and bearing for machine faults detection using machine learning model. J Vib Eng Technol. 2025;13. https://doi.org/10.1007/s42417-025-02041-y.
21. Breiman L. Random forests. Mach Learn. 2001; 45:5–32. https://doi.org/10.1023/A:1010950718922.
22. Rychlik A. Wheel rim state assessment using modal and geometrical parameters. MATEC Web Conf. 2018;182:01029. https://doi.org/10.1051/matecconf/201818201029.
23. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. 2001.
24. Wang Z, Zhang Q, Xiong J, Xiao M, Sun G, He J. Fault diagnosis of a rolling bearing using wavelet packet denoising and random forests. IEEE Sens J. 2017. https://doi.org/10.1109/JSEN.2017.2726011.

Volodymyr KLITNOI, PhD, Head of the Department of Machine Components and Hydropneumatic Systems of the Educational and Scientific Institute of Mechanical Engineering and Transport of the NTU «KhPI», Kharkiv, Ukraine. Main directions of scientific research: life extension of gears and sliding bearings, data-driven diagnostics, vibration isolation and vibration protection systems for units and machines, active vibration control systems.
e-mail: Volodymyr.Klitnoi@khpi.edu.ua

Serhii HUBSKYI, PhD, Associate Professor, Department of Computer Modeling and Integrated Forming Technologies of the Educational and Scientific Institute of Mechanical Engineering and Transport of the NTU «KhPI», Kharkiv, Ukraine. Main directions of scientific research: digital twins, intelligent monitoring systems for industrial machinery, machine learning in vibration diagnostics, and automation of production processes.
e-mail: Sergii.Gubskyi@khpi.edu.ua

Viktor KIRIK, PhD student of Department of Machine Components and Hydropneumatic Systems of the Educational and Scientific Institute of Mechanical Engineering and Transport of the NTU «KhPI», Kharkiv, Ukraine. Main directions of scientific research: sliding bearings, machine learning techniques.
e-mail: Viktor.Kirik@mit.khpi.edu.ua

Andrii YAVTUSHENKO, PhD student of the Department of Computer Modeling and Integrated Forming Technologies of the Educational and Scientific Institute of Mechanical Engineering and Transport of the NTU «KhPI», Kharkiv, Ukraine. Main directions of scientific research: investigation of the dependence of vibration on the technological parameters of the roll forming process for bent sections.
e-mail: Andrii.Yavtushenko@mit.khpi.edu.ua

Dmytro BORODIN, PhD, Associate Professor, Department of Machine Components and Hydropneumatic Systems of the Educational and Scientific Institute of Mechanical Engineering and Transport of the NTU «KhPI», Kharkiv, Ukraine. Main directions of scientific research: rolling bearings, bolted connections, CAD systems.
e-mail: Dmytro.Borodin@khpi.edu.ua