Wind Turbine Bearing Anomaly Detection Using CMMS Data and Machine Learning

Gabriel Freitas Santos, Helon Vicente Hultmann Ayala,
Florian Alain Yannick Pradelle and Paula Aida Sesini

October 18, 2024

Wind Turbine Bearing Anomaly Detection using CMMS data and Machine Learning

Gabriel Freitas Santos * **Helon Vicente Hultmann Ayala** *
Florian Pradelle * **Paula Sesini** *

* *Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil*

Abstract: The need to anticipate failures in wind turbines has become increasingly urgent. The exponential increase in the number of installed turbines, coupled with the aging of the generation fleet, has intensified the competition to reduce operation and maintenance costs, which means minimizing unplanned downtime and costly major repairs. The aim of this study is to utilize the vibration data available in the Condition Monitoring and Management Systems (CMMS) to identify turbines with significant condition deviations that pose a high risk of failure. The data processing approach using CNN and PCA in the pre-processing stage, along with SVM for health state classification, demonstrated excellent accuracy, above 90 %, for both single turbine and multi-turbine tests, making it suitable for managing wind farms with a large number of turbines.

Keywords: CNN, PCA, SVM, CMMS, Wind Turbine, Bearing, Machine Learning

1. INTRODUCTION

Wind energy consolidation worldwide as a power source in the last decade has put pressure on manufacturers, operators, and academia to advance in predictive fault diagnostics processes, aiming to reduce costs associated with corrective maintenance and unplanned downtime. A decrease of corrective maintenance impact positively the OPEX for Operational Wind Farms and this consideration and investment decrease the Leveled Cost of Energy (LCOE) for new projects (Dao et al., 2019) (Enevoldsen and Xydis, 2019). This advancement has led to increased investment in comprehensive condition monitoring systems for wind turbines, incorporating various sensors, particularly focused on vibration, distributed across the key components of the asset. The availability of a larger volume of data, which enables greater possibilities for fault identification, has also exponentially increased the complexity of managing a wind farm with multiple turbines.
While setting alarms for values outside the normal operating range has proven effective, there is significant potential to anticipate equipment degradation well before reaching the warning thresholds through the use of machine learning models. The academic community has increasingly focused on anticipating deviations in turbine operation. For instance, in (Wen, 2020), SCADA data is utilized for anomaly detection. In (Zhang et al., 2022), Support Vector Machine methods are employed to establish an anomaly score and subsequently classify the wind turbine based on a predefined threshold. Another significant contribution is presented in (Peng et al., 2022), demonstrating a strategy for anomaly identification by evaluating SCADA data windows associated with discrepancy analysis and recorded alarms. However, despite the references presented above, there is a lack of substantial scientific contributions aimed at defining a data pipeline from the asset management perspective. This pipeline seeks not to provide exact fault diagnosis but to focus on identifying deviations and failures with the goal of optimizing costs and enhancing operational performance. This study seeks to advance the field by leveraging data from manufacturers' Condition Monitoring and Management Systems (CMMS), with a focus on vibration analysis and fault diagnosis in Wind Turbine Drivetrains. It will utilize global acceleration and velocity data from sensors on key equipment, as well as pre-processed data such as amplitudes in specific frequency bands and harmonics of the rotational frequency. The article will outline a data pipeline designed to apply a classification model to new turbines with unseen data, enabling fault signature identification and categorization into three health categories. The second section will detail the specific fault examined, describe the fault identification data, and include a sensor location diagram. The third section will explain the tools, methods, and proposed pipeline in detail. The fourth section will present test results, including a discussion and interpretation of the findings. Finally, the fifth section will summarize the main results and contributions of the work.

2. CASE OF STUDY

The object of study in this work is the early detection of failures using the CMMS commonly installed in wind turbines where accelerometers and speed sensors are strategically positioned on the equipment, Figure 1 show all sensors available on the study. Typically, these data are integrated into supervision and monitoring systems that evaluate not only global measurement values but also frequencies defined by experts and pre-processed by the available hardware uptower. Actions are taken preferably after sensor responses exceed highly conservative measurement thresholds, minimizing false alarms managed by the manufacturer that usually deal with this system.

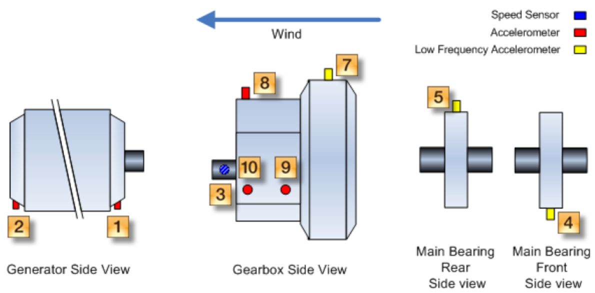


Fig. 1. Sensors Position Diagram

2.1 Available Datasets

Piezoelectric accelerometers with a sensitivity of 100 mV g^{-1}, frequency under 20 kHz, and ±3 dB accuracy (1.5 Hz – 13 kHz) were mounted on the turbine drive train. Synchronously sampled waveforms tracked speed changes, producing narrow spectral lines for variable speed machines, with FFT processing using a Hanning window (Pattabiraman et al., 2015a). Data measured by all sensors are made available every 2 minutes by the system. A total of 12 sensors positioned in the generator, gearbox, main shaft, and nacelle (main frame) provide a total of 112 variables related to vibration, along with an additional 18 process variables such as active power, rotor speed, gearbox oil temperature, bearing temperature. Figure 2 shows a small sample from available variables related to the sensors positioned in the generator as global values for the first to sixth harmonics of the rotational frequency. Among these variables, in addition to global measurements of acceleration and velocity, there is also information regarding the total bandpass between 300 Hz and 700 Hz, the first to sixth harmonics of the rotational frequency and crest factors for both existing generator bearings, gearbox stages and main bearing.

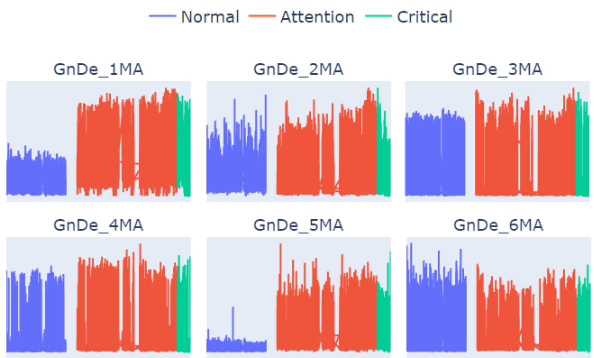


Fig. 2. Available features with health label sample plot

2.2 Generator Bearing Damaged

Fault considered on this study will be a failure diagnostic in the inner race of the generator-coupled bearing. The diagnosis was performed by the manufacturer for two turbines in the same site through frequency spectrum analysis, as shown in Figure 3, which demonstrates the increase in vibration component at a frequency of 96 Hz and its harmonics.

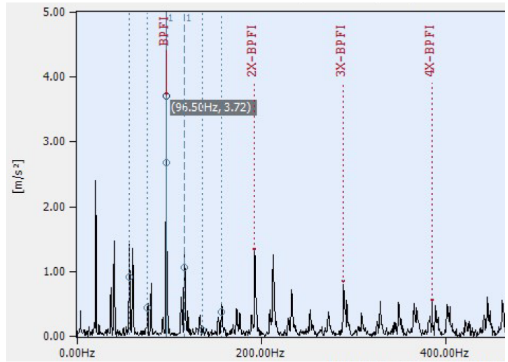


Fig. 3. Generator Bering Fault spectrum excerpt

Table 1. Number of instances for each category in available dataset

	Health State Labels			
Wind Turbine	Normal	Attention	Critical	Total
WTG A	48,912	79,443	11,115	139,470
WTG B	37,392	37,257	34,863	109,512

Dataset available for this study has around 10 months of data for each Wind Turbine Generator (WTG) between 2021 and 2022 and Table 1 shows the amount of instances for each health category. Instances were defined based on operational information and previous failure experiences, taking into account the characteristics of the presented data.

2.3 Background

The impact of this failure type is strongly related to the number of possible failures and consequently the challenge of replacing these components. Additionally, there is the need to maintain a stock of spare parts and the costs involved in these processes. Therefore, those responsible for the operation and maintenance of wind farms with hundreds of turbines need to effectively and quickly monitor the degradation of these components against their original operating condition. This allows for optimization of both planned labor and investments in spare parts.Therefore, the application of the following methodology is crucial for optimizing costs and increasing the performance and availability of wind assets.

3. METHODOLOGY

Large datasets often tend to be highly imbalanced, with a significant number of instances belonging to one category compared to the others. This imbalance can impact both the performance and confidence of the model, as it can bias the results towards identifying a specific category.

3.1 Condensed Nearest Neighbor

To address this issue, the Condensed Nearest Neighbor (CNN) methodology can be employed to randomly select instances from the available dataset, ensuring a balanced representation of all categories.
Based on the clustering algorithm called Nearest Neighbor (NN), CNN identifies clusters and recognizes that all

instances within a cluster share the same characteristics. It then discards the redundant instances, resulting in a condensed datasets that maintains same characteristics and build a balanced representation of all categories.
As showed in (Chou et al., 2006), algorithm begins by selecting the first element, denoted as x_0, from a set of instances grouped as X_n. These instances are randomly chosen. The Condensed Nearest Neighbor (CNN) method then iterates through all members of X_n, adding a member x to the initial group if its nearest neighbor does not belong to the same category as x. This process continues until the group contains an equal number of samples for each category, while still respecting the total number of samples in the minority category as defined by the user-specified parameter. This method proposed by (Hand and Batchelort, 1978) has been utilized even for classification or reducing and balancing datasets, aiming for the best possible performance and accuracy.

In this paper, a CNN was used to balance the proposed dataset while maintaining the previously defined labels.

3.2 Principal Component Analisys

Principal Component Analisys (PCA) transforms data to compare findings across data sets and determine the importance of components. It offers advantages in data analysis, visualization, and computational efficiency, making it valuable for speeding up machine learning algorithms and handling diverse types of data (Salih Hasan and Abdulazeez, 2021). PCA extracts important information from complex data sets, expresses it as principal components, and provides a flexible tool for summarizing data and overcoming duplication in features, while increasing the interpretability and efficiency of the analysis. In some cases, PCA allows for the simplification of large datasets with a high number of features, while retaining a minimum amount of information and achieving good classification results with reduced complexity.
PCA utilizes the covariance matrix and its eigenvalues and eigenvectors to extract uncorrelated information. By analyzing the eigenvectors, it identifies the extent to which information varies and provides insights into the underlying patterns and structure of the data (Hira and Gillies, 2015).

3.3 Support Vector Machine

Support Vector Machine (SVM) is a popular method used for classification or regression algorithms based on statistical learning theory as in (Berk, 2008). It is a machine learning algorithm that uses a training set consisting of data vectors with known class labels. These data vectors, characterized by unique features, are used to design a linear hyperplane that separates different classes. The goal of SVM is to find the optimal hyperplane that maximizes the margin between classes, allowing for effective classification of new, unseen data.
First we need to address the two-class problem. Let's consider a training set consisting of l feature vectors $x_i \in R^n$, where $i (= 1, 2, ..., n)$ represents the number of samples. Each sample is assigned a class label y_i, which can take the value of 1 for one class or -1 for the other class $(i.e., y_i \in -1, 1)$. If the two classes are linearly separable, it means that there exist linear separators called separating hyperplanes. The goal of SVM is to find the optimal hyperplane that maximizes the margin between classes, allowing for effective classification of new, unseen data (Kavzoglu and Colkesen, 2009).

3.4 Criteria and Performance Evaluation Metrics

The confusion matrix format will be used to evaluate the results of the experimental classification. This matrix has axes representing actual labels and predicted labels, and the value in each quadrant refers to the number of records of a specific category that are classified into a certain category, whether correctly or not. These quadrants have specific names according to their qualitative interpretation: True Positives (TP) - the number of positive instances correctly classified; True Negatives (TN) - the number of negative instances correctly classified; False Positives (FP) - the number of negative instances incorrectly classified as positive; and False Negatives (FN) - the number of positive instances incorrectly classified as negative.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

It is possible to calculate KPIs that aid in the interpretation of this matrix and also strengthen the focus on what is to be evaluated or is most important for assessing the solution to the proposed problem. For example, Accuracy defined in Equation 1, a KPI that will be used for evaluation in this work, represents the percentage of correct classifications versus the total number of samples tested.

3.5 Research Contributions

This research proposes preprocessing the dataset, based on the data explored in the case study section, with the aim of reducing redundant data using CNN. Subsequently, PCA is applied to extract a more concise and less complex set of features, resulting in a less complex classification model without compromising the effectiveness of SVM. Figure 4 illustrates the sequential order of the steps involved, from data collection to the evaluation of results. Two experiments will be conducted to compare the effects of using PCA strategy versus not using it. Specifically, the experiments will assess the impact of the PCA strategy on the results, with all other conditions remaining constant.
The routine will be divided into two main tests, each using two models. The first test involves training and validating the model using PCA with data from the same turbine, while the second test uses the original inputs for comparison. Although the validation data is new to the model, it is extracted from the same turbine and the same failure case. Another test will be conducted by training the model with the failure data from one wind turbine and validating the results with data from another turbine. This allows us to assess the scalability of this modeling approach, particularly in evaluating large wind farms.

4. RESULTS AND DISCUSSION

4.1 Original Features

Identification model applied to test data from WTG A, considering all available inputs, yielded satisfactory results, as shown in the Confusion Matrix represented in

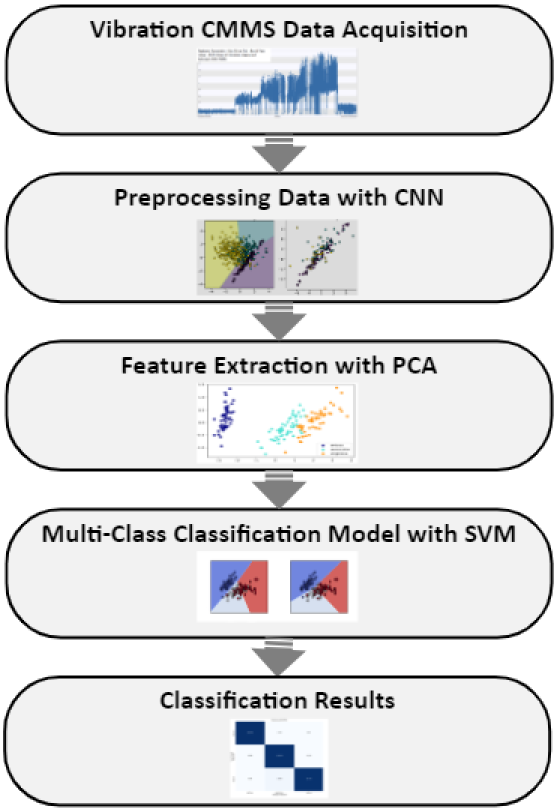


Fig. 4. Proposed Method for WTG Bearing fault diagnosis

Figure 5 of this section. The percentage of only 0.07% samples classified as False Positives demonstrates a high reliability for the applied model.
Additionally, it is important to note that the highest rate of incorrect classifications occurred between the two failure categories, indicating that despite the misclassification, there is a good differentiation between normal and healthy operation of the wind turbine versus degraded operation.

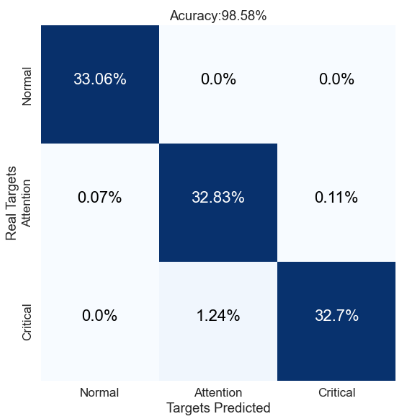


Fig. 5. WTG A Confusion Matrix

4.2 Principal Component Analysis Features

Despite the satisfactory result for the study, there remains the complexity of dealing with numerous inputs for training and classifying operational moments. Therefore, It was necessary to simplify and reduce the amount of input data, especially considering the ultimate goal of using this methodology for real-time identification of early failures in components across multiple wind turbines.
Consequently, the principal component analysis (PCA) was employed for training purposes, and the optimal number of principal components was identified to encapsulate 99% of the variance in the original training dataset. The visualization in Figure 6, portrays the scatter plot, illustrating the dispersion of data points among the chosen 6 components, while simultaneously highlighting the supervised classification employed as the target variable.
In addition to the components resulting from PCA, the Active Power Bins were also added as features to the model due to their correlation with vibration ranges and with the rotation of the wind turbine.

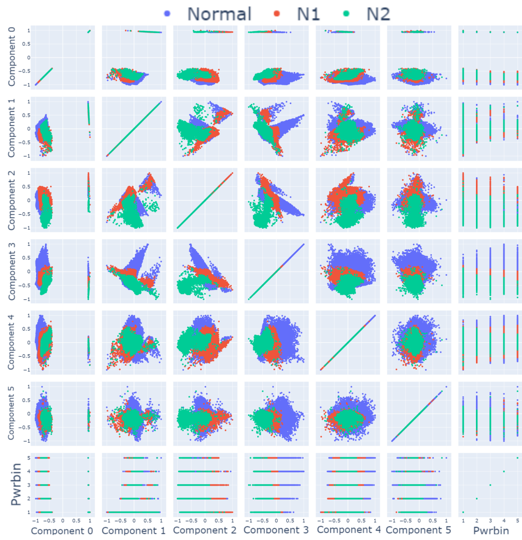


Fig. 6. Scatter Matrix Plot with PCA from WTG A

Even though the overall accuracy indicator being 10% lower than baseline case as confusion matrix on Figure 7 shows, the presented result is still satisfactory and within an acceptable range for the chosen application. This is primarily due to the reduction of model complexity, achieved by significantly reducing the number of defined inputs by 84%.

4.3 Hold Out Validation

Following the proposed methodology and applying the trained model with WTG A to the data from WTG B, which exhibits a similar failure pattern, we observed a high accuracy that follows the same pattern of correct classifications, false positives, and false negatives, as shown in Figure 8. This highlights the robustness of the trained model, especially considering that the data from WTG B was previously unseen. The low rate of false negatives was maintained, thus ensuring the model reliability.

The same procedure was applied to WTG B, considering a reduced model complexity. When applying the previously trained decomposition parameters to the new wind turbine, a remarkably similar dispersion pattern becomes evident, as exemplified in Figure 9. The graphs clearly show a

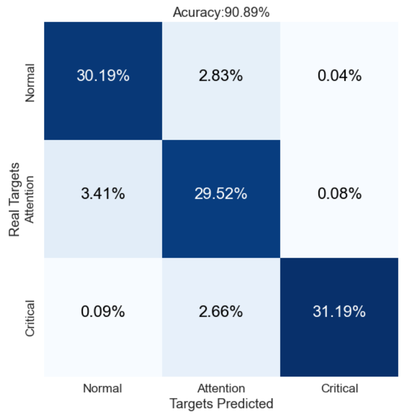


Fig. 7. WTG A Confusion Matrix from PCA Features

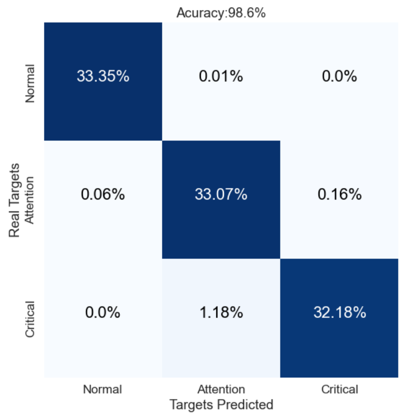


Fig. 8. WTG B Confusion Matrix trained with WTG A

higher degree of overlap between categories, indicating an expected degradation in performance due to this reason. Similar to the results obtained in the previous experiment, the hold out model also exhibited a slight decrease in overall classification performance, as expected due to the reduced model complexity. Despite the decrease in overall accuracy, the level of false positives remains within an acceptable range, still below 3% as Figure 10.

As previously mentioned and now concatenated in Figure 11a, the accuracy of all tests comparison, and in Figure 11b, the tested models percentage of false negatives. It reinforces the previous interpretation that the models incorporating the use of PCA have the lowest accuracies, but still above 90%. The percentage of false negatives, i.e., samples that the models failed to identify as faulty, can be interpreted as an indicator of the risk of non-detection by the models. The values around 3.5% are also considered acceptable.

5. CONCLUSION

This study demonstrated that with an accuracy of over 90%, it is possible to classify the degradation of the generator's bearing set in a wind turbine. The combined use of CNN and PCA resulted in a reduction in input complexity without compromising the identification of failure situations. It is important to highlight that, as the goal is to implement a real-time evaluation application, the impact of false positives is significant in the management of asset maintenance. Therefore, a very low number of false positives positively affects the process by avoiding unnecessary costs associated with unnecessary replacements.

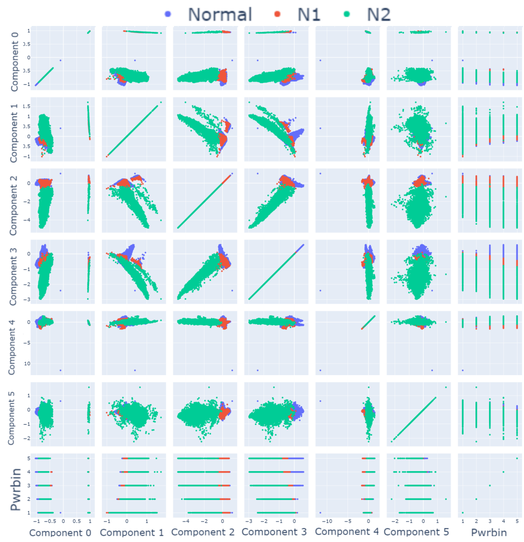


Fig. 9. Scatter Matrix Plot with PCA from WTG B

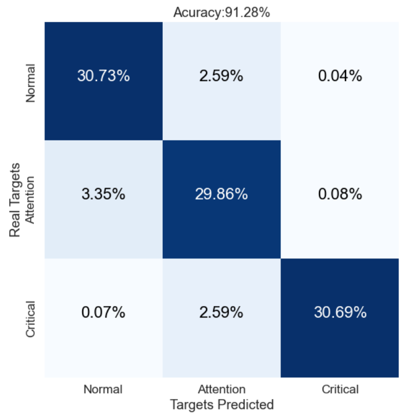
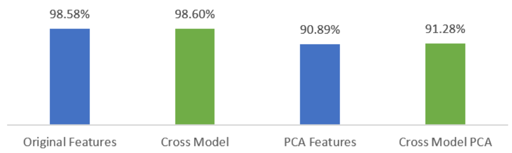
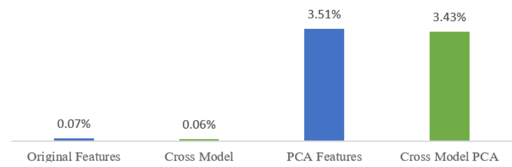


Fig. 10. WTG B Confusion Matrix from PCA Features trained with WTG A



(a) Models Accuracy



(b) Total Models False Negatives

Fig. 11. Measurements from Model Results

The results of the hold out test are also important to demonstrate that the failure behavior is quite similar regardless of the individual being evaluated. Therefore, in situations where CMS data availability is short for training, it is possible to use labeled data from another similar wind turbine while maintaining accuracy. The progression of this work would involve increasing the level of detail in the failure criticality assessment and even diagnosing the specific type of failure. Important to mentioned that the model is extremely dependent from the labeling done before training, there is a risk of a bad labeling on original data impact models performance. Furthermore, expanding to other types of failures and monitored equipment is also possible. From the aspect of classification algorithm, other methods would be tested as Decision Tree our clustering types as KNN for comparison.

REFERENCES

Berk, R.A. (2008). *Support Vector Machines*, 1–28. Springer New York, New York, NY.

Bisu, C., Olaru, A., Olaru, S., Alexei, A., Mihai, N., and Ushaq, H. (2024). Monitoring the wear trends in wind turbines by tracking fourier vibration spectra and density based support vector machines. *Mathematics*, 12(9). doi:10.3390/math12091307. URL https://www.mdpi.com/2227-7390/12/9/1307.

Borman, R.I., Napianto, R., Nugroho, N., Pasha, D., Rahmanto, Y., and Yudoutomo, Y.E.P. (2021). Implementation of pca and knn algorithms in the classification of indonesian medicinal plants. In *International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*.

Castellani, F., Natili, F., Astolfi, D., and Vidal, Y. (2024). Wind turbine gearbox condition monitoring through the sequential analysis of industrial scada and vibration data. *Energy Reports*, 12, 750–761. doi: https://doi.org/10.1016/j.egyr.2024.06.041.

Chang, C.C.W., Ding, T.J., Han, W., Chai, C.C., Bhuiyan, M.A.S., Choon-Yian, H., and Song, T.C. (2023). Recent advancements in condition monitoring systems for wind turbines: A review. *Energy Reports*, 9, 22–27.

Chou, C.H., Kuo, B.H., and Chang, F. (2006). The generalized condensed nearest neighbor rule as a data reduction method. In *18th International Conference on Pattern Recognition (ICPR'06)*, 556–559. Hong Kong, China. doi:10.1109/ICPR.2006.1119.

Clarke, B., Nicholas, G., Hart, E., Long, H., and Dwyer-Joyce, R. (2023). Loading on a wind turbine high-speed shaft gearbox bearing: Ultrasonic field measurements and predictions from a multi-body simulation. *Tribology International*, 181, 108319.

Dao, C., Kazemtabrizi, B., and Crabtree, C. (2019). Wind turbine reliability data review and impacts on levelised cost of energy. *Wind Energy*, 22, 1848–1871.

Enevoldsen, P. and Xydis, G. (2019). Examining the trends of 35 years growth of key wind turbine components. *Energy for Sustainable Development*, 50, 18–26.

Hand, D. and Batchelort, B. (1978). Experiments on the edited condensed nearest neighbor rule. *Information Sciences*, 14, 171–180.

He, R., Liu, W., Shan, M., and Wang, X. (2019). A new wind turbine health condition monitoring method based on vmd-mpe and feature-based transfer learning. *Measurement*, 148.

Hira, Z.M. and Gillies, D.F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in Bioinformatics*, 2015(1), 1–13.

Kavzoglu, T. and Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11, 352–359.

Kong, K., Dyer, K., Payne, C., Hamerton, I., and Weaver, P.M. (2023). Progress and trends in damage detection methods, maintenance, and data-driven monitoring of wind turbine blades – a review. *Renewable Energy Focus*, 44, 390–412.

Lydia, M. and Edwin Prem Kumar, G. (2023). Condition monitoring in wind turbines: a review. In F.P.G. Marquez, M. Papaelias, and V.L.J. Junior (eds.), *Wind Energy Engineering, Non-Destructive Testing and Condition Monitoring Techniques in Wind Energy*, 229–247. Academic Press.

Ouyang, T., He, Y., and Huang, H. (2019). Monitoring wind turbines' unhealthy status: A data-driven approach. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(2), 163–172.

Pattabiraman, T., Srinivasan, K., and Keppanan, M. (2015a). Assessment of sideband energy ratio technique in detection of wind turbine gear defects. *Case Studies in Mechanical Systems and Signal Processing*, 2.

Pattabiraman, T., Srinivasan, K., and Malarmohan, K. (2015b). Assessment of sideband energy ratio technique in detection of wind turbine gear defects. *Case Studies in Mechanical Systems and Signal Processing*, 2, 1–11.

Peng, J., Kimmig, A., Niu, Z., Wang, J., Liu, X., Wang, D., and Ovtcharova, J. (2022). Wind turbine failure prediction and health assessment based on adaptive maximum mean discrepancy. *International Journal of Electrical Power and Energy Systems*, 134.

Salih Hasan, B.M. and Abdulazeez, A.M. (2021). A review of principal component analysis algorithm for dimensionality reduction. *Journal of Soft Computing and Data Mining*, 2(1), 20–30.

Sun, S., Wang, T., and Chu, F. (2022). In-situ condition monitoring of wind turbine blades: A critical and systematic review of techniques, challenges, and futures. *Renewable and Sustainable Energy Reviews*, 160, 112326.

Wen, X. (2020). Modeling and performance evaluation of wind turbine based on ant colony optimization extreme learning machine. *Applied Soft Computing Journal*, 94.

Wen, X. and Xu, Z. (2021). Wind turbine fault diagnosis based on relieff-pca and dnn. *Expert Systems with Applications*, 178.

Yan, H., Mu, H., Yi, X., Yang, Y., and Chen, G. (2019). Fault diagnosis of wind turbine based on pca and gsa-svm. In *2019 Prognostics and System Health Management Conference (PHM-Paris)*, 13–17. Paris, France.

Zhang, J., Zeng, B., Shen, W., and Gao, L. (2022). A one class shapelet dictionary learning method for wind turbine bearing anomaly detection. *Measurement*, 197.