# Failure classification of wind turbine operational conditions using hybrid machine learning

Marcela Rodrigues Machado[1], Amanda Aryda Silva Rodrigues de Sousa[1], Jefferson da Silva Coelho[1,2], and Rafael de Oliveira Teloli[3]

[1] Department of Mechanical Engineering, University of Brasília, Campus Universitário Darcy Ribeiro, Brasília, 70910-900, Brazil

[2] Department of Mechanical Engineering, Federal University of Amazonas, 69080-900, Manaus, Brazil

[3] Université Marie et Louis Pasteur, SUPMICROTECH, CNRS, Institut FEMTO-ST, F-25000 Besançon, France

**Correspondence:** Marcela R. Machado (marcelam@unb.br)

**Abstract.** Wind turbines are complex electromechanical systems that require continuous monitoring to ensure operational efficiency, reduce maintenance costs, and prevent critical failures. Machine learning has shown great promise in structural health monitoring (SHM) by enabling automated fault detection through data-driven approaches. However, challenges remain in adapting SHM methods to complex environmental conditions while maintaining reliable fault detection and classification. This work proposes a hybrid model that combines supervised and unsupervised learning techniques for classifying operational failures in wind turbines. The proposed framework integrates multiphysics data, combining structural and environmental information, to monitor four distinct operational states. The approach begins with the analysis of sensor signals and the extraction of descriptive features that capture the dynamic behaviour of the turbine, taking into account the influence of temperature and wind speed. The unsupervised k-means is applied to label and cluster the dataset, while feature and sensor selection are performed using canonical correlation analysis to rank the most informative variables. A novel relative change damage index is introduced to normalise and scale features based on their relative variability, enhancing the accuracy of clustering and fault classification. Classification is conducted using six different machine learning algorithms, with an indication of the better model. Experimental results, also considering environmental conditions and sensor failures, demonstrate strong performance across both binary and multiclass tasks, including the detection of pitch-drive faults and the accurate identification of rotor icing and aerodynamic imbalance. The model achieved from 85 to 98% classification accuracy, highlighting its effectiveness in diagnosing wind turbine conditions and improving the overall reliability and operational analysis of these systems.

## 1 Introduction

Wind turbines are complex electromechanical systems that operate in hazardous environments and contain critical components, including blades, the main bearing, main shaft, gearbox, nacelle, tower, foundation, yaw system, and connecting bolts. Continuous monitoring of these components is essential to ensure their integrity, reduce costs, and enhance the operational efficiency of those turbines (Veers et al., 2023). Like any other electromechanical system, wind turbines are subject to several unforeseen and serious failures that can result in fatal disasters (Asian et al., 2017; Machado and Dutkiewicz, 2024). Mechanical,

electrical, and environmental failures are the primary causes of issues in wind turbines (Thomas, 2024). Mechanical failures typically affect components such as the rotor blades, gearbox, bearings, and main shaft, which can compromise the turbine's performance. In the electrical system, which includes generators, converters, and control systems, failures can result from insulation degradation, thermal voltages, and electrical transients, impacting the efficiency and safety of operation (Liu et al., 2024). Environmental factors, such as extreme temperatures, atmospheric discharges, humidity, wind direction and speed, and corrosion, pose significant challenges to turbine reliability and service life (Figueiredo et al., 2010; Morozovska et al., 2024). An effective wind turbine monitoring strategy combines Condition Monitoring (CM) of mechanical subsystems and SHM of structural elements (Civera and Surace, 2022). The authors highlight the growing integration of both into AI-based systems, aiming to improve reliability and operational safety. Research on fault detection of turbine components using machine learning (ML) techniques has advanced significantly over the past decade. It has significantly enhanced the efficiency and accuracy of SHM techniques by automating data analysis and improving damage detection (Smarsly et al., 2016; Flah et al., 2020; Farrar and Worden, 2012). These algorithms process large volumes of data, identifying patterns and anomalies that may indicate structural deterioration. Thus, structural health monitoring with machine learning (SHM-ML) has emerged as a powerful data-driven approach, enhancing multi-sensor systems and intelligent algorithms to detect, localise, and quantify structural damage under complex environmental and operational conditions.

Deep learning (DL) and machine learning models have emerged as leading approaches for wind turbine fault diagnosis, outperforming traditional methods in handling complex, nonlinear, and noisy operational data. Their applications span vision-based approaches (Li et al., 2024; Dwivedi et al., 2024; Li et al., 2025; Zhao et al., 2025), acoustic-based techniques (Regan et al., 2017; Kuo et al., 2022; Lei et al., 2025; Carmona-Troyo et al., 2025), and vibration-based monitoring, among others. Vision-based methods, however, remain limited to identifying superficial or external damage and are strongly affected by lighting and weather conditions, while being unable to detect subsurface faults. Acoustic techniques are highly sensitive to environmental noise and often require expensive sensor maintenance, adding complexity to the monitoring process. Vibration-based monitoring is among the most established SHM strategies, as variations in natural frequencies, mode shapes, and damping ratios are closely associated with structural degradation. Nevertheless, its effectiveness can be reduced by operational variability, which makes it more difficult to detect early-stage or localised damage. Despite these limitations, vibration monitoring remains a widely applied technique for various wind turbine components.

Hybrid approaches involve strategies that combine multiple machine learning models, integrate multimodal or heterogeneous data, or fuse complementary algorithms with multiphysics information, thereby enhancing model generalisation, interpretability, and reliability in complex monitoring scenarios. In this context, Baquerizo et al. (2022) proposed a Siamese Neural Network using image-like matrices from 24-channel triaxial accelerometer data for jacket structures, achieving 100% accuracy in crack and bolt-looseness detection. Building on this, Antoniadou et al. (2015b) combined vibration analysis, SCADA, and ML algorithms (SVM, neural networks, Gaussian Processes) for blade and gearbox monitoring, extracting features from time, frequency, and time-frequency domains using diverse sensors, including accelerometers, fiber optics, and thermography. Hybrid intelligent models for floating offshore turbines using a CNN-LSTM-ATT with Chebyshev decomposition to forecast mooring-line tension under harsh seas were proposed in (Tang et al., 2025), a Gaussian Process system for predicting crack location,

severity, and remaining useful life composite blades in (Krishna et al., 2025), and a Transformer-based PatchTST virtual sensor, trained on Alpha Ventus data, accurately reconstructing strain in jacket foundations (Ángel Encalada-Dávila et al., 2025). A digital twin for mooring-line monitoring (Mousavi et al., 2024) used a hybrid CNN-LSTM with CEEMD/FDD-derived natural frequencies, achieving 97.65% accuracy in damage localisation and severity classification, outperforming DCNN, BPNN, and SVM under noise and environmental variability.

Wind turbine blades monitoring has increasingly leveraged hybrid machine learning models. Early approaches combined neural networks and Gaussian processes using frequency response functions and SCADA data (Antoniadou et al., 2013, 2015a), while Dervilis et al. (2014b, a) employed auto-associative neural networks with RBF models to achieve high sensitivity with few false positives. Rule-based and ensemble methods were explored in (Joshuva. and Sugumaran., 2017; Joshuva et al., 2019), which integrated MLPs and logistic model trees with ARMA features, achieving 94.75% accuracy in crack localisation. More recent studies focus on hybrid architectures that fuse different algorithms or modalities. Tsiapoki et al. (2018) combined environmental and operational effects with unsupervised clustering, Movsessian et al. (2022) integrated Mahalanobis Distance, XGBoost, and SHAP for explainable SHM, Song et al. (2024) and Sharma and Nava (2024) employed CNN-GRU and AR-CNN hybrids for offshore blade and mooring system monitoring with over 93% accuracy, and Ashkarkalaei et al. (2025b, b) combined One-Class SVM, autoencoders, and DSFs with ANOVA-selected features with Random Forest, achieving F1 >95%. Bayesian-optimised kNN/SVM frameworks (Korolis et al., 2025) and Relevance Vector Machines (Kuai et al., 2024) further demonstrate the advantages of hybrid and sparse models, offering high accuracy and computational efficiency under complex environmental variability.

The vulnerability of towers and foundations has driven the development of hybrid ML-based SHM strategies. Nguyen et al. (2018) combined artificial neural networks (ANNs) with modal parameters, such as mode shapes and frequencies, to localise and quantify tower damage in simulations. Hoxha et al. (2020) extended this approach by integrating vibration responses with multiple classifiers, including kNN, quadratic SVM, and Gaussian SVM, for tower fault detection. For jacket-type foundations, Vidal et al. (2020) and Leon-Medina et al. (2021) employed PCA for dimensionality reduction combined with classifiers such as kNN, SVM, and XGBoost to detect and localise cracks. Ren and Yong (2022) advanced unsupervised hybrid methods using dynamically weighted k-means clustering to categorise tower fault types. In contrast, Milani et al. (2025) demonstrated that Random Forest classifiers could effectively monitor blade pitch misalignment, capturing subtle deviations. For drivetrain monitoring, Elforjani (2020) combined ANNs, Gaussian processes, and SVM with feature extraction and PCA to enhance gearbox fault detection. Zarrin et al. (2021) developed a neuromorphic model for classifying accelerometer data from healthy and damaged gearboxes, while Gao et al. (2021) applied CNNs for bearing and gearbox fault detection. Praveen et al. (2022) improved vibration monitoring by segmenting gearbox signals and validating them with decision trees, SVM, and deep neural networks. For bearing faults, Vives (2022); Vives et al. (2022) demonstrated the effectiveness of combining kNN, SVM, and deep learning, whereas Meyer (2022) and Amin et al. (2023) employed unsupervised and cyclostationary-feature CNNs for accurate health state classification. These studies highlight the advantages of hybrid approaches that integrate multiple algorithms and preprocessing techniques to enhance robustness in fault detection and classification.

Data quality, imbalance, generalisation, feature selection, and explainability remain challenges in wind turbine monitoring, with ongoing research focused on improving applicability, addressing data limitations, and enhancing interpretability. Despite recent advances, SHM techniques still face difficulties in generalising across operational and environmental conditions while ensuring reliable detection. While many studies employ supervised, unsupervised, or hybrid learning for anomaly detection and classification, these methods are still limited to specific operations and the detection of local component faults. Hybrid machine learning models, particularly those that account for environmental variability, demonstrate strong potential for wind turbine diagnostics by, enabling accurate, early fault detection and adaptability. However, a research gap remains in designing hybrid frameworks that balance computational efficiency and interpretability without compromising diagnostic performance.

This study proposes a hybrid monitoring framework that integrates multiple machine learning models with multimodal data to enhance the reliability, accuracy, and interpretability of fault detection in the Aventa wind turbine. Unlike conventional single-model or single-source approaches, the proposed framework effectively captures cross-domain correlations and accounts for environmental variability, enabling dependable monitoring under complex operational conditions. The methodology employs unsupervised k-means clustering to partition the data into homogeneous clusters, thereby facilitating pattern recognition without predefined labels. These clusters are subsequently analysed using multiple supervised machine learning algorithms for binary and multiclass fault classification, based on structural-temporal response signals and SCADA data (wind speed and temperature) from the Aventa onshore wind turbine. The framework simultaneously performs fault classification, feature ranking, and sensor selection, demonstrating flexibility in handling diverse input features and reinforcing SHM decision-making. The main contributions of this study are: (i) the development of a hybrid ML framework for operational fault assessment combining multiple supervised and unsupervised learning technique and multimodal data, (ii) the formulation of a feature relative change damage index strategy for feature normalisation and scaling, and (iii) the implementation of a canonical correlation-based feature and sensor selection process. By integrating multiphysics information, the proposed model enhances diagnostic capability, computational efficiency, and prediction accuracy. Experimental validation on the Aventa 6.7 kW onshore wind turbine confirms high performance in both binary and multiclass fault classification. Six ML algorithms were compared, and the best-performing model achieved classification accuracies between 85% and 98%. The framework reliably detected pitch-drive faults and accurately identified rotor icing and aerodynamic imbalance, even under varying environmental conditions and sensor failure scenarios, demonstrating its effectiveness and potential for real-world SHM applications.

## 2 Model description

The proposed monitoring model consists of seven steps, as illustrated in Fig. 1, including: (1) receiving the acquired data; (2) data processing and organisation; (3) feature extraction, normalisation, and grouping for similarity pattern; (4) unsupervised feature labelling and clustering; (5) feature and sensor selection; (6) data splitting, and ML failure identification and classification; and (7) Fault classification and model evaluation. The final step also outputs the operational failure and identifies the best-performing ML algorithm based on its performance metric. The novelties of the presented study are associated with feature and sensor selection, data normalisation, and multimodal and multiple fault classification, which are detailed in sec-

tions 2.2, 2.3, and 3.2. The pseudocode developed for the failure classification, presented in Algorithm 1, has the following structure:

**Algorithm 1** Hybrid ML model for Aventa wind turbine failure classification

1: **Data Acquisition:** Receive time-domain structural response from accelerometer sensors and environmental conditions from the SCADA.

2: **Data Processing:** Involves correcting raw data by checking for missing values and applying signal processing techniques.

3: **Feature extraction, normalisation, and grouping:** Fifteen formulations are employed to extract features from the time-domain signal, which is then grouped by similarity, see section 2.3.

4: **Pattern Recognition, labelling, and Clustering:** Use the k-means unsupervised algorithm to cluster the data: $J = \sum_{i=1}^{n} min_k \left( \|x_i - \overline{x}_k\|^2 \right)$

5: **Feature and sensor selection** performed via Canonical-correlation-based selection, and final dataset organisation.

6: **Data Splitting, and ML classification :** Split the dataset into training and testing as listed in Table 4. Further, applying ML algorithms (SVM, K-NN, RF, NB, DT and XGBoost) for failure identification and classification.

7: **Model Evaluation:** Calculate the Performance for the ML classifiers (Cross-validation, Accuracy, Precision, Recall, and F1-score), identification of the best ML model(s), and generation of hte confusion matrix.

**Final decision:** Information about failure state based on classification algorithm outcomes.

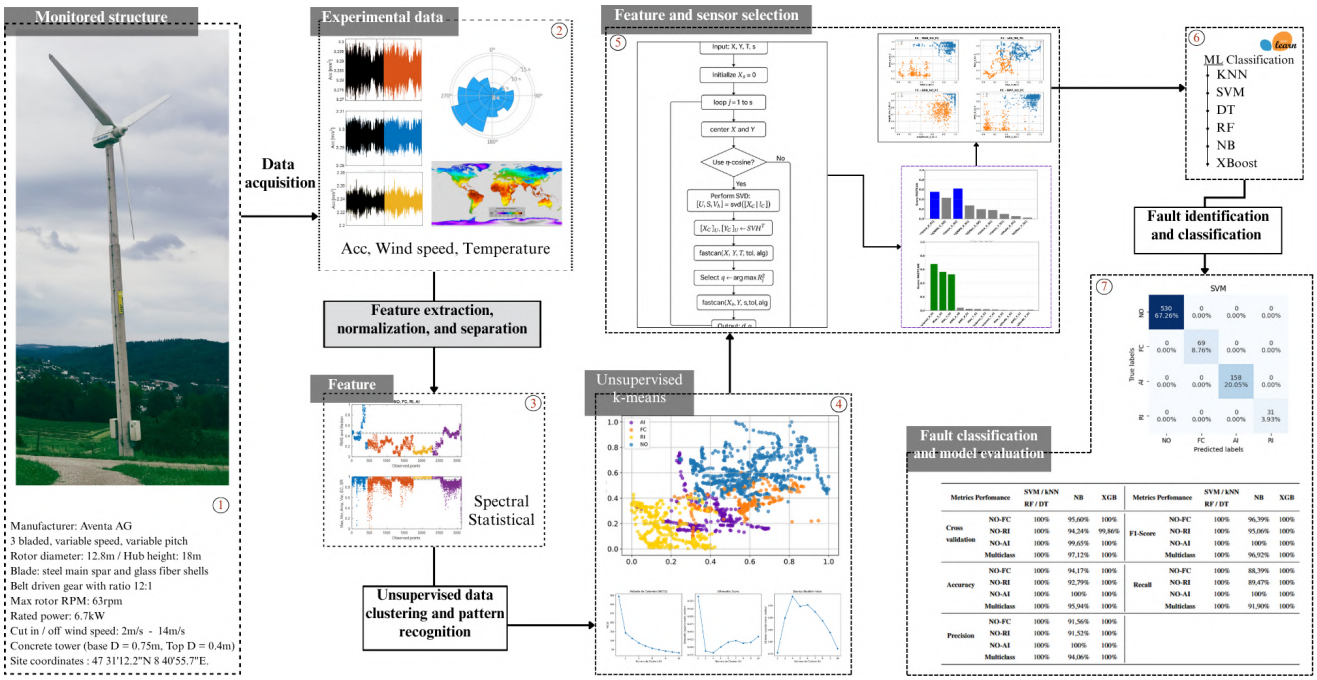



**Figure 1.** Pipeline of the hybrid machine learning model for fault classification on the Aventa 6.7 kW wind turbine.

### 2.1 Data processing, signal and sensor analysis

The wind turbine monitored in this study is the Aventa AV-7 model, manufactured by Aventa AG (Switzerland) and commissioned by ETH. This turbine has a rated power of 6.7 kW and operates using a belt-driven generator coupled with a frequency converter and a variable-speed drive. It begins generating power at a wind speed of 2 m/s and has a cut-off speed of 14 m/s. The rotor has a diameter of 12.8 m with three blades and is mounted at a hub height of 18 m. The maximum rotational speed reaches 63 RPM. Turbine control is managed through a variable-speed and -pitch mechanism. The turbine is installed in Taggenberg, Switzerland (coordinates: 47°31'12.2"N, 8°40'55.7"E). Instrumentation and data acquisition were provided in (Eleni Chatzi et al., 2023) and include 14 accelerometers strategically placed along the tower, on the nacelle main frame, the main bearing, and the generator, as shown in Fig. 2. Additionally, two full-bridge strain gauges are mounted at the tower base to measure fore-aft and side-to-side strain, which can be converted into bending moments. Acceleration and strain signals are sampled at 200 Hz. Environmental measurements, including temperature and humidity, are collected at the tower base with a sampling rate of 1 Hz. Operational performance data (SCADA), including wind speed, nacelle yaw orientation, rotor RPM, power output, and turbine status, are recorded at a rate of 10 Hz. To ensure data quality and enhance the reliability of fault detection, a comprehensive supervised data preprocessing was implemented. This included a supervised sensor analysis, outlier removal, signal segmentation, detrending, feature extraction, normalisation, and similar features identification and grouping.

**Table 1.** Information on SCADA data and accelerometers used in the monitoring, with the respective sensor channel and location.

| Location | Channel name | | | |
|---|---|---|---|---|
| Tower | L5_X_01 | L5_Y_01 | L5_X_02 | L5_Y_02 |
| Nacelle | NMF_Y_01 | NMF_X_02 | NMF_Y_02 | |
| Main Shaft Bearing | MSH_X_01 | MSH_Z_01 | - | - |
| Generator | GEN_X_01 | GEN_Y_01 | GEN_Z_01 | |
| SCADA | Temperatute | Wind speed | - | |

The data preprocessing stage begins by analysing the sensors' physics and evaluating each sensor's contribution to failure identification. The pre-established failures are the rotor icing event (RI), the flexible coupling of the linear drive of the collective pitch system (FC), and aerodynamic imbalance on one blade (AI). The Aventa dataset provides three-axis acceleration signals. Table 1 lists the SCADA data and accelerometers used for monitoring, along with their respective sensor channels and locations. The $x$-axis captures side-to-side turbine motion, and the $y$-axis captures fore-aft turbine motion. Figure 2(a) illustrates the sensor locations within the turbine. Aside from the SCADA data (temperature and wind velocity), accelerometers marked in orange, blue, and yellow are used for monitoring. Specifically, three accelerometers are located in the nacelle, GEN_ACC (orange), NMF_ACC (blue), and MSB_ACC (yellow), and one is located at the tower top, F5_ACC. The sensors record signals under the different operational conditions, including normal operation (NO). Figure 2 also presents the raw temporal data available in the dataset. In the acceleration graphs, black lines indicate normal turbine operation, whereas coloured lines

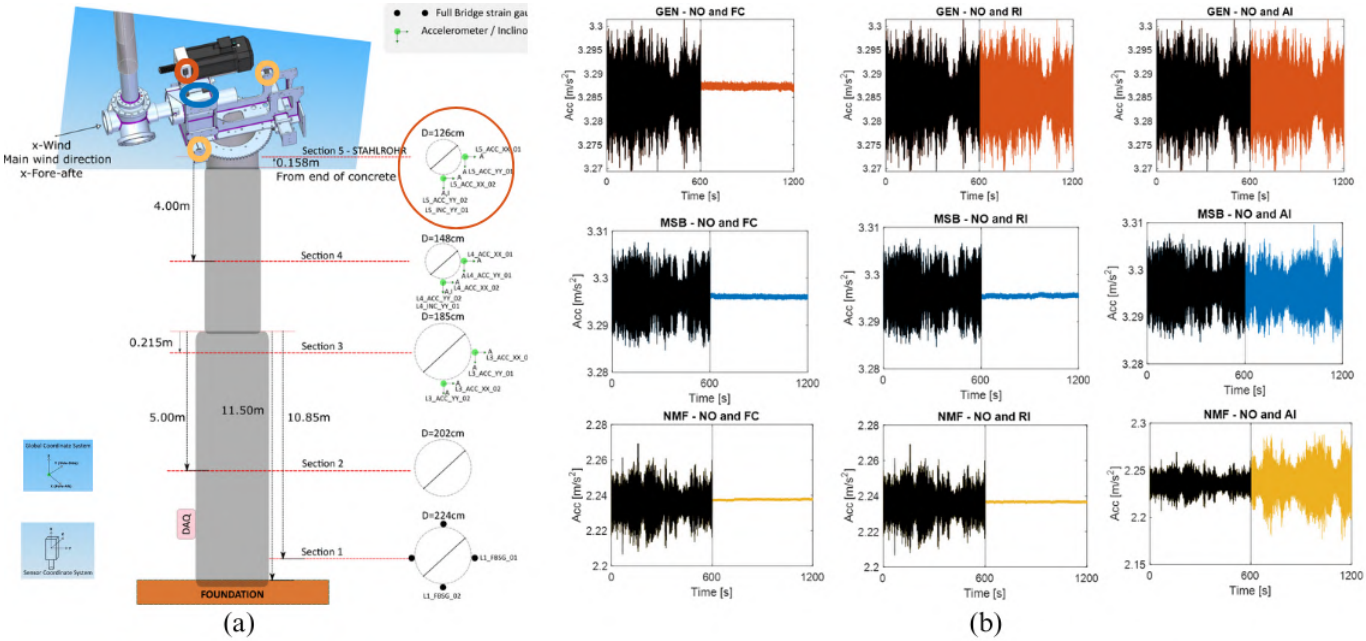



**Figure 2.** Schematic representation of the turbine sensor channel and location Chatzi et al. (2023). SCADA data and accelerometers marked in orange, blue, and yellow colours are adopted in the monitoring (a). Temporal signal data of the turbine's NO and operation with failures (b): (Top) sensor GEN NO-FC, NO-RI, and NO-AI placed from left to right, respectively. (Middle) sensor MSB NO-FC, NO-RI, and NO-AI; and (Bottom) sensor NMF NO-FC, NO-RI, and NO-AI.

represent various failure conditions. To ensure consistent comparisons, all signals were selected from the same day, minimising the influence of changing environmental conditions.

The top row of graphs shows the x-axis response from the generator-mounted sensor (GEN), comparing NO with failures FC, RI, and AI, from left to right. The middle row displays the x-axis response from the main shaft bearing sensor (MSB), and the bottom row shows the same for the nacelle main frame sensor (NMF), following the same failure comparison pattern. Both x- and y-axis signals were analysed, revealing consistent trends: FC failure caused a pronounced amplitude reduction, whereas RI and AI failures resulted in only minor amplitude changes. These variations may fluctuate daily due to environmental conditions. To ensure effective fault detection, data from accelerometer sensors placed at the tower top, nacelle, main shaft bearing, and generator were used in the model, considering their sensitivity to structural, mechanical, and aerodynamic failures.

## 2.2 Dataset

During data analysis, duplicate files were identified, prompting a data redistribution strategy to eliminate redundancies and streamline the process. The data were categorised and colour-coded by operational conditions and events, making patterns, such as seasonality, frequency, and duration, more apparent and visually accessible. These cleaning and reorganisation steps

ensure that the same data is not reused across different turbine operations, thereby reducing the risk of misclassification and misinterpretation during monitoring.

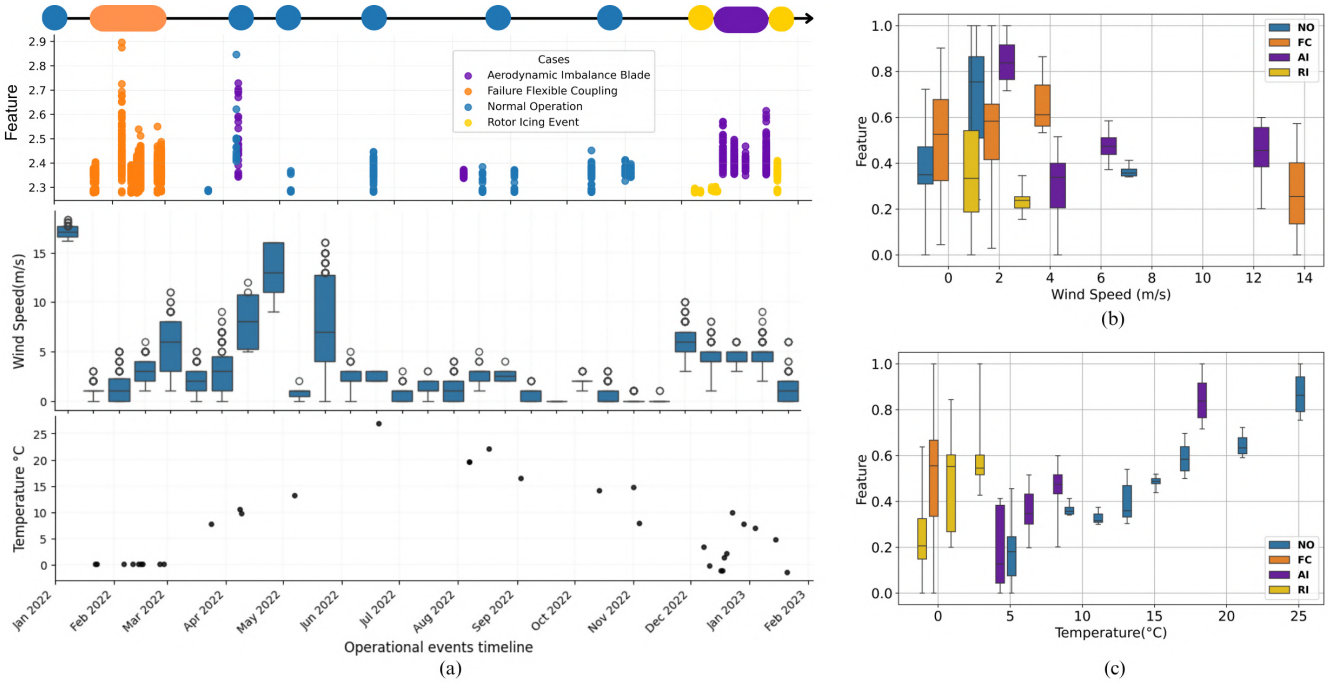



**Figure 3.** The feature dataset volume over the monitoring timeline categorised by yellow indicates RI, blue denotes NO, purple corresponds to AI, and orange represents the FC condition (a-top). Wind speed boxplot (a-middle), and mean temperature (a-bottom). Boxplot of L05 sensor feature by wind speed variation (b), and feature by temperature variation (c).

Feature extraction was performed after this reorganisation. Aside from the accelerometer sensors, the supervisory control and data acquisition (SCADA) data includes the temperature, wind speed, humidity, and power output. Temperature, wind speed, and acceleration responses are used for monitoring. Figure 3(a) shows the maximum values of the temporal spectrum for each operational condition mapped over the monitoring timeline. Each condition is colour-coded for clarity: yellow for RI, blue for NO, purple for AI, and orange for flexible FC. Events are plotted along the horizontal axis to highlight their temporal distribution. The timeline begins with a NO sample, followed by a pitch drive coupling failure recorded on 2022-02-16, which required replacement. From April to November 2022, the turbine operated normally, during which AI was intentionally simulated using roughness tape on the blades. Additionally, naturally occurring RI events were observed between December 2022 and February 2023. Wind speed fluctuates throughout the monitoring period, as indicated by the boxplot, which displays the minimum, maximum, median, and first and third quartiles of the wind speed data. Meanwhile, ambient temperature ranges from -2°C to 26°C throughout the year (see Fig 3c).

Both environmental conditions, wind speed and temperature, significantly influence the dynamic response of the wind turbine over time. Figures 3(b–c) present boxplots showing the feature values corrected by wind speed and temperature, respectively.

This analysis focuses on the RMS value of sensor NMF, which is selected for its highest-scoring features and well-defined failure patterns, representing the general behaviour of the features. When correlating feature values with wind speed, the FC and AI conditions exhibit the greatest sensitivity to increasing wind speeds, though with relatively low dispersion. Conversely, RI, NO, and FC show higher maximum, minimum, and quartile values at lower wind speeds. Regarding temperature, rising temperatures primarily affect NO and AI conditions, whereas lower temperatures have a greater impact on FC and RI. These findings underscore the importance of incorporating environmental variables into the monitoring framework to improve the robustness and accuracy of fault classification. The SCADA data in the general dataset are considered the mean temperature and mean wind speed, both acquired over the 24 hours of the respective day.

### 2.3 Feature extraction, normalisation, and dataset organisation

Feature selection involves choosing a subset of variables from the original dataset. This step transforms the spectrum data into new variables to create a refined dataset (Christopher M. Bishop, 2006). Fifteen features are formulated using the time-domain signal (Knittel et al., 2019; Barreto et al., 2021; Ashkarkalaei et al., 2025a), as presented in Table 2, where $x$ is the temporal spectrum signal, and $p$ is the sampling points value.

**Table 2.** temporal and statistical formulation used for feature extraction.

| Statistical | | Temporal | |
|---|---|---|---|
| **Feature** | **Formulation** | **Feature** | **Formulation** |
| Median value | Median $= (x)$ | Energy of the signal | Energy $= \sum_{k=1}^{P} x_k^2$ |
| Mean value | Mean $= \mu = \frac{1}{p}\sum_{i=1}^{p}(x_i)$ | Energy of the centred signal | Ec=$\sum_{k=1}^{P}(x_k - \mu)^2$ |
| Variance | Var $= \frac{1}{p}\sum_{i=1}^{p}(x_i - \mu)^2$ | Shannon Entropy | $E_S$(x)=-$\sum_{k=1}^{P} x_k^2 * \log_2(x_k^2)$ |
| Skewness | Skewness $= \frac{1}{p}\sum_{i=1}^{p}\frac{(x_i-\mu)^3}{\sigma^3}$ | Signal rate | SR $= \frac{\max(x_{k=1:P})-\min(x_{k=1:P})}{\mu}$ |
| Standard deviation | STD $= \sigma = \sqrt{\frac{1}{p}\sum_{i=1}^{p}(x_i - \mu)^2}$ | Maximum value | Max $= \max(x)$ |
| Moment order | Mo$_i = \frac{E(x-\mu)^i}{\mathrm{Var}^{i/2}}$ $(i = 5:10)$ | Minimum value | Min $= \min(x)$ |
| Kurtosis | Kurtosis $= \frac{1}{p}\sum_{i=1}^{p}\frac{(x_i-\mu)^4}{\sigma^4} - 3$ | Amplitude range (Peak-to-peak) | Amp : Max − Min |
| | | Root mean square | RMS $= \sqrt{\frac{1}{p}\sum_{i=1}^{p} x_i^2}$ |

Temporal features are extracted from a signal's frequency or time content, such as maximum, minimum, amplitude range, energy, energy centre, skewness, signal rate, and RMS. It captures how the signal's energy is distributed over time. Statistical features, in turn, are computed from the signal's time-domain values, including the median, mean, variance, skewness, kurtosis, higher-order moments, and standard deviation, which characterise the overall distribution and variability of the signal. Together, these features are compiled, analysed, and evaluated for their similarity and sensitivity to the dynamic behaviour of the system.

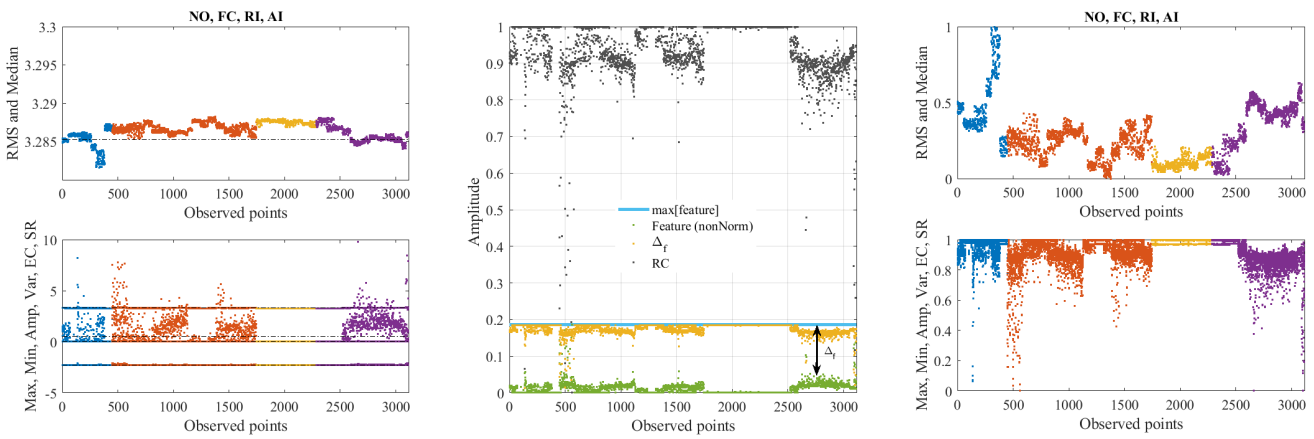
Despite feature refinement and grouping, the remaining features still pose a challenge due to their minimal amplitude variation, often differing by only a few digits. Such small variations can hinder accurate machine learning classification. To address this issue, the proposed relative change damage index incorporates a normalisation step based on the relative change of each

feature, defined as

$$\Delta_f = \max(\mathbf{X}_i) - \mathbf{X}_i \tag{1a}$$

$$RC = \frac{\Delta_f}{\max(\Delta_f)} \tag{1b}$$

where $\mathbf{X}$ is a table containing the selected features, and $\Delta_f$ measures each feature's deviation from its maximum value within the selected feature set to each feature vector point ($\mathbf{X}_i$), the normalization factor $RC$ rescales these deviations to the [0, 1] interval. This linear transformation preserves the original ordering and proportionality of the data, ensuring that the intrinsic dynamic pattern is maintained. At the same time, it amplifies subtle differences that would otherwise be masked by small numerical variations, thereby enhancing the sensitivity of the features for fault detection and classification. Hence, the $RC$ normalisation converts the data into dimensionless features that, before normalisation, differed only slightly (e.g., 3.280 vs. 3.286), rendering these differences numerically negligible for machine learning models. Thus, the proposed approach amplifies these small variations while preserving the intrinsic dynamic pattern. Features with values close to unity correspond to undamaged conditions, whereas values approaching zero indicate faults. Hence, the method enhances fault sensitivity and enables consistent comparison across features of different magnitudes, making the damage feature more sensitive to anomalies under varying operating states.




(a) (b) (c)

**Figure 4.** (a) Non-normalised feature grouped by similarity under different operating conditions, NO-blue, FC-orange, RI-yellow, AI-purple colour. (b) Demonstration of the scaling and normalisation on the AMP-feature related to Eq. (1). Scaled and normalised features grouped by similarity under different conditions (c).

The relative change normalisation provides a theoretical foundation based on relative scaling and dimensionless feature representation. Figure 4(a-c) illustrates the process: (a) the non-normalised features, (b) the scaling and normalisation applied to the Amp feature (extended to all others), and (c) the normalised features obtained using the proposed relative change technique. Features such as energy, kurtosis, higher-order moments, and Shannon entropy exhibited low sensitivity to failures

and were excluded during normalisation and analysis. This first supervised analysis revealed two feature groups with similar behaviour. Group 1 consists of RMS and median (Fig. 4 c-Top), while Group 2 comprises maximum, minimum, amplitude range, variance, energy centre, and signal rate (Fig. 4 c-Bottom). The formation of two feature groups indicates a high degree of correlation within each group, with RMS and median primarily capturing the signal's central tendency. In contrast, the features of group 2 describe its variability and extremes. Therefore, using all features would introduce redundancy, increase computational cost, and risk overfitting due to the curse of dimensionality. Selecting only representative features from each group ensures that the classifier retains the essential discriminatory information while improving efficiency and generalisation.

From the two correlated feature groups, four representative features (RMS, maximum, variance, and amplitude range) were selected to compose the global dataset. Those four features were chosen over the other in their respective group due to near-equivalence, thus avoiding redundancy. The initial global dataset, therefore, consisted of the four selected features extracted from each sensor axis, resulting in a structure of measurement_rows $\times$ 48 columns, plus the SCADA data, as detailed in Table 1. However, including all these features may appear arbitrary and degrade model performance due to their high similarity. To refine the global dataset, a canonical correlation analysis was applied to identify the most sensitive features for each sensor, accounting for all axes simultaneously within the sensor group. The highest feature correlation scores were selected, reducing the dataset to 4 representative features plus SCADA. This process yielded the final datasets: the binary classification dataset has six columns (four features plus SCADA), and the multiclass dataset has nine columns (selected features plus SCADA). Identifying the most informative sensor is also relevant, as it indicates which sensor contributes most to classification accuracy and provides resilience in the event of sensor loss. To validate this study, the most sensitive feature from the best-performing sensor was removed in subsequent tests, thereby demonstrating the robustness and consistency of the proposed feature selection strategy. Feature and sensor selection are detailed in Section 3.2.

#### 2.3.1 Structured final datasets

The final refined dataset organisation is a crucial step for enabling machine learning algorithms to classify operational failures accurately. The deployed dataset exhibits a multimodal nature, integrating structural, temperature, and wind velocity information. The structural data are directly acquired from the accelerometers, which capture the system's physical response and potential failure signatures. In contrast, temperature and wind speed data are obtained from the SCADA system and represent indirect but relevant variables that influence the dynamic behaviour measured by the accelerometers. However, they do not directly describe the dynamics of the damage. The final dataset is organised into four configurations, which are evaluated and their relevance demonstrated in Section 4, yielding

- *Fe-kms-Sc*: dataset composed of the selected features (Fe) extracted from the accelerometer time signals, followed by the SCADA data (Sc), and beelining provided by the k-means clustering (kms).
- *Fe-Sc-kms*: dataset including the selected features, the SCADA data, and the corresponding k-means labels.
- *Fe-kms*: dataset including only the selected features and the k-means labels, without incorporating the SCADA information.

- *Fe-kms-LoseSensor*: dataset including the selected features excluding those obtained from the most sensitive sensor identified by the Fast CCA method, together with the k-means labels and SCADA data. This configuration simulates the loss of the most sensitive sensor in each analysis scenario.

The k-means labelling of the datasets *Fe-kms-Sc* and *Fe-kms-LoseSensor* is applied exclusively to the features that contain the most relevant physical information about the damage. The SCADA data are then organised by the corresponding day and hour for each feature sample. This ensures that the SCADA records follow the same row reordering imposed by the k-means clustering, while preserving the correct temporal(day and hour) correspondence between the SCADA measurements and the related accelerometer data.

## 3 Hybrid machine learning model

The monitoring model integrates sequential stages, starting with data processing and feature extraction, followed by pattern recognition through clustering, and extending to feature and sensor selection to strengthen dataset reliability. By coupling unsupervised clustering with canonical-correlation-based feature selection, the framework ensures preservation of the most informative variables, thereby enhancing the accuracy of the classification task.

### 3.1 Pattern recognition and clustering

Unsupervised pattern recognition, labelling, and clustering of the normalised features were performed using the k-means algorithm, with the elbow method applied to determine the optimal number of clusters. The silhouette coefficient and Davies-Bouldin index are the internal validation metrics used to evaluate the number and quality of clustering identified by the K-means. Table 3 lists the k-means metrics for each sensor and case of study.

**Table 3.** k-means clustering identification for each sensor related to the failure.

| **Sensors** | **GEN** | **L05** | **MSH** | **NMF** | **GEN** | **L05** | **MSH** | **NMF** | **GEN** | **L05** | **MSH** | **NMF** | **GEN** | **L05** | **MSH** | **NMF** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Multiclass** | | | | **NO-FC** | | | | **NO-AI** | | | | **NO-RI** | | | |
| Elbow method ($k$) | 4 | 4 | 4 | 4 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Silhouette score | 0.602 | 0.521 | 0.589 | 0.563 | 0.865 | 0.867 | 0.873 | 0.869 | 0.596 | 0.564 | 0.604 | 0.576 | 0.596 | 0.508 | 0.583 | 0.559 |
| Davies-Bouldin | 0.504 | 0.597 | 0.509 | 0.539 | 0.131 | 0.123 | 0.121 | 0.121 | 0.582 | 0.642 | 0.592 | 0.615 | 0.751 | 0.863 | 0.729 | 0.755 |

For binary cases (k=2), the results indicate good separability between operating conditions, with consistent metrics across different components and failure types. In NO-FC scenarios, Silhouette Score values remain high (between 0.86 and 0.87) and Davies-Bouldin indices are low (between 0.12 and 0.13), indicating that points within each group are very similar to each other and that the clusters are well separated. In NO-RI and NO-AI conditions, the results still exhibit a satisfactory clustering structure, with Silhouette values ranging from 0.56 to 0.60 and Davies-Bouldin scores between 0.58 and 0.75, indicating good cluster quality. In general, all binary-scenario cases yielded metrics superior to the reference values recommended in the litera-

ture (Silhouette > 0.5 and Davies-Bouldin < 1.0), confirming that the clustering method can adequately distinguish the different operational states of the turbines. For the multi-class case with k = 4, the validation metrics indicate good clustering quality. Silhouette Score values range from 0.52 to 0.60, and the Davies-Bouldin index values range from 0.50 to 0.60, indicating compact and well-separated clusters. These results confirm that the clustering process identified variations in the operational states of the turbines. Therefore, these metrics provide strong evidence that the clustering step is reliable and that the labels assigned during the unsupervised phase are consistent.

**Table 4.** Explicit description of the datasets used for training, validation, and testing.

| Classification | Cases | Training/Validation | Test |
|---|---|---|---|
| Binary | NO and FC | 1308 | 436 |
| | NO and RI | 941 | 314 |
| | NO and AI | 699 | 234 |
| Multiclass | NO, FC, RI and AI | 2363 | 788 |

The outputs are subsequently used as inputs for the feature and sensor selection. Furthermore, the samples were randomly divided, and the clustered dataset was split into training/validation 75%, and test 25%, which were used for model construction and evaluation, as shown in Table 4. Each classifier was assessed using a 5-fold cross-validation procedure to enhance training accuracy, which was achieved by randomly partitioning the training dataset into five distinct subsets.

### 3.2 Feature and sensor fast selection

The supervised feature analysis indicates that changes in the turbine's dynamic behaviour are associated with specific events. For the subsequent procedures, feature and sensor selection are performed using the canonical-correlation-based feature selection method (Fastcan) introduced by Zhang et al. (2025), which was initially applied only for feature selection and is extended in this work to include sensor selection. In this study, the selection process is carried out in two steps: first, identifying the best feature(s) for each sensor, and second, selecting the most relevant sensor(s). The pseudocode for this process is given in Algorithm 2, which is repeated for each fault diagnosis. The Fastcan method is based on the canonical correlation coefficient (CCA), a criterion used to measure the linear association between two multivariate random variables. It evaluates the relevance of features with respect to the response matrix. Given the feature matrix $\mathbf{X} \in \mathbb{R}^{N \times n}$ and the response matrix $\mathbf{Y} \in \mathbb{R}^{N \times m}$, the canonical correlation coefficient is defined as

$$R_i(\mathbf{X}, \mathbf{Y}) = \max_{\alpha_i, \beta_i} r(\mathbf{X}_C \alpha_i, \mathbf{Y}_C \beta_i), \tag{2}$$

where $\mathbf{X}_C$ and $\mathbf{Y}_C$ are column-centred matrices and $r(\cdot,\cdot)$ denotes the Pearson correlation. The projection vectors $\alpha_i \in \mathbb{R}^n$ and $\beta_i \in \mathbb{R}^m$ are obtained by solving the following eigenvalue problems

$$\left(\mathbf{X}_C^\top \mathbf{X}_C\right)^{-1} \mathbf{X}_C^\top \mathbf{Y}_C \left(\mathbf{Y}_C^\top \mathbf{Y}_C\right)^{-1} \mathbf{Y}_C^\top \mathbf{X}_C \, \alpha_i = R_i^2(\mathbf{X},\mathbf{Y}) \, \alpha_i, \tag{3}$$

$$\left(\mathbf{Y}_C^\top \mathbf{Y}_C\right)^{-1} \mathbf{Y}_C^\top \mathbf{X}_C \left(\mathbf{X}_C^\top \mathbf{X}_C\right)^{-1} \mathbf{X}_C^\top \mathbf{Y}_C \, \beta_i = R_i^2(\mathbf{X},\mathbf{Y}) \, \beta_i. \tag{4}$$

The eigenvalues $R_i^2(\mathbf{X},\mathbf{Y})$ correspond to the squared canonical correlation coefficients, while $\alpha_i$ and $\beta_i$ are the canonical weight vectors that define the canonical variates $\mathbf{X}_C\alpha_i$ and $\mathbf{Y}_C\beta_i$. There are at most $\min(n,m)$ non-zero canonical correlation coefficients $R_1(\mathbf{X},\mathbf{Y}),\ldots,R_{\min(n,m)}(\mathbf{X},\mathbf{Y})$. The overall criterion for feature selection is the sum of squared canonical correlations (SSC)

$$SSC(\mathbf{X},\mathbf{Y}) = \sum_{k=1}^{\min(n,m)} R_k^2(\mathbf{X},\mathbf{Y}). \tag{5}$$

To accelerate computation, the SSC can be reformulated using two equivalent decompositions: (i) the $h$-correlation form (valid when $N \leq n+m$),

$$SSC(\mathbf{X},\mathbf{Y}) = \sum_{i=1}^{n}\sum_{j=1}^{m} r^2(w_i, v_j), \tag{6}$$

where $\{w_i\}$ and $\{v_j\}$ are orthogonal bases of the feature and response spaces; and (ii) the $\theta$-angle form (valid when $N > n+m$),

$$SSC(\mathbf{X},\mathbf{Y}) = \sum_{i=1}^{n}\sum_{j=1}^{m} \cos^2\big(\angle(w_i, v_j)\big), \tag{7}$$

where the canonical correlations are expressed as squared cosines of the principal angles between the subspaces of $\mathbf{X}$ and $\mathbf{Y}$. In the greedy selection step, the next feature $x_{ri}$ is chosen to maximise the incremental SSC

$$d = \arg\max_i \sum_{k=1}^{\min(p+1,m)} R_k^2\big((\mathbf{X}_s, x_{ri}), \mathbf{Y}\big), \tag{8}$$

where $\mathbf{X}_s$ denotes the already selected features. This formulation allows efficient feature ranking by reducing computational complexity while preserving the theoretical properties of canonical correlation analysis. This process involves looping through the features for each sensor in an interaction. In sequence, a similar procedure is performed for sensor ranking, with the feature matrix defined as $\mathbf{X}_s \in \mathbb{R}^{S\times s}$, where $\mathbf{X}_s$ contains the highest ranking feature(s) of each sensor.

### 3.3 Classification and model evaluation

ML algorithms are applied to classify the operational condition of the wind turbine by analysing datasets. As an automated approach, ML identifies patterns in data through various algorithms and uses these learned patterns for predictive analysis and decision-making. The proposed model follows a hybrid unsupervised-supervised ML-based framework, incorporating classification strategies to improve the structure of data attributes. The data-driven ML models are fed selected features extracted

**Algorithm 2** Pseudocode for the feature and sensor selection with canonical correlation

**Input:** $\mathbf{X} \in \mathbb{R}^{N \times n}$ Candidate feature matrix
**Input:** $\mathbf{Y} \in \mathbb{R}^{N \times m}$ Target multi-label matrix
**Output:** $\mathbf{d} \in \mathbb{R}^{1 \times \mathcal{T}}$, where $\mathbf{d} = (d_1, d_2, \ldots, d_t)$ Index of the feature with the highest ranking score
**Output:** $\mathbf{q} \in \mathbb{R}^{1 \times s}$, where $\mathbf{q} = (q_1, q_2, \ldots, q_s)$ Index of the feature with the highest ranking score associated to the sensor

$\mathcal{T} \in N$ Number of features to be selected
s $\in$ S Number of sensor
$\mathrm{X}_s \leftarrow 0$ Candidate feature matrix coating the best feature(s) of each sensor. The threshold is Score $\geq 0.6$
**for** $j \leftarrow 1$ to s **do**
  $\mathbf{X}_C \leftarrow [\mathrm{X} - \mathrm{mean}(\mathrm{X})]$, $\mathbf{Y}_C \leftarrow [\mathbf{Y} - \mathrm{mean}(\mathbf{Y})]$
  **if** use $\eta$-cosine **then**
    $[\mathbf{U}, \mathbf{S}, \mathbf{V}_h^\top] \leftarrow \mathrm{svd}([\mathbf{X}_C \mid \mathbf{Y}_C])$
    $[\mathbf{X}_C]_U \mid [\mathbf{Y}_C]_U \leftarrow \mathbf{S}\mathbf{V}_h^\top$
  **end if**
  $R_i^2 \leftarrow 0$ for $i = 1, \ldots, n$ Feature ranking scores
  $d_i \leftarrow 0$ for $i = 1, \ldots, t$ Output indices vector, where $t \in N$ is the number of features to be selected
  **procedure** FASTCAN(X, Y, t, tol, alg) Function described in (Zhang et al., 2025)
    $d \leftarrow \arg\max_j R_j^2$
  **return d**
  $\mathrm{X}_s \leftarrow \mathbf{X[:, \max(d)]}$ **Highest score feature**
**end for**
**procedure** FASTCAN($\mathrm{X}_s$, Y, s, tol, alg)
  q $\leftarrow \arg\max_j R_j^2$
**return** q

from the multiphysics database, incorporating vibration responses and environmental data, and the output corresponds to the wind turbine's operational classification state.

For the classification task, six supervised ML algorithms are utilised, and the unsupervised k-means algorithm is used for initial clustering. At the same time, the supervised classifiers naive bayes (NB), Decision tree (DT), random forest (RF), kNN, SVM, and extreme gradient boosting (XGB) are employed to detect the system's operational conditions. The ML algorithms embedded in the framework are based on the open-source Scikit-learn library. These algorithms perform final classification based on the initial clustering results, producing outputs including confusion matrices and performance metrics. The selection of hyperparameters follows the recommendations in (de Sousa et al., 2023; Coelho et al., 2024, 2025), which identified optimal configurations for this application. Specifically, the SVM model employs a linear kernel with a penalty parameter of C = 100, a one-vs-one multiclass strategy, and a tolerance of $1 \times 10^{-3}$. For kNN, the number of neighbours is set to k = 3, using the Euclidean distance metric, uniform weights, and a leaf size of 30. The RF and DT algorithms both use 100 trees, a maximum

depth of 3, and the Gini splitting criterion. The Naive Bayes classifier employs a Gaussian model, while XGBoost uses the XGBClassifier implementation. These configurations have shown high accuracy and robustness in prior studies, and are adopted here to optimise classification performance.

The final step involves evaluating model performance on a previously separated test dataset. Hyperparameters may be fine-tuned to improve Accuracy, Precision, Recall, and F1-score. A confusion matrix is also analysed to provide further diagnostic insight. To prevent overfitting and ensure generalisability, a 5-fold cross-validation scheme is employed. Performance metrics and confusion matrices are generated for all algorithms, enabling the framework to identify the most accurate model and guide users in selecting the most suitable ML algorithm for wind turbine monitoring.

## 4 Results and discussion

During the monitoring process, data from various sensors are available. To ensure accurate failure identification and classification, it is essential to evaluate the most informative sensor(s) data. In the proposed framework, both feature and sensor are ranked and identified to compose the reduced dataset, ensuring that only the most sensitive information related to the specific failure is selected and further input to the ML algorithm for wind turbine condition classification. The feature and sensor scores, normalised between zero and one, are returned from the ranking process performed by the CCA. Further, the dataset for the ML algorithms was built by selecting the top-ranked features from each sensor, as well as the SCADA wind speed and temperature. The evaluation of the turbine's operational condition includes both binary and multiclass classification tasks.Table 5 summarises the failure cases related to the most sensitive feature of each sensor ($X_s$), and the most sensitive sensor associated with its features for each failure.

**Table 5.** Explicit description of the datasets used for training, validation, and testing.

| Cases | CCA selected feature of each sensor $X_s$ | Ranked sensor | Ranked sensor feature |
|---|---|---|---|
| NO-FC | $RMS_Y01^{MSH}$, $Max_Y01^{L05}$, $Amp_Z01^{GEN}$, $RMS_X02^{NMF}$ | NMF | $RMS_X02^{NMF}$ |
| NO-RI | $RMS_X01^{MSH}$, $Var_Y01^{L05}$, $RMS_Y01^{GEN}$, $RMS_X02^{NMF}$ | MSH | $RMS_X02^{MSH}$ |
| NO-AI | $RMS_Y01^{MSH}$, $RMS_X01^{L05}$, $Var_Z01^{GEN}$, $RMS_X02^{NMF}$ | NMF | $RMS_X02^{NMF}$ |
| Multiclass | $RMS_X01^{MSH}$, $Var_Y01^{MSH}$, $Var_X01^{L05}$, $Max_X02^{L05}$, $Max_Y01^{L05}$, $Max_Z01^{GEN}$, $Var_Z01^{GEN}$, $RMS_X02^{NMF}$, $Var_X02^{NMF}$ | GEN<br>NMF | $Max_Z01^{GEN}$<br>$RMS_X02^{NMF}$<br>$Var_Z01^{GEN}$ |

For binary classification between NO-FC failure,the individual feature scores for each accelerometer sensor are shown in Fig. 5(a-d). Subsequently, in the sensor ranking, the NMF sensor ($RMS_X02^{NMF}$) yields the highest score, indicating it as the most sensitive to FC-related failures, as shown in Fig. 8(a). The k-means algorithm is applied to label and organise the data based on similarities among the selected features. Incorporating a larger number of informative features enhances both pattern recognition and classification accuracy. Each feature is visualised individually for clarity, the final classification integrates all

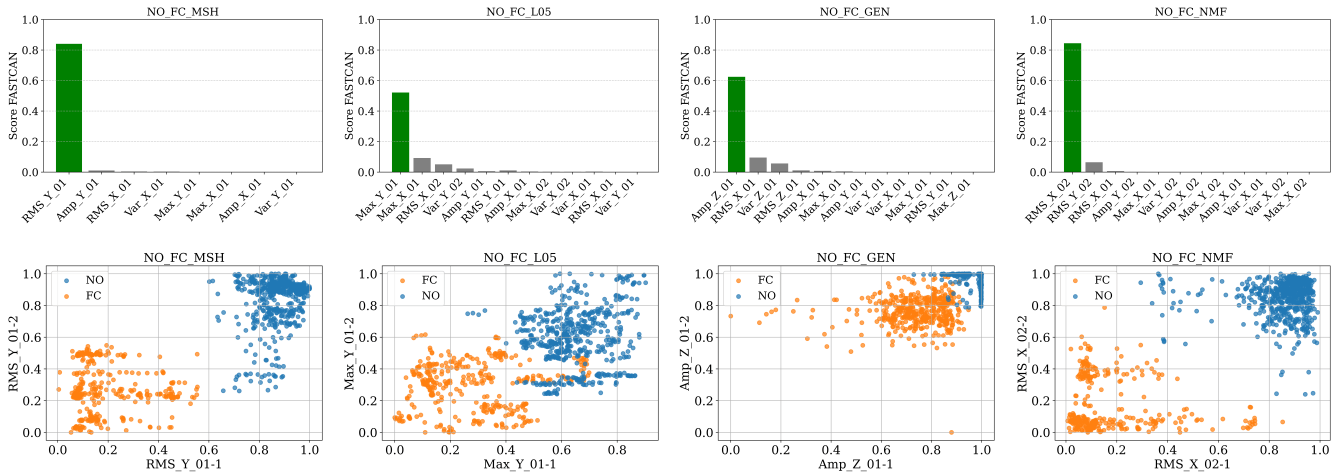



**Figure 5.** CCA feature score, and dispersion diagrams of the highest-scoring feature: (a) NO-FC sensor MSH, (b) NO-FC sensor L05 (1 and 2), (c) NO-FC sensor GEN, and (d) NO-FC sensor NMF (1 and 2).

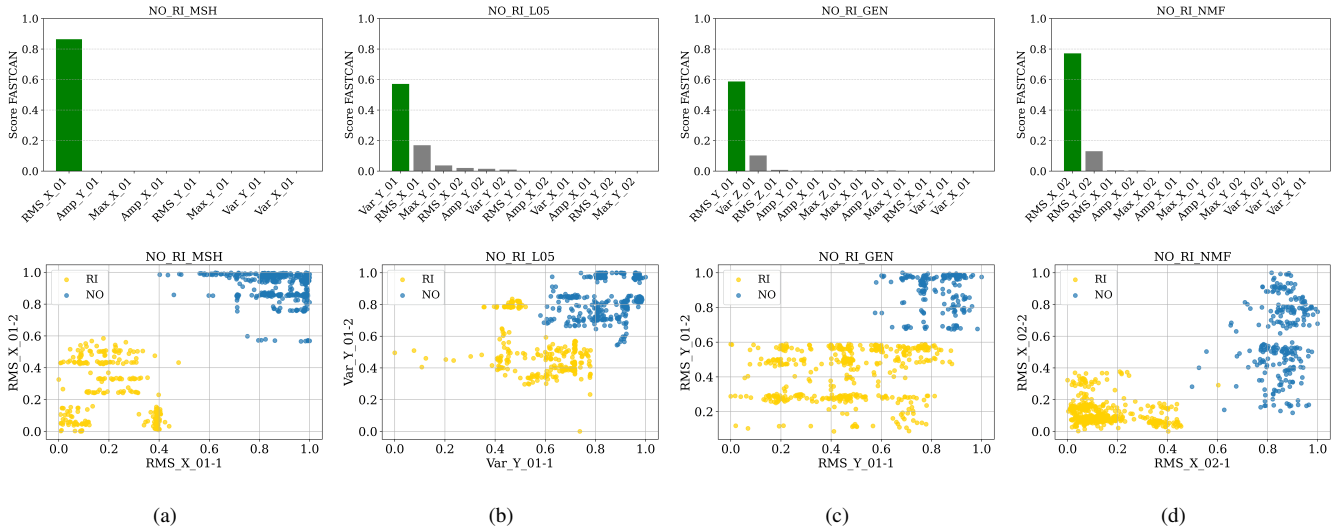



**Figure 6.** CCA feature score, and dispersion diagrams of the highest-scoring feature: (a) NO-RI sensor MSH, (b) NO-RI sensor L05 (1 and 2), (c) NO-RI sensor GEN, and (d) NO-RI sensor NMF (1 and 2).

selected features along with SCADA data to strengthen the overall decision-making process. Features with minimal overlap between operational states are especially valuable, as they contribute to clearer distinctions and improved ML performance in identifying fault conditions. Scatter plots in Fig. 5(e-h) display the correlation between Feature-1 and Feature-2, which

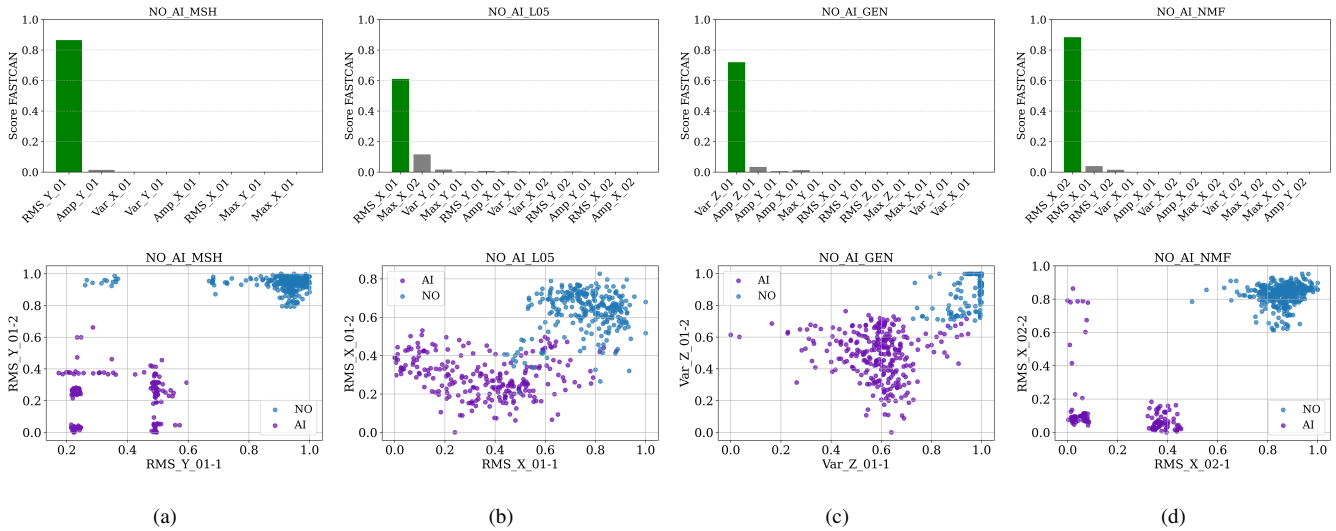



**Figure 7.** CCA feature score, and dispersion diagrams of the highest-scoring feature: (a) NO-AI sensor MSH, (b) NO-AI sensor L05 (1 and 2), (c) NO-AI sensor GEN, and (d) NO-AI sensor NMF (1 and 2).

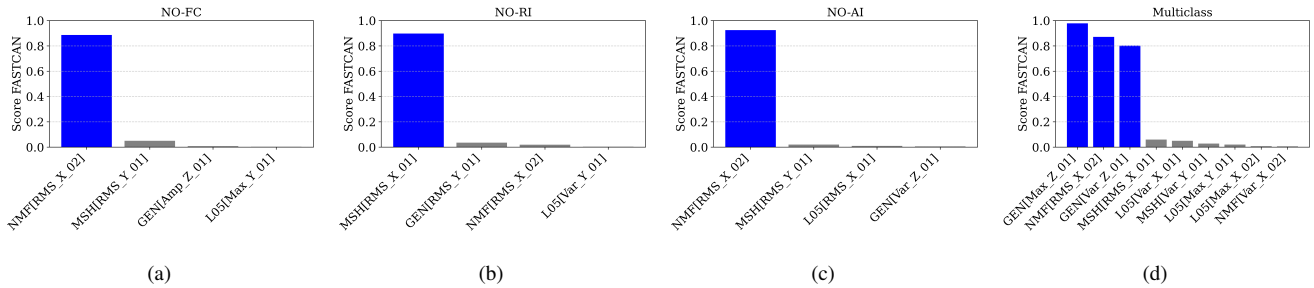


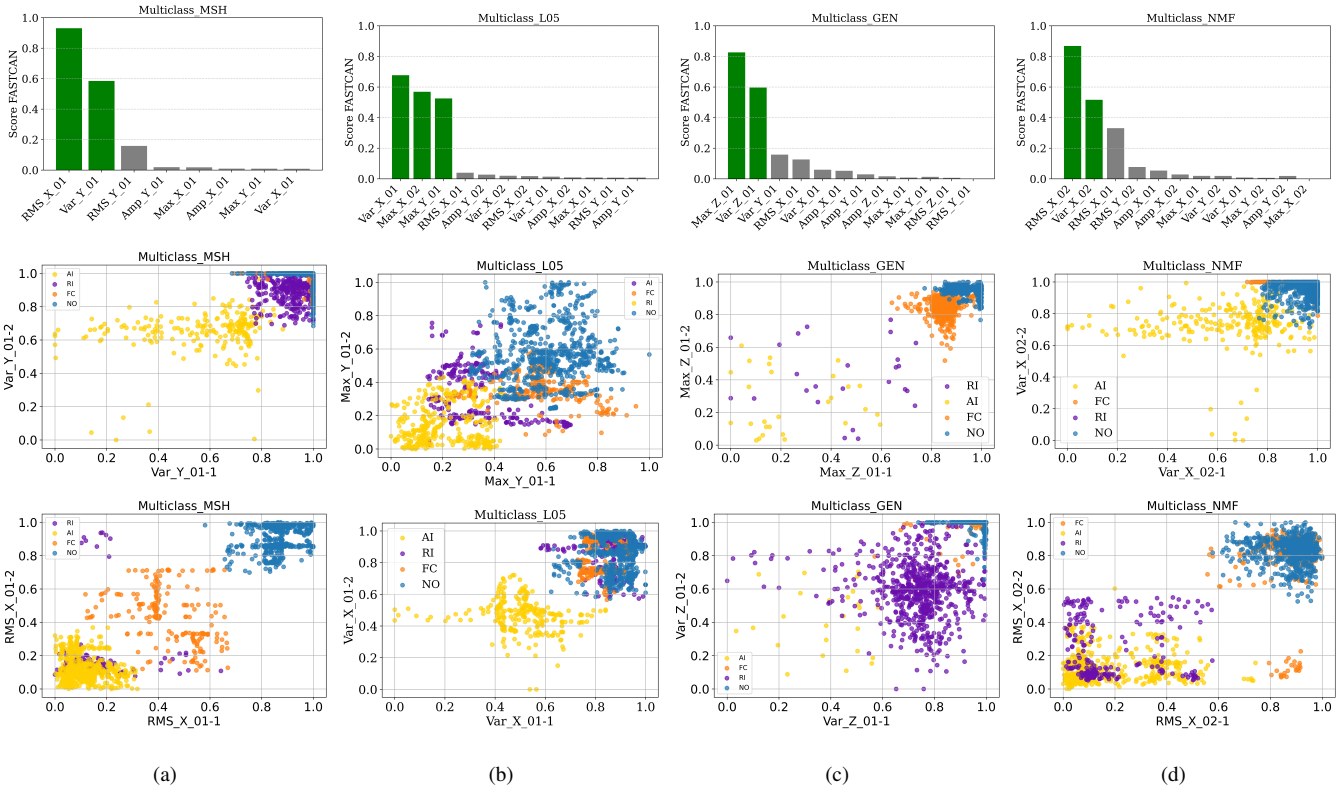

**Figure 8.** Sensor score based on the top-performing features of each operation condition: (a) NO-FC, (b) NO-AI, (c) NO-RI, and (d) multi.

represents two groups derived from splitting each selected feature column. These diagrams highlight how high-scoring features cluster across sensors, revealing patterns that differentiate between normal and faulty operating conditions. The identified clusters exhibit consistent grouping and dispersion, thereby enhancing the interpretability of each operational condition.

For classification between NO-RI, the feature scores for each accelerometer are shown in Fig. 6(a-d) and Table 5, with the MSH sensor (RMS_X02$^{MSH}$) achieving the highest score, indicating strong sensitivity to RI-related anomalies, as indicated in Fig. 8(b). The k-means clustering is applied to the selected features to identify patterns and assign labels. Scatter plots in Fig. 6(e–h) display the highest-scoring feature pairs (feature-1 vs. feature-2) per sensor, revealing clear cluster separation between NO and RI. These dispersion diagrams highlight how specific failures impact dynamic behaviour, with minimal overlap between classes indicating highly informative features for classification. For the NO-AI, the feature ranking scores of each accelerometer are presented in Fig. 7(a-d), where the NMF sensor (RMS_X02$^{NMF}$) shows the highest relevance, indicating

strong responsiveness to RI-related conditions, Fig. 8(c). k-means clustering is applied to the selected features to identify patterns and assign labels. The resulting scatter plots in Fig. 7(e-h) show the top-ranked feature pairs (Feature-1 vs. Feature-2) for each sensor, revealing distinct cluster boundaries between NO and AI. These diagrams illustrate how different faults affect the turbine's dynamic behaviour, with low class overlap indicating that the selected features are highly effective for classification.

In the multiclass classification, the k-means algorithm identifies four clusters corresponding to NO, FC, RI, and AI conditions, organising the dataset accordingly into response clusters. These clusters support further validation of classification through performance metrics. Figures 9 display the highest-scoring features of each sensor and their related dispersion diagrams, labelled by k-means and ranked by Algorithm 2, respectively. Each class is colour-coded to facilitate quick visual identification and analysis of overlaps between operational states.




(a) (b) (c) (d)

**Figure 9.** CCA feature score, and dispersion diagrams of the highest-scoring feature: (a) sensor MSH, (b) sensor L05 (1 and 2), (c) sensor GEN, and (d) sensor NMF (1 and 2).

Figure 9(a-d) gives the top-ranked features from each sensor, MSH, L05, NMF, and GEN, respectively, highlighting the most sensitive sensors selected for multiclass classification. Likewise, in the binary case, features with higher scores are included in the global dataset for the classification, where the GEN and NMF sensors show the highest relevance, indicating strong sensi-

tivity, as shown in Fig. 8(d). Each selected feature is visualised through its dispersion correlation in Fig. 9(a-c), corresponding to results from sensors GEN, MSH, and NMF, respectively. Figures 9(d-f), in turn, show the results for L05 sensors. Overall, k-means effectively labels classes associated with each operational fault. Certain features exhibit a clear separation between normal conditions and fault types, although overlap among fault classes is evident in all cases. This suggests that while individual features may help detect abnormal behaviour, they are not self-sufficient alone to discriminate between specific failure modes. These individual feature analyses reveal important patterns. However, when classification models are applied, features are considered in combination. This integrated approach improves both classification accuracy and the robustness of the performance metrics. While some features aid in distinguishing normal from faulty states, reliable fault classification generally requires the combined contribution of multiple informative features.

### 4.1 Operational condition classification and metrics evaluation

The fault classification uses six machine learning algorithms: kNN, SVM, DT, RF, Naive Bayes, and XGBoost. These algorithms perform the final classification using the input dataset, producing confusion matrices and performance metrics. The selection of hyperparameters is based on the investigations of (de Sousa et al., 2023), which specify optimal configurations for this assignment and are briefly described in Section 3.3. Figure 10 presents the performance of ML models for binary classification (NO-FC, NO-RI, and NO-AI) and multiple failure conditions, evaluated via cross-validation and metrics such as accuracy, precision, recall, and F1-score. For metric estimation, micro-, macro-, and weighted averages were tested, assuming Micro as the standard because it computes classification metrics by globally summing true positives, false positives, and false negatives across all classes, resulting in similar Precision, Recall, and F1-score.

The ML models were tested on the *Fe-kms-Sc*, *Fe-Sc-kms*, *Fe-kms*, and *Fe-kms-LoseSensor* datasets to assess the influence of environmental dependencies (SCADA data) and the sensitivity of each sensor in failure evaluation. The metrics for the *Fe-Sc-kms* dataset, shown as black ($*$) in Fig. 10(a-d), indicate that, except for the NB classifier, all ML models achieved 100% accuracy. Such a perfect performance across multiple models can suggest potential issues, such as data leakage, overfitting, or improper dataset splitting. However, cross-validation with multiple random partitions was performed to ensure statistical robustness. In this case, the consistently high accuracy reflects the physical consistency and strong discriminative power of the selected features rather than methodological flaws.

The k-means clustering results and associated metrics also reveal clear class separation. The feature scores derived from the CCA, which quantify the linear association between the selected features and the k-means clusters (Fig. 11), highlight the strong correlation between SCADA data and the structural sensors (accelerometers). The SCADA system provides environmental variables, such as daily temperature and wind speed, but lacks important structural information more directly related to damage states. In the *Fe-Sc-kms* dataset configuration, both features and SCADA inputs are used in the k-means clustering, which is predominantly influenced by the SCADA parameters. Therefore, the perfect ML metrics are attributed to dataset bias, where the models primarily classify failure conditions based on environmental variations rather than the structural response itself.

The models' metrics for the *Fe-kms-Sc* dataset, shown as blue ($\bullet$) in Fig. 10(a-d), indicate the sensitivity of the ML models to the damage. The k-means is performed on the structural features, and the SCADA follows its data reorganisation, but it is

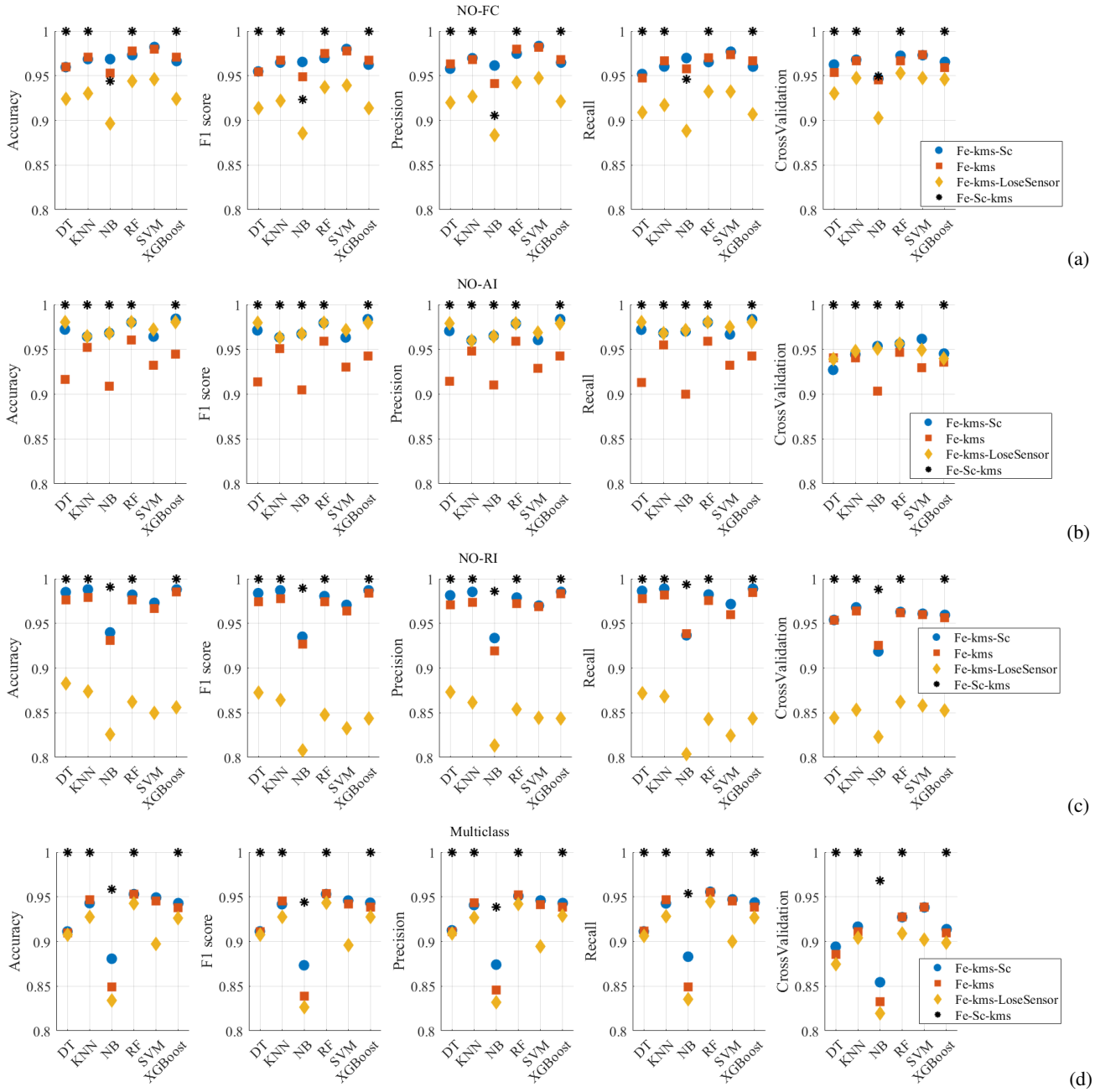



**Figure 10.** Comparison of the metrics (accuracy, F1-score, precision, recall, and cross-validation) for the six ML algorithms and each arranged final dataset. (a) Binary NO-FC, (b) Binary NO-AI, (c) Binary NO-RI, and (d) multiclass failure study cases

not directly considered in the k-means labelling. For the binary case, the model's metrics range from 0.93 to 0.99. The SVM model reached 0.98 on the NO-FC case, XGBoost reached 0.98 for NO-AI, and XGBoost reached 0.99 for NO-RI. Multiclass

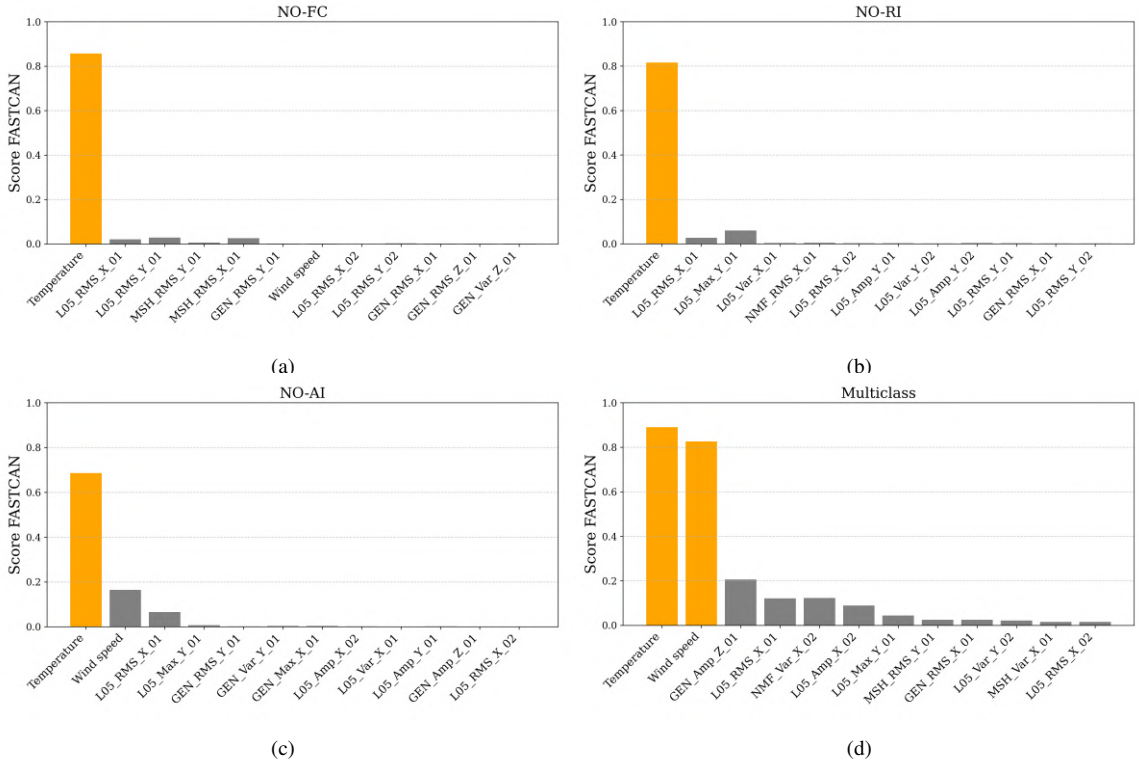
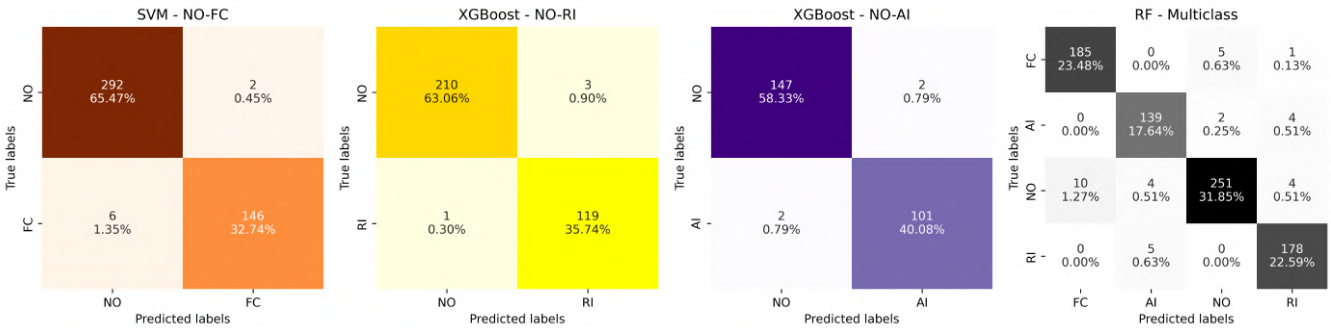



**Figure 11.** Feature score of the *Fe-kms-Sc* dataset for (a) NO-FC, (b) NO-RI, (c) NO-AI, and (d) multiclass.

case varied from 0.86 given by NB to 0.955 by RF. The SCADA data are not included in the *Fe-kms* dataset, shown as orange (■) in Fig. 10(a-d). The performance metrics reached 0.98 with the SVM model for the NO-FC case, 0.955 with RF for the NO-AI case, 0.98 with XGBoost for the NO-RI case, and 0.95 with RF in the multiclass analysis. When the most sensitive sensor for each failure case was removed from the *Fe-kms-LoseSensor* dataset (yellow ♦), performance metrics decreased, except for NO-AI, indicating the importance of this sensor's information for monitoring accuracy and model performance. This also reinforces that the most sensitive sensor is typically located near the damage site. In the NO-AI case, all sensors were positioned on the nacelle and tower. Although these sensors can capture the dynamic effects of blade aero-imbalance, they are distant from the local damage, which reduces their sensitivity to this phenomenon. Consequently, the sensor group primarily captured global structural responses to the fault, rather than local damage effects, which explains the variation in performance when one of the sensors was removed.

The confusion matrices shown in Figs. 12(a-c) present the results of binary classification using the respective best-performing ML algorithms using the Fe-kms-Sc dataset. In the NO-FC case, the SVM model achieved optimal accuracy and cross-validation, correctly classifying 292 samples as NO (65.47%) and 146 as FC (32.74%). For the NO-RI classification, the XGBoost model correctly classifies 210 samples as NO (63.06%) and 119 as RI (35.74%). In the NO-AI scenario, the XG-




(a) (b) (c) (d)

**Figure 12.** Confusion matrix of binary operational classification for (a) NO-FC, (b) NO-RI, (c) NO-AI, and (d) for multi-classification. The dataset assumes the best features plus SCADA data (temperature and wind speed) in each analysis.

Boost model outperformed the others, correctly classifying 101 samples as AI (40.08%) and 147 as NO (58.33%). Figure 12(d) displays the confusion matrix for the multiclass classification involving AI, FC, NO, and RI. The RF model correctly identified 139 samples as AI (17.64%), 185 as FC (23.48%), 178 as RI (22.59%), and 251 as NO (31.85%).

Among the evaluated configurations, the *Fe-kms-Sc* dataset provided the most balanced and reliable results for the treated failure identification, including structural and environmental condition information. The dataset *Fe-Sc-kms* yielded perfect metrics across all models, primarily because environmental parameters, such as temperature and wind speed, were dominant, biasing the classification toward environmental conditions rather than structural responses. In contrast, excluding SCADA data (*Fe-kms*) allowed the models to capture the structural dynamics associated with damage more accurately, but environmental variability was not considered. The *Fe-kms-LoseSensor* results highlighted the critical role of the most sensitive sensor, typically located near the damage, in maintaining model accuracy. Overall, the *Fe-kms-Sc* dataset, which integrates structural features with aligned but non-dominant environmental information, yielded the most physically consistent and discriminative performance for the wind turbine failure classification.

## 5 Conclusions

This study proposes a data-driven hybrid framework for classifying operational conditions of a wind turbine, including normal operation, pitch-drive faults, rotor icing, and aerodynamic imbalance, integrating in a hybrid approach multimodal data and multiple ML algorithms. The monitoring process follows the pipeline: data processing and organisation, feature extraction, normalisation, and grouping, clustering, feature and sensor selection, data splitting, machine learning classification, and model evaluation. A novel relative change damage index was introduced to enhance scalability and normalise features extracted from structural dynamic responses and environmental conditions. Canonical correlation analysis was used to identify and rank the most sensitive features among the fifteen extracted from temporal responses and SCADA data (wind speed and temperature), and sensor ranking. Thus, multimodal information, including vibration signals from six accelerometers distributed across

the turbine and environmental parameters(wind speed and temperature), was incorporated into the framework. Unsupervised k-means clustering enabled the discovery of homogeneous the need for data groups, supporting robust pattern recognition without predefined labels.

Failure classification was implemented as both binary and multiclass tasks using kNN, SVM, DT, RF, Naive Bayes, and XGBoost. Models were tested on the *Fe-kms-Sc*, *Fe-Sc-kms*, *Fe-kms*, and *Fe-kms-LoseSensor* datasets to evaluate environmental influence and sensor sensitivity. Using the *Fe-kms-Sc* dataset, the models performed best, where SVM achieved the highest accuracy for NO-FC (0.98), XGBoost for NO-RI (0.99) and NO-AI (0.98), and RF for multiclass classification (0.955). The *Fe-kms-Sc* dataset, which combines structural features with aligned SCADA data, yielded the most reliable failure detection. Excluding SCADA data (*Fe-kms*), the models focused on structural changes, with binary accuracies of 0.93-0.99 and multiclass accuracies of 0.86-0.95. Including SCADA alone (*Fe-Sc-kms*) produced perfect metrics due to environmental dominance, whereas removing the most sensitive sensor (*Fe-kms-LoseSensor*) reduced performance, confirming the importance of sensor placement near the damage and identification of the most sensitive sensor in each analysis.

The proposed hybrid model, developed and tested on a 6.7kW Aventa turbine, effectively manages the complexity of multi-source turbine data and representative fault variations. Its success relies on careful feature and sensor selection, ML model selection, inclusion of environmental data, dataset multimodality, and thoughtful dataset arrangement, which together enhance discriminative power and classification reliability. However, validation is limited by the small size of the wind turbine, the use of induced faults to simulate aero-imbalance in the blades, and differences between small-scale experimental turbines and large-scale operational turbines or wind farms. Extreme environmental and operational events, such as strong turbulence, large-scale icing, and mixed-fault scenarios, were not included in this study due to the controlled nature and limited scope of the experimental datasets. Ongoing studies aim to evaluate the methodology on larger, more diverse datasets, further explore environmental effects, extend the study to offshore wind farms, and assess reliability, generalisation, and applicability to utility-scale systems.

*Author contributions.* All authors contribute equally to this research and to the preparation of this paper.

*Competing interests.* The authors declare that no competing interests are present

*Acknowledgements.* This research is part of project No. 444595/2024-4 funded by the National Council for Scientific and Technological Development-CNPq, call CNPq/MCTI/FNDCT Nº 22/2024. M.R.M acknowledges the sponsors of projects (CNPq, Capes Finance Code 001, and FAPDF.00193-00002143/2023-02).

## References


Amin, A., Bibo, A., Panyam, M., and Tallapragada, P.: Vibration based fault diagnostics in a wind turbine planetary gearbox using machine learning, Wind Engineering, 47, 175–189, https://doi.org/10.1177/0309524X221123968, 2023.

Antoniadou, I., Dervilis, N., Barthorpe, R. J., Manson, G., and Worden, K.: Advanced tools for damage detection in wind turbines, Key Engineering Materials, 569-570, 547–554, https://doi.org/10.4028/www.scientific.net/KEM.569-570.547, 2013.

Antoniadou, I., Dervilis, N., Papatheou, E., Maguire, A. E., and Worden, K.: Aspects of structural health and condition monitoring of offshore wind turbines, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, https://doi.org/10.1098/rsta.2014.0075, 2015a.

Antoniadou, I., Dervilis, N., Papatheou, E., Maguire, A. E., and Worden, K.: Aspects of structural health and condition monitoring of offshore wind turbines, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20140075, https://doi.org/10.1098/rsta.2014.0075, 2015b.

Ashkarkalaei, M., Ghiasi, R., Pakrashi, V., and Malekjafarian, A.: Feature selection for unsupervised defect detection of a wind turbine blade considering operational and environmental conditions, Mechanical Systems and Signal Processing, 230, 112568, https://doi.org/https://doi.org/10.1016/j.ymssp.2025.112568, 2025a.

Ashkarkalaei, M., Ghiasi, R., Pakrashi, V., and Malekjafarian, A.: Feature selection for unsupervised defect detection of a wind turbine blade considering operational and environmental conditions, Mechanical Systems and Signal Processing, 230, 112568, https://doi.org/https://doi.org/10.1016/j.ymssp.2025.112568, 2025b.

Ashkarkalaei, M., Ghiasi, R., Pakrashi, V., and Malekjafarian, A.: Optimum feature selection for the supervised damage classification of an operating wind turbine blade, Structural Health Monitoring, 0, 1475, https://doi.org/10.1177/14759217251313815, 2025b.

Asian, S., Ertek, G., Haksoz, C., Pakter, S., and Ulun, S.: Wind Turbine Accidents: A Data Mining Study, IEEE Systems Journal, 11, 1567–1578, https://doi.org/10.1109/JSYST.2016.2565818, 2017.

Baquerizo, J., Tutivén, C., Puruncajas, B., Vidal, Y., and Sampietro, J.: Siamese Neural Networks for Damage Detection and Diagnosis of Jacket-Type Offshore Wind Turbine Platforms, Mathematics, 10, https://doi.org/10.3390/math10071131, 2022.

Barreto, L. S., Machado, M. R., Santos, J. C., Moura, B. B. d., and Khalij, L.: Damage indices evaluation for one-dimensional guided wave-based structural health monitoring, Latin American Journal of Solids and Structures, 18, e354, https://doi.org/10.1590/1679-78256292, 2021.

Carmona-Troyo, J. A., Trujillo, L., Enríquez-Zárate, J., Hernandez, D. E., and Cárdenas-Florido, L. A.: Classification of Damage on Wind Turbine Blades Using Automatic Machine Learning and Pressure Coefficient, Expert Systems, 42, e70024, https://doi.org/https://doi.org/10.1111/exsy.70024, 2025.

Chatzi, E., Abdallah, I., Hofsäß, M., Bischoff, O., Barber, S., and Marykovskiy, Y.: Aventa AV-7 ETH Zurich Research Wind Turbine SCADA and High-Frequency Structural Health Monitoring (SHM) Data, https://doi.org/10.5281/zenodo.8229750, 2023.

Christopher M. Bishop: Pattern Recognition and Machine Learning, Springer, 2nd edn., ISBN 978-1493938438, 2006.

Civera, M. and Surace, C.: Non-Destructive Techniques for the Condition and Structural Health Monitoring of Wind Turbines: A Literature Review of the Last 20 Years, Sensors, 22, https://doi.org/10.3390/s22041627, 2022.

Coelho, J. S., Machado, M. R., Dutkiewicz, M., and Teloli, R. O.: Data-driven machine learning for pattern recognition and detection of loosening torque in bolted joints, https://doi.org/10.1007/s40430-023-04628-6, 2024.

Coelho, J. S., Machado, M. R., and Dutkiewicz, M.: Integrating Virtual Sensor Data Augmentation Into Machine Learning for Damage Quantification of Bolted Structures Under Assembly Uncertainty, Structural Control and Health Monitoring, 2025, 8030 303, https://doi.org/https://doi.org/10.1155/stc/8030303, 2025.

de Sousa, A. A. S. R., da Silva Coelho, J., Machado, M. R., and Dutkiewicz, M.: Multiclass Supervised Machine Learning Algorithms Applied to Damage and Assessment Using Beam Dynamic Response, Journal of Vibration Engineering and Technologies, 11, 2709–2731, https://doi.org/10.1007/s42417-023-01072-7, 2023.

Dervilis, N., Choi, M., Antoniadou, I., Farinholt, K. M., Taylor, S. G., Barthorpe, R. J., Park, G., Farrar, C. R., and Worden, K.: Machine learning applications for a wind turbine blade under continuous fatigue loading, Key Engineering Materials, 588, 166–174, https://doi.org/10.4028/www.scientific.net/KEM.588.166, 2014a.

Dervilis, N., Choi, M., Taylor, S., Barthorpe, R., Park, G., Farrar, C., and Worden, K.: On damage diagnosis for a wind turbine blade using pattern recognition, Journal of Sound and Vibration, 333, 1833–1850, https://doi.org/https://doi.org/10.1016/j.jsv.2013.11.015, 2014b.

Dwivedi, D., Babu, K. V. S. M., Yemula, P. K., Chakraborty, P., and Pal, M.: Identification of surface defects on solar PV panels and wind turbine blades using attention based deep learning model, Engineering Applications of Artificial Intelligence, 131, 107 836, https://doi.org/https://doi.org/10.1016/j.engappai.2023.107836, 2024.

Eleni Chatzi, and Imad Abdallah, a. M. H., Bischoff, O., Barber, S., and Marykovskiy, Y.: Aventa AV-7 ETH Zurich Research Wind Turbine SCADA and high frequency Structural Health Monitoring (SHM) data , Zenodo plataform, https://doi.org/10.5281/zenodo.8229750, 2023.

Elforjani, M.: Diagnosis and prognosis of real world wind turbine gears, Renewable Energy, 147, 1676–1693, https://doi.org/https://doi.org/10.1016/j.renene.2019.09.109, 2020.

Farrar, C. R. and Worden, K.: Structural Health Monitoring: A Machine Learning Perspective, Wiley, ISBN 9781118443118, https://doi.org/10.1002/9781118443118, 2012.

Figueiredo, E., Park, G., Farrar, C. R., Worden, K., and Figueiras, J.: Machine learning algorithms to damage detection under operational and environmental variability, Health Monitoring of Structural and Biological Systems, p. 7650, https://doi.org/10.1117/12.849189, 2010.

Flah, M., Nunez, I., Chaabene, W. B., and Nehdi, M.: Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review, Archives of Computational Methods in Engineering, 28, 2621 – 2643, https://doi.org/10.1007/s11831-020-09471-9, 2020.

Gao, Y., Zhai, P., and Mosalam, K. M.: Balanced semisupervised generative adversarial network for damage assessment from low-data imbalanced-class regime, Computer-Aided Civil and Infrastructure Engineering, 36, 1094, https://doi.org/10.1111/mice.12741, 2021.

Hoxha, E., Vidal, Y., and Pozo, F.: Damage Diagnosis for Offshore Wind Turbine Foundations Based on the Fractal Dimension, Applied Sciences, 10, https://doi.org/10.3390/app10196972, 2020.

Joshuva., A. and Sugumaran., V.: A data driven approach for condition monitoring of wind turbine blade using vibration signals through best-first tree algorithm and functional trees algorithm: A comparative study, ISA Transactions, 67, 160–172, https://doi.org/https://doi.org/10.1016/j.isatra.2017.02.002, 2017.

Joshuva, A., Vishnuvardhan, R., Deenadayalan, G., Sathishkumar, R., and Sivakumar, S.: Implementation of rule based classifiers for wind turbine blade fault diagnosis using vibration signals, International Journal of Recent Technology and Engineering, 8, 320–331, https://doi.org/10.35940/ijrte.B1050.0982S1119, 2019.

Knittel, D., Makich, H., and Nouari, M.: Milling diagnosis using artificial intelligence approaches, Software Impacts, 20, 809, 2019.

Korolis, J. S., Bourdalos, D. M., and Sakellariou, J. S.: Machine Learning-Based Damage Diagnosis in Floating Wind Turbines Using Vibration Signals: A Lab-Scale Study Under Different Wind Speeds and Directions, Sensors, 25, https://doi.org/10.3390/s25041170, 2025.

Krishna, S. R., Sathish, J., Tarun, M., Rahul Mani Datta, T., and Vamsi, S. Raghu, a. S. S. J.: Machine Learning-Enabled Crack Diagnosis and Prognosis in Glass/Carbon Fiber Composite Structures, Iranian Journal of Science and Technology, Transactions of Mechanical Engineering, 49, 1511–1530, https://doi.org/https://doi-org.ez54.periodicos.capes.gov.br/10.1007/s40997-025-00855-5, 2025.

Kuai, H., Civera, M., Coletta, G., Chiaia, B., and Surace, C.: Cointegration strategy for damage assessment of offshore platforms subject to wind and wave forces, Ocean Engineering, 304, 117 692, https://doi.org/https://doi.org/10.1016/j.oceaneng.2024.117692, 2024.

Kuo, J.-Y., You, S.-Y., Lin, H.-C., Hsu, C.-Y., and Lei, B.: Constructing Condition Monitoring Model of Wind Turbine Blades, Mathematics, 10, https://doi.org/10.3390/math10060972, 2022.

Lei, Z., Lin, H., Tang, X., Xiong, Y., and Wen, H.: Wind Turbine Blade Fault Detection Method Based on TROA-SVM, Sensors, 25, https://doi.org/10.3390/s25030720, 2025.

Leon-Medina, J. X., Anaya, M., Parés, N., Tibaduiza, D. A., and Pozo, F.: Structural Damage Classification in a Jacket-Type Wind-Turbine Foundation Using Principal Component Analysis and Extreme Gradient Boosting, Sensors, 21, https://doi.org/10.3390/s21082748, 2021.

Li, W., Z., W., W., T., and D., Y.: Surface Defect Detection and Evaluation Method of Large Wind Turbine Blades Based on an Improved Deeplabv3+ Deep Learning Model, Structural Durability & Health Monitoring, 18, 553, https://doi.org/10.32604/sdhm.2024.050751, 2024.

Li, W., Zhao, W., and Du, Y.: Large-scale wind turbine blade operational condition monitoring based on UAV and improved YOLOv5 deep learning model, Mechanical Systems and Signal Processing, 226, 112 386, 2025.

Liu, H., Wang, Y., Zeng, T., Wang, H., Chan, S.-C., and Ran, L.: Wind turbine generator failure analysis and fault diagnosis: A review, IET Renewable Power Generation, 18, 3127–3148, https://doi.org/https://doi.org/10.1049/rpg2.13104, 2024.

Machado, M. and Dutkiewicz, M.: Wind turbine vibration management: An integrated analysis of existing solutions, products, and Open-source developments, Energy Reports, 11, 3756–3791, https://doi.org/https://doi.org/10.1016/j.egyr.2024.03.014, 2024.

Meyer, A.: Vibration Fault Diagnosis in Wind Turbines Based on Automated Feature Learning, Energies, 15, https://doi.org/10.3390/en15041514, 2022.

Milani, S., Leoni, J., Cacciola, S., Croce, A., and Tanelli, M.: A machine-learning-based approach for active monitoring of blade pitch misalignment in wind turbines, Wind Energy Science, 10, 497–510, https://doi.org/10.5194/wes-10-497-2025, 2025.

Morozovska, K., Bragone, F., Svensson, A. X., Shukla, D. A., and Hellstenius, E.: Trade-offs of wind power production: A study on the environmental implications of raw materials mining in the life cycle of wind turbines, Journal of Cleaner Production, 460, 142 578, https://doi.org/https://doi.org/10.1016/j.jclepro.2024.142578, 2024.

Mousavi, Z., Varahram, S., Ettefagh, M. M., Sadeghi, M. H., Feng, W.-Q., and Bayat, M.: A digital twin-based framework for damage detection of a floating wind turbine structure under various loading conditions based on deep learning approach, Ocean Engineering, 292, 116 563, https://doi.org/https://doi.org/10.1016/j.oceaneng.2023.116563, 2024.

Movsessian, A., Cava, D. G., and Tcherniak, D.: Interpretable Machine Learning in Damage Detection Using Shapley Additive Explanations, ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg, 8, 021 101, https://doi.org/10.1115/1.4053304, 2022.

Nguyen, C. U., Huynh, T. C., and Kim, J. T.: Vibration-based damage detection in wind turbine towers using artificial neural networks, Structural Monitoring and Maintenance, 5, 507–519, https://doi.org/10.12989/smm.2018.5.4.507, 2018.

Praveen, H. M., Sabareesh, G., Inturi, V., and Jaikanth, A.: Component level signal segmentation method for multi-component fault detection in a wind turbine gearbox, Measurement, 195, 111 180, https://doi.org/https://doi.org/10.1016/j.measurement.2022.111180, 2022.

Regan, T., Beale, C., and Inalpolat, M.: Wind Turbine Blade Damage Detection Using Supervised Machine Learning Algorithms, Journal of Vibration and Acoustics, 139, 061 010, https://doi.org/10.1115/1.4036951, 2017.

Ren, L. and Yong, B.: Wind Turbines Fault Classification Treatment Method, Symmetry, 14, https://doi.org/10.3390/sym14040688, 2022.

Sharma, S. and Nava, V.: Condition monitoring of mooring systems for Floating Offshore Wind Turbines using Convolutional Neural Network framework coupled with Autoregressive coefficients, Ocean Engineering, 302, 117 650, https://doi.org/https://doi.org/10.1016/j.oceaneng.2024.117650, 2024.

Smarsly, K., Dragos, K., and Wiggenbrock, J.: Machine learning techniques for structural health monitoring, European Workshop on Structural Health Monitoring (EWSHM 2016), 2016.

Song, F., Han, Y., William Heath, A., and Hou, M.: Structural damage detection of floating offshore wind turbine blades based on Conv1d-GRU-MHA network, Engineering Failure Analysis, 166, 108 896, https://doi.org/https://doi.org/10.1016/j.engfailanal.2024.108896, 2024.

Tang, X., Huang, W., Li, X., and Ma, G.: Prediction of mooring tension of floating offshore wind turbines by CNN-LSTM-ATT and Chebyshev polynomials, Ocean Engineering, 331, 121 327, https://doi.org/https://doi.org/10.1016/j.oceaneng.2025.121327, 2025.

Thomas, D.: Unveiling Wind Turbine Failures Causes, Detection, and Prevention for Enhanced Reliability, Journal of Failure Analysis and Prevention, 24, 2051–2053, https://doi.org/10.1007/s11668-024-02026-1, 2024.

Tsiapoki, S., Häckell, M. W., Grießmann, T., and Rolfes, R.: Damage and ice detection on wind turbine rotor blades using a three-tier modular structural health monitoring framework, Structural Health Monitoring, 17, 1289–1312, https://doi.org/10.1177/1475921717732730, 2018.

Veers, P., Bottasso, C. L., Manuel, L., Naughton, J., Pao, L., Paquette, J., Robertson, A., Robinson, M., Ananthan, S., Barlas, T., Bianchini, A., Bredmose, H., Horcas, S. G., Keller, J., Madsen, H. A., Manwell, J., Moriarty, P., Nolet, S., and Rinker, J.: Grand challenges in the design, manufacture, and operation of future wind turbine systems, Wind Energy Science, 8, 1071–1131, 2023.

Vidal, Y., Aquino, G., Pozo, F., and Gutiérrez-Arias, J. E. M.: Structural Health Monitoring for Jacket-Type Offshore Wind Turbines: Experimental Proof of Concept, Sensors, 20, https://doi.org/10.3390/s20071835, 2020.

Vives, J.: Vibration analysis for fault detection in wind turbines using machine learning techniques, Advances in Computational Intelligence, 2, 1–12, https://doi.org/10.1007/s43674-021-00029-1, 2022.

Vives, J., Roses Albert, E., Quiles, E., Palací, J., and Fuster, T.: Vibration Analysis for Fault Detection of Wind Turbines by Combining Machine-Learning Techniques and 3D Scanning Laser, Genetics Research, 2022, https://doi.org/10.1155/2022/2093086, 2022.

Zarrin, P. S., Martin, C., Langendoerfer, P., Wenger, C., and DIaz, M.: Vibration Analysis of a Wind Turbine Gearbox for Off-cloud Health Monitoring through Neuromorphic-computing, IECON Proceedings (Industrial Electronics Conference), 2021-October, https://doi.org/10.1109/IECON48115.2021.9589879, 2021.

Zhang, S., Wang, T., Worden, K., Sun, L., and Cross, E. J.: Canonical-correlation-based fast feature selection for structural health monitoring, Mechanical Systems and Signal Processing, 223, 111 895, https://doi.org/https://doi.org/10.1016/j.ymssp.2024.111895, 2025.

Zhao, W., Li, W., and Du, Y.: Wind turbine blade rotational condition monitoring based on RBs-YOLO deep learning model, Mechanical Systems and Signal Processing, 230, 112 641, https://doi.org/https://doi.org/10.1016/j.ymssp.2025.112641, 2025.

Ángel Encalada-Dávila, Vidal, Y., and Tutivén, C.: Strain virtual sensor for offshore wind turbine jacket supports: A time series transformer approach validated with Alpha Ventus wind farm data, Mechanical Systems and Signal Processing, 231, 112 653, 2025.