https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

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. The k-means algorithm 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. Experimental results demonstrate strong performance in 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 up to 100% classification accuracy, highlighting its effectiveness in diagnosing wind turbine conditions and improving the overall reliability and operational safety 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). Failure mechanisms in wind turbines can be classified into three main categories (Thomas, 2024): mechanical, electrical and environ-

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

mental causes. 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). Current research into fault detection of turbine components employing machine learning (ML) techniques has advanced over the past decade, but it remains an ongoing area of study. The integration of structural health monitoring (SHM) with ML approaches can effectively detect structural defects and operational failures, enabling precise and automated monitoring processes by extracting features from sensor data and classifying them to detect faults.

Machine learning techniques have significantly enhanced the efficiency and accuracy of SHM systems 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 vibration data, identifying patterns and anomalies that may indicate structural deterioration. Unlike traditional methods that rely on hand-crafted features, machine learning enables real-time monitoring and early issue detection by learning from historical and newly acquired data. Beyond vibration analysis (Ciang et al., 2008; Kim et al., 2017; Weijtjens et al., 2017; Nguyen et al., 2017). ML can integrate data from various sources, including spectral, temperature, and moisture sensors, as well as visual inputs from cameras or drones, among others. By analysing these diverse data streams jointly, the model can sort out correlations between different physical effects, for instance, distinguishing temperature-induced variations from actual fault signatures. This multidimensional approach not only enhances anomaly detection but also improves predictive capabilities by capturing the underlying system dynamics. Additionally, ML algorithms continuously refine their performance, adapting to new data and increasing the accuracy and efficiency of SHM systems over time. These advancements contribute to early intervention, reducing the risk of structural failures.

Recent studies have investigated the integration of ML and SHM techniques for system monitoring, mainly by using vibration data to detect faults (Jia and Li, 2023). Wind turbine blade monitoring have been investigated in Antoniadou et al. (2013, 2015) by neural networks and Gaussian processes for pattern recognition and fault discrimination, initially using Frequency Response Function measurements with multilayer perceptrons (MLPs) and radial basis function (RBF) networks, then later focusing on data-driven vibration analysis and SCADA data. Tsiapoki et al. (2018) tested a three-layer monitoring framework on rotor blade data, examining environmental and operational impacts on structural dynamics and using unsupervised clustering to enhance damage detection accuracy. A vibration-based fault classification method that addresses blade cracks, erosion, and connection issues, finding functional trees to be highly effective, was investigated in (Joshuva. and Sugumaran., 2017). Further, the authors tested twelve rule-based classifiers, demonstrating their adaptability to complex fault patterns (Joshuva et al., 2019). Khazaee et al. (2022) proposed a convolutional neural network (CNN) for analysing tower vibrations with high detection accuracy, and Dervilis et al. (2014b, a) introduced neural networks, such as auto-associative Neural Networks (AANN) and RBF networks, which effectively identify structural anomalies with high sensitivity and minimal false positives. Random forest classifier was investigated in (Milani et al., 2025) for active monitoring of blade pitch misalignment in wind turbines, successfully detecting small misalignments.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The turbine support structures, such as towers and foundations, are also vulnerable to damage, necessitating reliable SHM strategies to prevent catastrophic failures. Vidal et al. (2020) proposed a data-driven methodology for diagnosing cracks at specific locations in jacket-type foundations, utilising accelerometer data for vibration responses. Principal component analysis (PCA) was applied for dimensionality reduction and feature extraction, followed by classification using k-nearest neighbors (kNN) and support vector machines (SVM) to detect and localise damage. Similarly, Leon-Medina et al. (2021) developed an SHM method for offshore turbine foundations, incorporating data preprocessing, principal component analysis (PCA), and classification using XGBoost. Tested on a small-scale model with varying structural states, this method effectively distinguished intact from damaged conditions. Ren and Yong (2022) further advanced SHM for turbine towers by using a k-means clustering model optimised with a dynamic weighting algorithm to categorise fault types, highlighting the adaptability of unsupervised clustering methods for classifying fault patterns under complex structural configurations and operational conditions. Artificial neural networks (ANNs) to assess damage in wind turbine towers, utilising modal parameters such as mode shapes and frequencies to identify damaged elements and quantify their severity in a simulated tower structure, were investigated in (Nguyen et al., 2018), and a strategy utilising vibration responses to detect damage proposed in (Hoxha et al., 2020) employing classifiers such as kNN, quadratic SVM, and Gaussian SVM.

Regarding drivetrain components addressed by ML-based SHM strategies. Zarrin et al. (2021) developed a neuromorphic ML model to classify accelerometer data from healthy and damaged gearbox conditions. Praveen et al. (2022) introduced signal segmentation to isolate specific gearbox vibration stages, which was validated using machine learning algorithms, including decision trees, SVM, and deep neural networks. Elforjani (2020) utilised ANN, Gaussian Processes, and SVM for gearbox fault detection, enhancing classification accuracy with feature extraction and PCA. Gao et al. (2021) proposed a CNN solution for detecting mechanical faults in bearings and gearboxes. For bearing fault detection, Vives et al. (Vives, 2022) demonstrated the effectiveness of kNN and SVM, later integrating deep learning for advanced monitoring (Vives et al., 2022). Amin et al. (2023) employed a CNN with cyclostationary features for fault detection, whereas Meyer (2022) investigated unsupervised CNN-based learning for health state classification. Bolt connections, crucial for structural stability but prone to loosening, are monitored using ML. Studies using Gaussian Process Regression (GPR) highlight that high vibration amplitudes increase loosening risk, underscoring SHM's role in maintaining integrity and preventing failures (Mehmanparast et al., 2020).

Despite recent advances, significant challenges remain in adapting SHM techniques to complex operational and environmental conditions, as well as in improving detection accuracy and reliability. While many studies employ supervised or unsupervised learning approaches to enhance anomaly detection or optimise classifiers for specific turbine components, integrating both techniques into a unified SHM framework remains a major research hurdle. In this study, the unsupervised clustering algorithm k-means is employed to label and group data into homogeneous clusters, enabling pattern recognition without the need for predefined labels, and six supervised ML algorithms for binary or multiple classification. The proposed PyMLDA open-source tool (Coelho et al., 2024a) is utilised to support damage assessment based on dynamic sensoring information received from the wind turbine. It combines hybrid machine learning algorithms with signal processing, feature extraction, and sensor selection to enhance SHM applications. The model's flexibility in accepting diverse input features reinforces SHM decision-making by facilitating the identification, classification, and quantification of damage levels. Therefore, the novelty of

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

this work lies in two proposed contributions: (i) a feature relative change strategy for data normalisation, and (ii) a feature and sensor selection process based on canonical correlation analysis. These enhancements improve failure classification accuracy and allow for the inclusion of multiphysics information within the dataset. The integration of these hybrid methods significantly strengthens diagnostic capabilities, increasing both the efficiency and precision of structural condition monitoring. The results demonstrated the model's high performance in both binary and multiclass classification of operational failures of wind turbines, confirming its robustness and effectiveness in real SHM scenarios.

2 Method description

The proposed framework is based on the PyMLDA open-code [1](Coelho et al., 2024a), which aims to predict damage assessment based on numerical or experimental data obtained from the dynamic response of systems. The proposed monitoring model consists of eight steps, as illustrated in Fig. 1, including: (1) receiving the acquired data; (2) data processing; (3) feature extraction and normalisation; (4) unsupervised pattern recognition; feature labelling and clustering; (5) feature and sensor selection; (6) data splitting;, (7) operational classification; and (8) model evaluation. The final step also outputs the estimated damage state(operational failure) and identifies the best-performing ML algorithm based on its classification accuracy. The main novelty introduced in this study, and implemented in the improved PyMLDA framework, lies in the feature and sensor selection, normalisation procedures, and the integration of unsupervised classification and clustering techniques. The Algorithm 1 pipeline can be stated as follows:

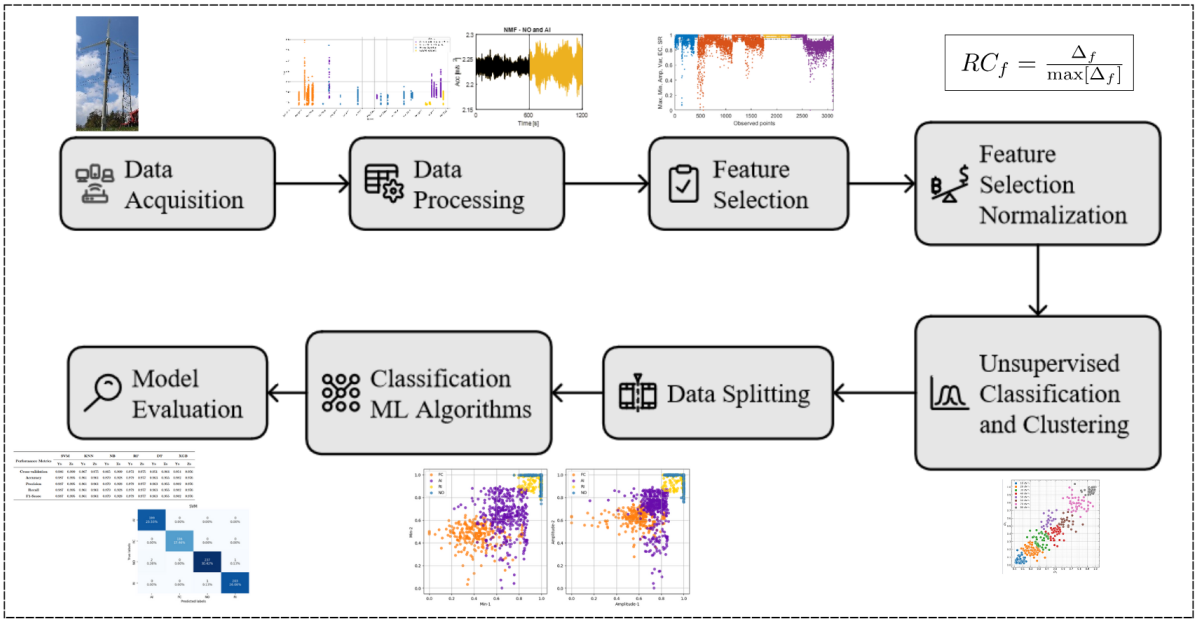


Figure 1. Pipeline representation of a machine learning model for fault classification on the Aventa 6.7 kW wind turbine.

[1]https://github.com/mromarcela/wedowind-challenge-ASCE-EMI

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Algorithm 1 PyMLDA for wind turbine failure classification

1: **Data Acquisition:** Receive structural from the time-domain responses.

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

3: **Feature extraction and normalisation:** Fifteen techniques are applied to extract features from the time-domain signal, see section 2.3.

4: **Pattern Recognition 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

6: **Data Splitting:** Split the dataset into training and testing as listed in Table 3, respectively.

7: **Classification ML algorithms:** Applied ML algorithms for detection, such as SVM, K-NN, RF, NB, DT and XGBoost.

8: **Model Evaluation:** Calculate the Performance for the ML classifiers (Cross-validation, Accuracy, Precision, Recall, and F1-score).

Final decision: Information about damage state based on classification and regression algorithm outcomes.

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. 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 10 Hz. To ensure data quality and enhance the reliability of anomaly detection, a comprehensive data preprocessing pipeline was implemented. This included a supervised sensor analysis, outlier removal, signal segmentation, detrending, feature extraction, and feature normalisation.

The data preprocessing stage begins by analysing the sensors' physics and evaluating the contribution of each one 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. For this study, the x-axis (side-to-side turbine motion) and y-axis (fore-aft turbine motion) signals were included in the monitoring to capture the most information about the damage dynamic behaviour. Figure 2 illustrates the sensor locations within the nacelle, where three accelerometers, GEN_ACC (orange), NMF_ACC (blue), and MSB_ACC (yellow), record signals under different operational conditions, including normal operation (NO). Figure 2 also presents the raw spectral data available in

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

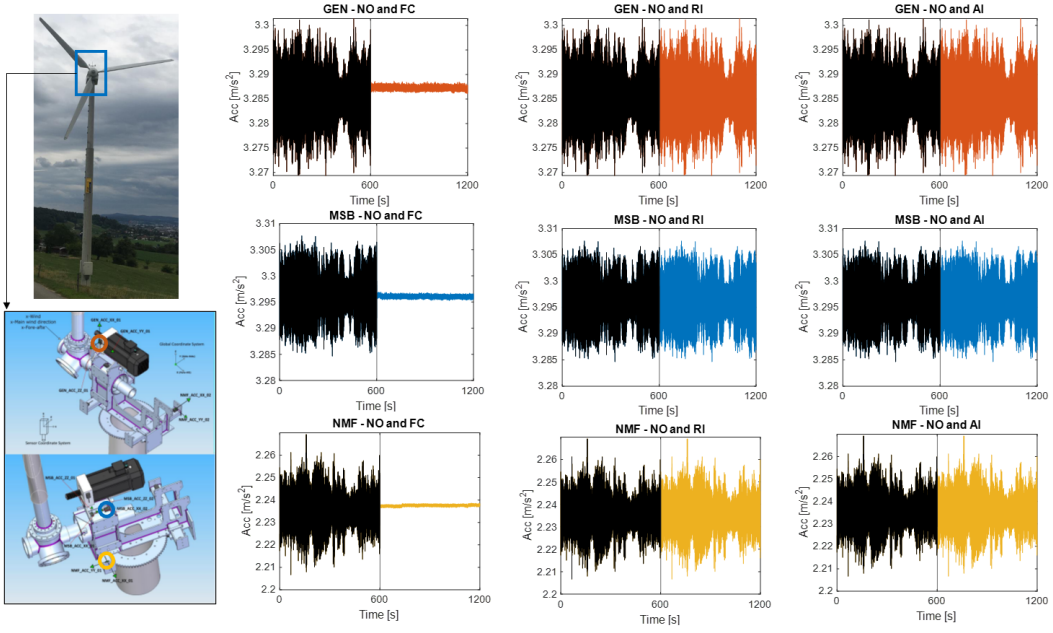


Figure 2. Schematic representation of the turbine and sensors position in the nacelle. Temporal section comparison of the turbine's normal operation and operation with failures: (Top) sensor GEN NO-FC, NO-RI, and NO-AI placed from left to right, respectively. (Middle) sensor MSB NO-FC, NO-RI, and NO-AI; and (Bottom) sensor NMF NO-FC, NO-RI, and NO-AI.

Table 1. Sensor channel and corresponding location of accelerometers and SACDA adopted in the analysis.

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 ACC_Y_01	GEN_Z_01	
SCADA	Temperatute	Wind speed	-	

the dataset. In the acceleration graphs, black lines indicate normal turbine operation, whereas coloured lines 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
135 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 vibration sensors placed

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

at the tower top, nacelle, main shaft bearing, and generator were used in the model, based on their sensitivity to structural, mechanical, and aerodynamic failures, as well as vibration. Wind speed and temperature are from the SCADA system. Table 1 lists the adopted sensors, channel names and SCADA information.

2.2 Dataset organisation

During data analysis, duplicate files were identified, prompting a data redistribution strategy to eliminate redundancies and streamline the process. The data was categorised and colour-coded according to 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.

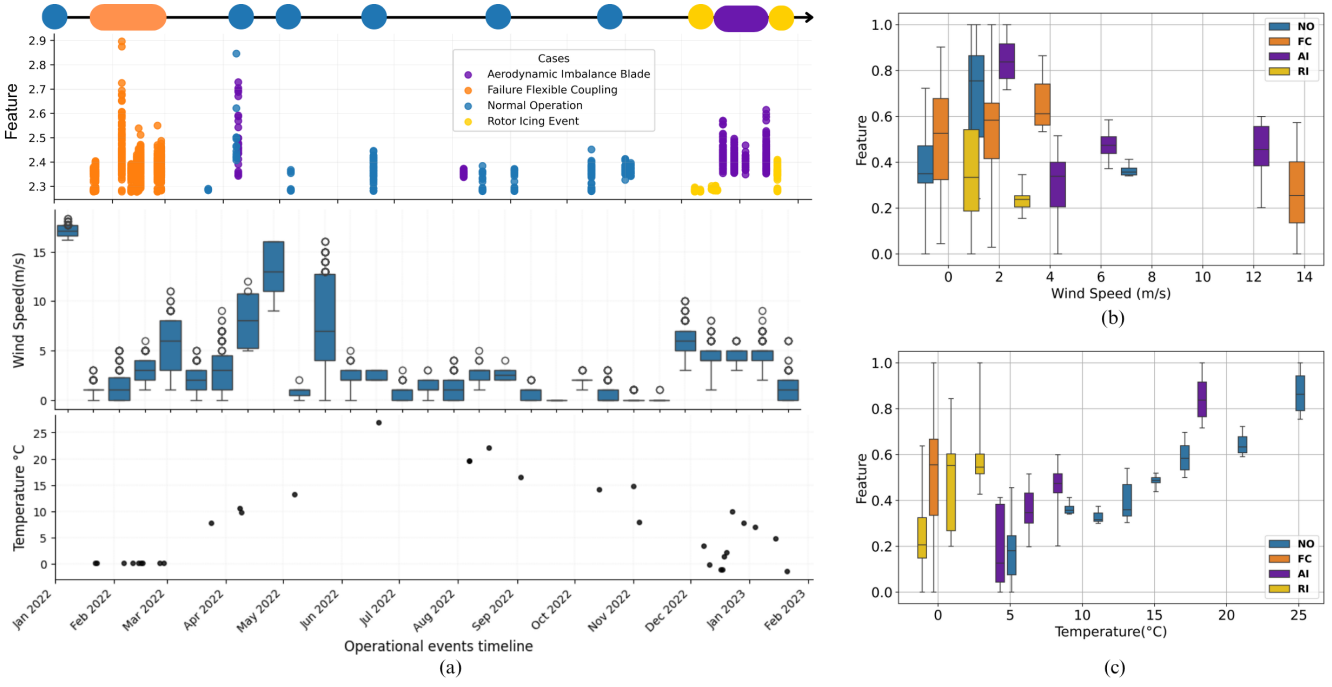


Figure 3. The feature dataset volume over the monitoring timeline categorised by yellow indicates RI, blue denotes NO, purple corresponds to AI, and orange represents the FC condition (a-top). Wind speed boxplot (a-middle), and mean temperature (a-bottom). Boxplot of F05 sensor feature by wind speed 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 response are used in the 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 tempo-

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

ral 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 maximum value of sensor L05-1 (Fig. 9e), 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.

2.3 Feature extraction and normalisation

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), as presented in Table 2, where x is the spectrum signal, and p is the sampling points value.

Table 2. Spectral and statistical formulation used for feature extraction.

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

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

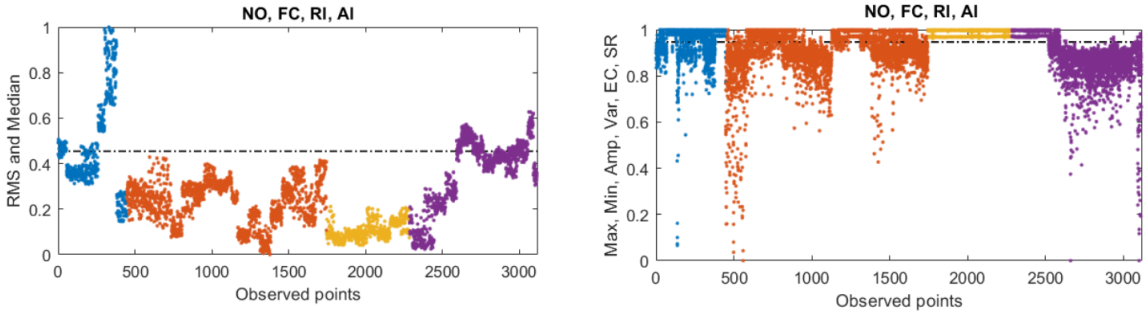


Figure 4. Features extracted groups combining the four turbine operation conditions. The black dashed-dot line is the threshold of the NO condition. (RHS) RMS and median, and (LHS) Max, Min, Amp, Var, EC, and SR.

The formulation used to extract the feature includes maximum, minimum, amplitude range, median, mean, variance, energy, energy centre, skewness, kurtosis, higher-order moments, Shannon entropy, signal rate, and RMS. These features are organised into a dataset encompassing all four operational states. However, some features, such as energy, kurtosis, higher-order moments, and Shannon entropy, exhibited low sensitivity to failures and were therefore discarded during the subsequent normalisation and analysis stages. Despite this refinement, the remaining features still present a challenge because their amplitude variation is minimal, with differences of less than three digits. This small variation poses a significant obstacle for accurate ML classification. To address this issue, the proposed damage index integrates a normalisation step based on the relative change of each feature, defined as

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

where $\Delta_f = (\max(\mathrm{X}) - \mathrm{X})$ represents the difference between each element in the feature's vector, $\mathrm{X} = \{\mathrm{Max}, \mathrm{RMS}, ..., n\}$, and RC is the feature's relative change, calculated by dividing Δ_f by its maximum value. By applying the normalised damage index related to the relative change proposed in Eq. (1), one can scale the Δf values without losing the intrinsic dynamic behaviour for time and normalise the features between unity and zero, preserving their essential characteristics imposed by the operational condition while enabling consistent feature comparison. The features most sensitive to failure include the maximum (max) and minimum (min) values, amplitude range, root mean square (RMS), variance, energy centre, and signal rate. These eight normalised features are grouped by similarity, as illustrated in Fig. 4. From these groups, four features, RMS, max, variance, and amplitude range, are selected to create the global dataset and serve as input for the unsupervised k-means clustering algorithm.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

3 Hybrid machine learning model

3.1 Pattern Recognition and clustering

The next steps in the monitoring process are unsupervised pattern recognition, labelling and clustering of the selected normalised features using the k-means algorithm, with the elbow method used to determine the optimal number of clusters. For validation, the elbow method was confirmed using $\kappa = 2$ for binary classification and $\kappa = 4$ for multi-classification, where κ represents the number of classes to identify clusters within unlabelled data. 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 56.20%, test 25%, and validation 18.80%, which were used for model construction and evaluation, as shown in Table 3. 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.

Table 3. Explicit description of the datasets used for training, validation, and testing.

Classification	Cases	Training	Test	Validation
Binary	NO and FC	981	436	327
	NO and RI	705	314	236
	NO and AI	524	234	175
Multiclass	NO, FC, RI and AI	1772	788	591

3.2 Feature and sensor fast selection

The supervised features analysis indicates that changes in the turbine's dynamic behaviour are related to specific events. For subsequent procedures, a feature and sensor selection is performed using the Canonical-correlation-based selection method introduced by Zhang et al. (2025). In this work, feature and sensor selection is a two-step procedure, beginning with the selection of the best feature(s) of each sensor, followed by the indication of the best sensor(s). The pseudocode for this process is given in Algorithm 2, which is repeated for each fault diagnosis.

The feature matrix is defined as $\mathbf{X} \in \mathbb{R}^{N\times n}$, comprising n features over N instances, and the target matrix as $\mathbf{Y} \in \mathbb{R}^{N\times m}$, consisting of m label variables, defined by the k-means. The objective is to identify a subset of $\mathcal{T}$ informative features from $\mathbf{X}$ that contribute most significantly to modelling $\mathbf{Y}$. The selection process is guided by the sum of squared canonical correlation coefficients(SSC), which serves as the ranking criterion by quantifying the shared variance (i.e., coefficient of determination $R^2(\mathbf{X},\mathbf{y})$) between the features and each target vector $\mathbf{y}$. At each iteration $i \in 0, 1, ..., \mathcal{T}-1$, a search is executed to select the $(i+1)$th most relevant feature $\mathbf{x}_d \in \mathbb{R}^{N\times 1}$, conditioned on the previously selected set $\mathbf{X}_s \in \mathbb{R}^{N\times i}$. This iterative process continues until the best $\mathcal{T}$ features are selected, ensuring maximum explanatory power to the target response, where

$$d = \arg\max_{j} = \sum_{k=1}^{(i+1)\wedge m} R_k^2((\mathbf{X}_s, \mathbf{x}_j)|\mathbf{Y}) \tag{2}$$

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

for the canonical correlation coefficients based on the η-cosine in the SSC is given by

$$\sum_{k=1}^{n\wedge m} R_k^2(\mathbf{X},\mathbf{Y}) = \sum_{k=1}^{n\wedge m} \cos^2(\Theta_k(\mathbf{W},\mathbf{V})) = \sum_{i=1}^{n}\sum_{j=1}^{m} \cos^2(\angle(\mathbf{w}_i,\mathbf{v}_i)) \tag{3}$$

for $\mathbf{W}$ and $\mathbf{V}$ whose columns form an orthogonal basis, and $\Theta_k(\mathbf{W},\mathbf{V})$ the angle between both bases (for further details, please see (Zhang et al., 2025)). 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 $\mathbf{X}_s$ contain the highest ranking feature(s) of each sensor.

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 ∈ S Number of sensor
$\mathrm{X}_s \leftarrow 0$ Candidate feature matrix coating the best feature(s) of each sensor. The threshold is Score ≥ 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 η-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
  $\mathrm{X}_s \leftarrow \mathbf{X}[:, \mathbf{max(d)}]$ **Highest score feature**
end for
procedure FASTCAN(X_s, Y, s, tol, alg)
  $\mathrm{q} \leftarrow \arg\max_j R_j^2$
return q

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

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

decision-making. The proposed model follows a hybrid unsupervised-supervised ML-based framework, incorporating classification strategies to improve structured data attributes. The data-driven ML models are fed with selected features extracted from the multiphysics database, incorporating vibration responses and environmental data, and the output corresponds to the wind turbine's classification operational state.

For the classification task, six ML algorithms are utilised. The unsupervised k-means algorithm is used for initial clustering. At the same time, the supervised classifiers naive bayes (NB), 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 the final classification based on the initial clustering results, generating outputs that include confusion matrices and performance metrics. The selection of hyperparameters follows the recommendations in (de Sousa et al., 2023; Coelho et al., 2024b), 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 1×10^{-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, various sensor data are available from the acquisition system. 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 scores are used to rank the inputs, 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. Features with scores above 0.6 are appointed to build the dataset used in the ML algorithms. Hence, the dataset comprises the high-score features of 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. For binary classification between normal operation and flexible coupling failure, the feature matrix is composed of $X_{(NO\text{-}FC)} = \{RMS\_Y01^{MSH}, Max\_Y01^{L05}, Amp\_Z01^{GEN}, RMS\_X02^{NMF}\}$ identified in the selection step. The individual feature scores for each accelerom-

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

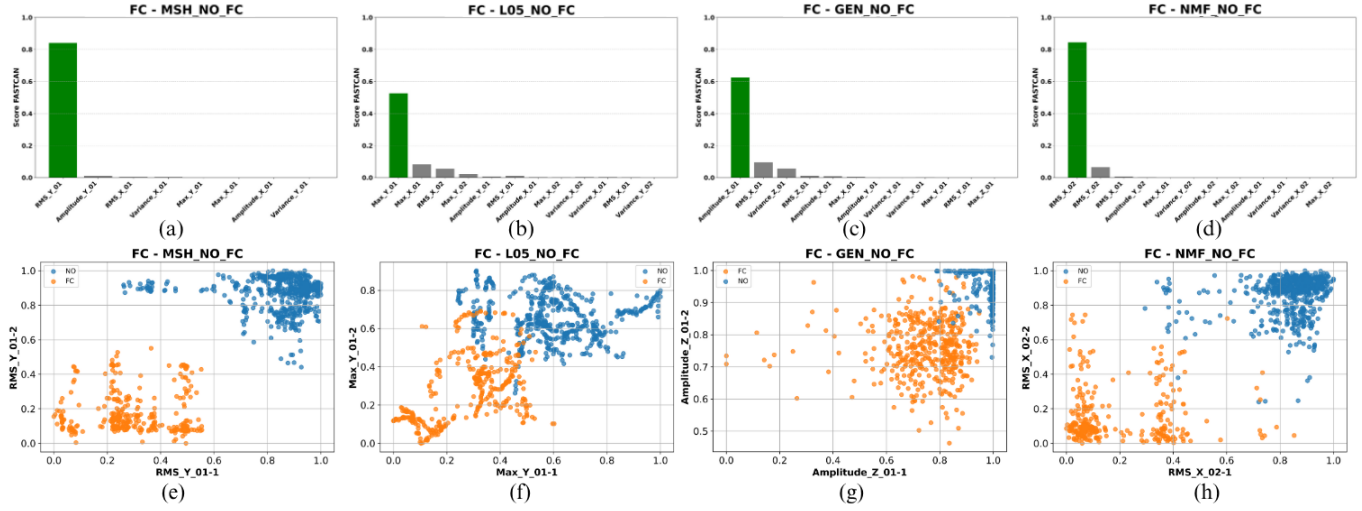


Figure 5. Feature scores for the accelerometer in the x-y-z directions and the corresponding dispersion diagrams for the highest-scoring features: (a,e) MSH; (b,f) L05 (sensors 01 and 02); (c,g) GEN; and (d,h) NMF (sensors 01 and 02).

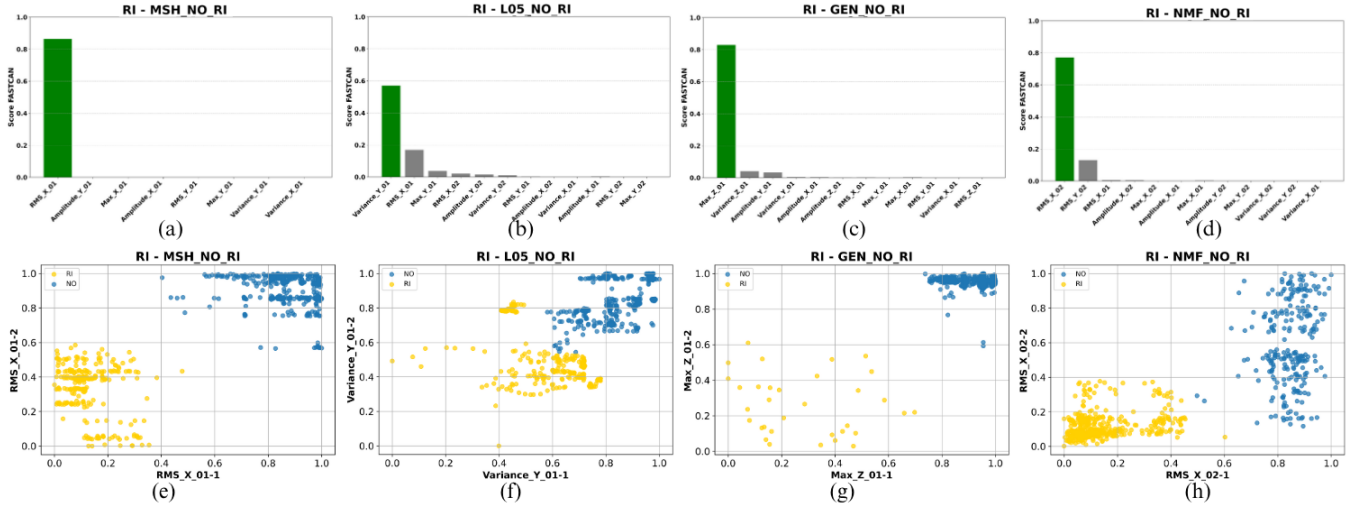


Figure 6. Feature scores for the accelerometer in the x-y-z directions and the corresponding dispersion diagrams for the highest-scoring features: (a,e) MSH; (b,f) L05 (sensors 01 and 02); (c,g) GEN; and (d,h) NMF (sensors 01 and 02).

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. Although each

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

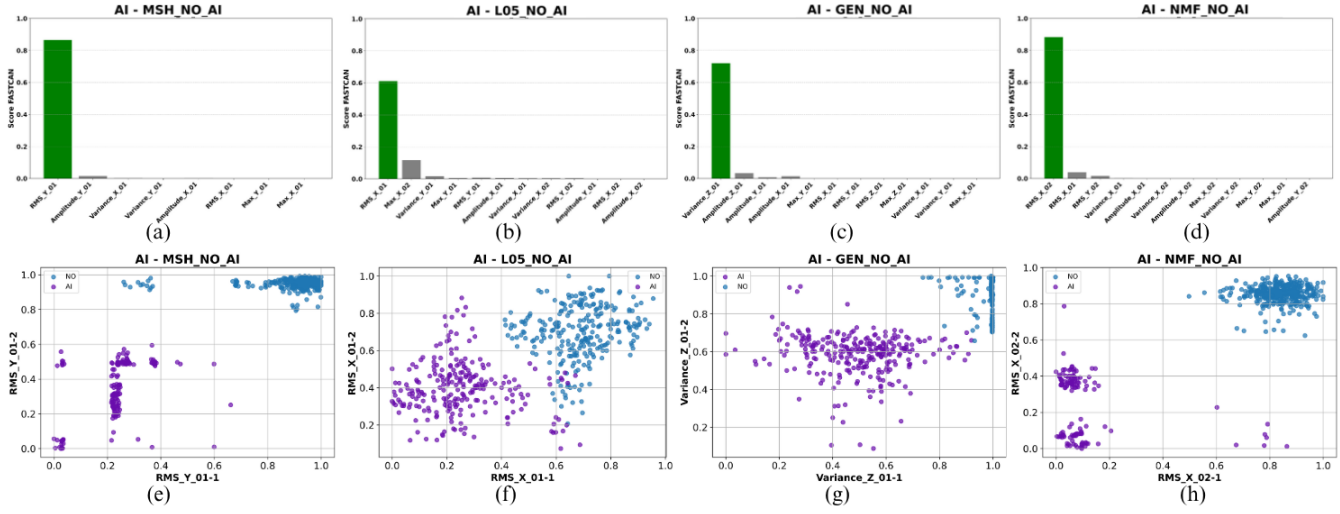


Figure 7. Feature scores for the accelerometer in the x-y-z directions and the corresponding dispersion diagrams for the highest-scoring features: (a,e) MSH; (b,f) L05 (sensors 01 and 02); (c,g) GEN; and (d,h) NMF (sensors 01 and 02).

feature is visualised individually for clarity, the final classification integrates all 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 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 and RI, the feature matrix is defined as $X_{(NO\text{-}RI)} = \{RMS\_X01^{MSH}, Var\_Y01^{L05}, Max\_Z01^{GEN}, RMS\_X02^{NMF}\}$. Feature scores for each accelerometer are shown in Fig. 6(a-d), with the GEN sensor (Max_Z01^{GEN}) 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.

The feature matrix for the NO and AI is $X_{(NO\text{-}AI)} = \{ RMS\_Y01^{MSH}, RMS\_X01^{L05}, Var\_Z01^{GEN}, RMS\_X02^{NMF}\}$, where the feature ranking scores of each accelerometer are presented in Fig. 7(a-d), where the GEN sensor (Max_Z01^{GEN}) 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 RI. These diagrams illustrate

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

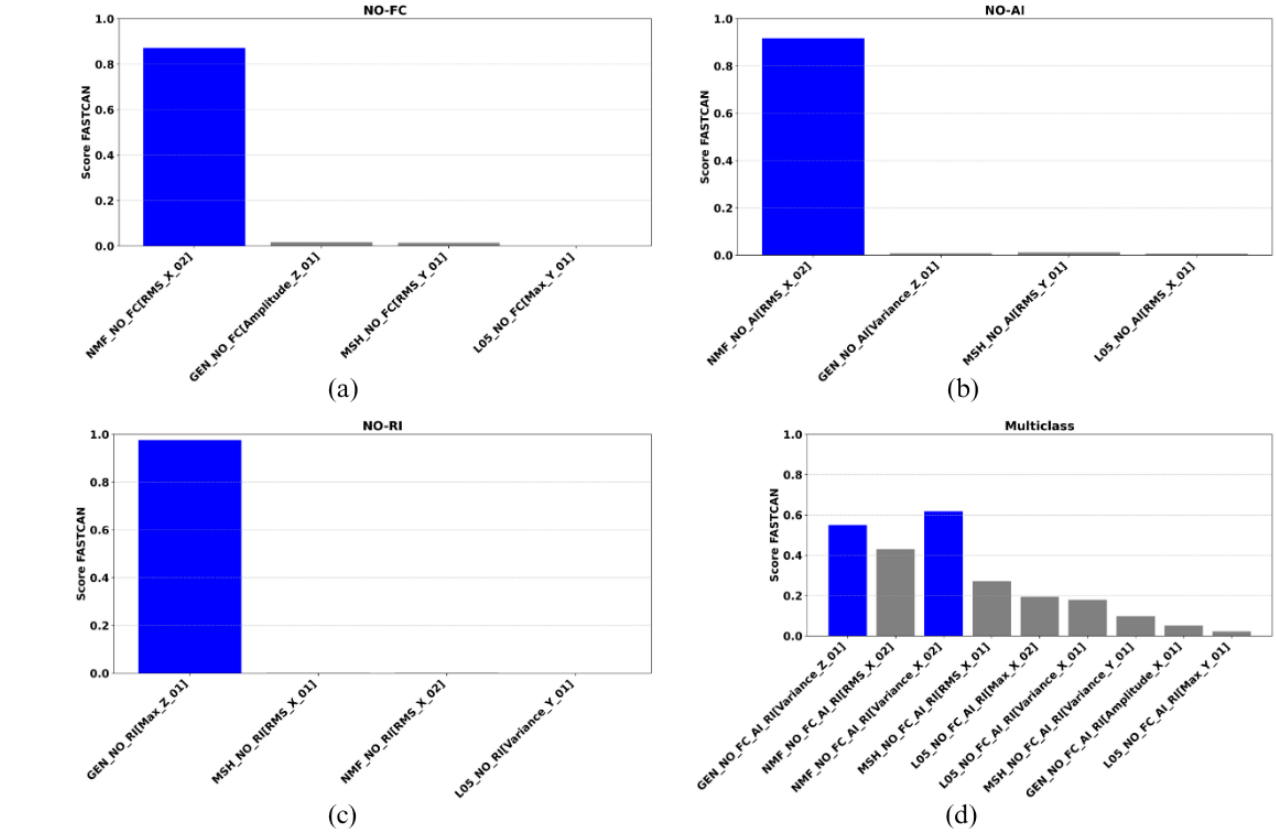


Figure 8. Sensor score based on the top-performing features of each operation condition: (a) NO-FC, (b) NO-AI, (c) NO-RI, and (d) multi.

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 and 10 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.

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 scores higher than 0.5 are included in the global dataset for the classification task, where the feature matrix is defined as $X_{multi} = \{RMS\_X01^{MSH}, Var\_Y01^{MSH}, Var\_X01^{L05}, Max\_X02^{L05}, Max\_Y01^{L05}, Amp\_X01^{GEN}, Var\_Z01^{GEN}, RMS\_X02^{NMF}, Var\_X02^{NMF}\}$. The GEN and MNF sensors show the highest relevance, indicating strong sensitivity, as shown in Fig. 8(d).

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

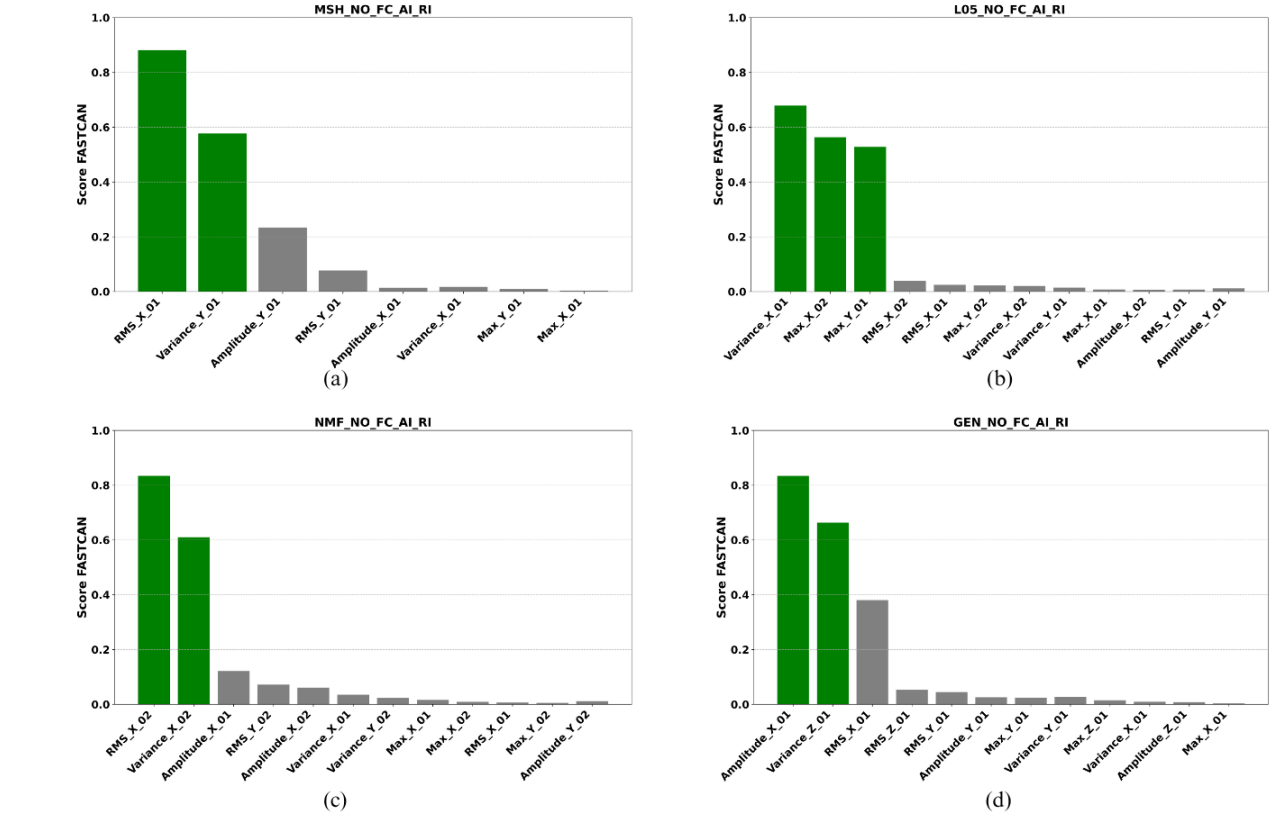


Figure 9. Feature selection based on the top-performing features of each sensor: (a) MSH, (b) L05, (c) MNF, and (d) GEN.

Each selected feature is visualised through its dispersion correlation in Fig. 10(a-c), corresponding to results from sensors GEN, MSH, and NMF, respectively. Figures 10(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 classification machine learning algorithms, including kNN, SVM, DT, RF, Naive Bayes, and XGB. These algorithms perform the final classification based on the input dataset, providing outputs as confusion matrices and

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

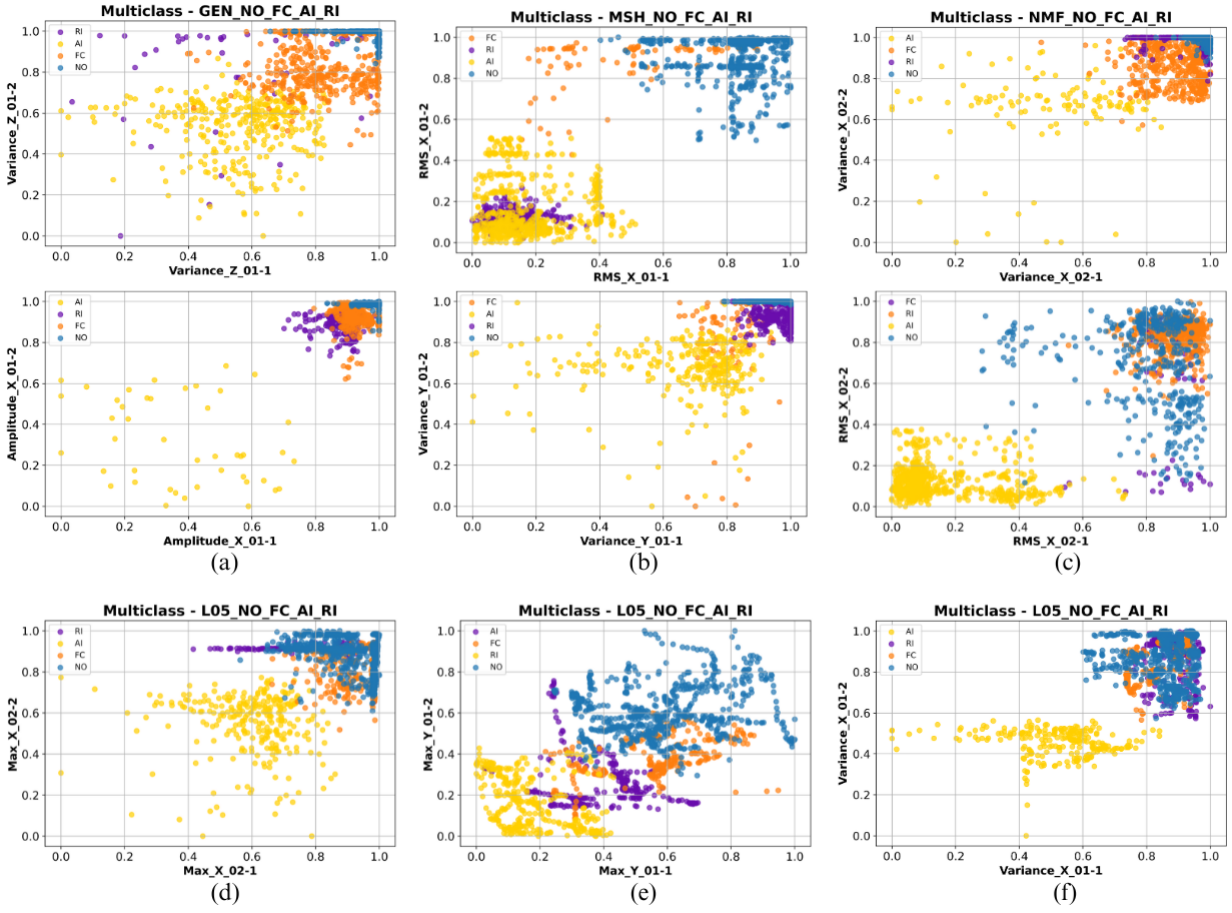


Figure 10. Correlation plots of multiple conditions failure (NO, FC, RI, and AI) for the highest scored features of each sensor: (a) GEN, (b) MSH, (c) NMF, and (d-f) L05.

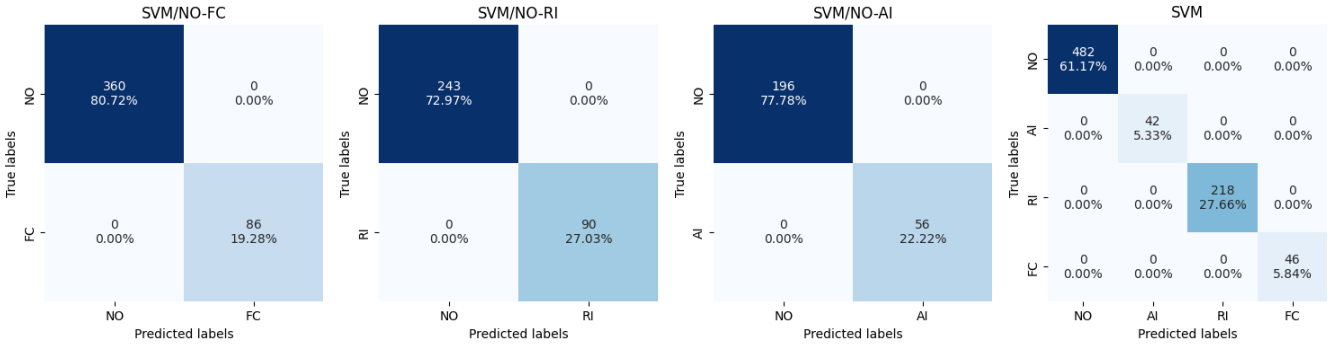
performance metrics. The selection of hyperparameters is based on the investigations of (de Sousa et al., 2023), which specify optimal configurations for this assignment and are briefly described in Section 3.3.

Table 4 presents the performance of ML models in binary classification (NO-FC, NO-RI, and NO-AI) and multiple failure conditions evaluated through 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. Micro-averaging computes classification metrics by globally summing true positives, false positives, and false negatives across all classes, resulting in similar Precision, Recall, and F1-score. Macro-averaging, in turn, calculates these metrics for each class separately and then averages them, treating all classes equally. Weighted averaging follows the same approach as macro-averaging but adjusts for class frequencies, ensuring that more frequent classes contribute proportionally to the final score.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

Table 4. Performance metrics of the ML algorithms SVM, kNN, NB, RF, DT, and XGB for binary and multiple classification

Metrics Perfomance		SVM / kNN RF / DT	NB	XGB	Metrics Perfomance		SVM / kNN RF / DT	NB	XGB
Cross validation	**NO-FC**	100%	95,60%	100%	**F1-Score**	**NO-FC**	100%	96,39%	100%
	NO-RI	100%	94,24%	99,86%		**NO-RI**	100%	95,06%	100%
	NO-AI	100%	99,65%	100%		**NO-AI**	100%	100%	100%
	Multiclass	100%	97,12%	100%		**Multiclass**	100%	96,92%	100%
Accuracy	**NO-FC**	100%	94,17%	100%	**Recall**	**NO-FC**	100%	88,39%	100%
	NO-RI	100%	92,79%	100%		**NO-RI**	100%	89,47%	100%
	NO-AI	100%	100%	100%		**NO-AI**	100%	100%	100%
	Multiclass	100%	95,94%	100%		**Multiclass**	100%	91,90%	100%
Precision	**NO-FC**	100%	91,56%	100%					
	NO-RI	100%	91,52%	100%					
	NO-AI	100%	100%	100%					
	Multiclass	100%	94,06%	100%					



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

Figure 11. 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.

Among the ML methods used for classification of turbine state operation, the SVM algorithm achieves excellent metrics, reaching a value of 1 across all evaluated measures (Accuracy, Precision, Recall, and F1-score) in the three binary classification scenarios. The performance of SVM is especially remarkable in NO-RI, where it attains peak scores. Although kNN, NB, DT, RF, and XGB also show strong metric results, with less pronounced variations, SVM stands out, offering greater interpretation and precision. This analysis highlights the crucial importance of considering feature and sensor selection in dataset construction for training classification models. Confusion matrices are generated for all algorithms, but based on the results, the SVM

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

is selected as the model for detailed analysis across the three scenarios. This selection allows for a focused presentation, particularly on the educational aspects of the findings, given the extensive number of graphical results obtained.

The performance metrics of the SVM, kNN, NB, RF, DT, and XGB algorithms for a multiclass classification are presented in Table 4. The evaluation is based on cross-validation, accuracy, precision, recall, and F1-score, providing a comprehensive assessment of the effectiveness of each algorithm. SVM stood out as the best-performing model overall, achieving superior results. kNN, XGB, and RF also showed consistent performance. In contrast, the NB model demonstrated lower performance metrics. Based on SVM's exceptional results, it will be used for further analysis, including the evaluation of confusion matrices for the features of selected sensors, as well as SCADA, where it achieved the highest performance.

The confusion matrices shown in Figs. 11(a–c) present the results of binary classification using the SVM algorithm. In the NO–FC case, the model achieved high accuracy, correctly classifying 360 samples as NO (80.72%) and 86 as FC (19.28%). For the NO–RI classification, the performance remained strong, with 243 correct predictions for NO (72.97%) and 90 for RI (27.03%), and notably, no false positives were observed. In the NO–AI scenario, the model correctly classified 196 samples as AI (77.78%) and 56 as NO (22.22%), indicating a slight performance decrease, which may be attributed to the increased complexity or class imbalance in the AI dataset. Figure 11(d) displays the confusion matrix for the multiclass classification involving AI, FC, NO, and RI. The model correctly identified 42 samples as AI (5.33%), 46 as FC (5.84%), 218 as RI (27.66%), and 482 as NO (61.17%). Notably, no false positives were detected for the NO class, highlighting the model's high precision in identifying normal operating conditions.

5 Conclusions

This work proposes a framework for classifying operational conditions (normal operation, pitch drive faults, rotor icing, and aerodynamic imbalance) of the wind turbine using experimental data. The monitoring follows an eight-step process, including data processing, feature and sensor selection, normalisation, data splitting, unsupervised clustering, machine learning classification, and model evaluation. A novel feature normalisation strategy was introduced to enhance the differentiation of dynamic responses. Among the fifteen extracted features, eight were identified as the most sensitive and ranked using canonical correlation analysis. In addition to vibration signals acquired from six accelerometers positioned throughout the turbine, environmental parameters such as temperature and wind speed were incorporated into the monitoring process. The operational conditions were labelled and grouped using the unsupervised k-means algorithm to uncover distinct clusters in the data. For binary classification tasks, the SVM consistently achieved perfect performance metrics (accuracy = 1.0), demonstrating exceptional capability in distinguishing between normal and faulty conditions across various sensor orientations. Confusion matrices confirmed the model's reliability, showing minimal misclassifications and no false positives. While other algorithms, including kNN, NB, and XGB, also performed well, SVM stood out for its consistency and precision. In multiclass classification involving multiple operational conditions, SVM again delivered the best results, achieving high accuracy and strong class discrimination. Overall, SVM proved to be the most robust and effective model for both binary and multiclass fault detection. These results highlight the potential of the proposed framework, implemented as PyMLDA open-code, for structural health monitoring of complex systems, which demonstrated outstanding performance in wind turbine condition diagnostics.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

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).

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

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, 2015.

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.

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.

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

Ciang, C. C., Lee, J.-R., and Bang, H.-J.: Structural health monitoring for a wind turbine system: a review of damage detection methods, Measurement Science and Technology, 19, 122 001, https://doi.org/10.1088/0957-0233/19/12/122001, 2008.

Coelho, J., Machado, M. R., and de Souza, A. A. S.: Python open-source code for Machine Learning Damage Assessment , Software Impacts, 19, 100 628, https://doi.org/10.1016/j.simpa.2024.100628, 2024a.

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, 2024b.

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.

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/doi.org/10.1117/12.849189, 2010.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

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 – 1113, 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.

Jia, J. and Li, Y.: Deep Learning for Structural Health Monitoring: Data, Algorithms, Applications, Challenges, and Trends, Sensors, 23, 8824, https://doi.org/https://doi.org/10.3390/s23218824, 2023.

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.

Khazaee, M., Derian, P., and Mouraud, A.: A comprehensive study on Structural Health Monitoring (SHM) of wind turbine blades by instrumenting tower using machine learning methods, Renewable Energy, 199, 1568–1579, https://doi.org/https://doi.org/10.1016/j.renene.2022.09.032, 2022.

Kim, W., Yi, J.-H., Kim, J.-T., and Park, J.-H.: Vibration-based Structural Health Assessment of a Wind Turbine Tower Using a Wind Turbine Model, Procedia Engineering, 188, 333–339, https://doi.org/https://doi.org/10.1016/j.proeng.2017.04.492, structural Health Monitoring - From Sensing to Diagnosis and Prognosis, 2017.

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

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.

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.

Mehmanparast, A., Lotfian, S., and Vipin, S. P.: A Review of Challenges and Opportunities Associated with Bolted Flange Connections in the Offshore Wind Industry, Metals, 10, https://doi.org/10.3390/met10060732, 2020.

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.

https://doi.org/10.5194/wes-2025-131
Preprint. Discussion started: 7 August 2025

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.

Nguyen, T. C., Huynh, T. C., Yi, J. H., and Kim, J. T.: Hybrid bolt-loosening detection in wind turbine tower structures by vibration and impedance responses, Wind and Structures, An International Journal, 24, 385–403, https://doi.org/10.12989/was.2017.24.4.385, 2017.

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.

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

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

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, https://doi.org/10.5194/wes-8-1071-2023, 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.

Weijtjens, W., Verbelen, T., Capello, E., and Devriendt, C.: Vibration-based structural health monitoring of the substructures of five offshore wind turbines, Procedia Engineering, 199, 2294–2299, https://doi.org/https://doi.org/10.1016/j.proeng.2017.09.187, x International Conference on Structural Dynamics, EURODYN 2017, 2017.

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.