Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2025) 4866-4874

ECCOMAS
Proceedia

COMPDYN 2025
10th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Rhodes Island, Greece, 15-18 June 2025

SEMI-SUPERVISED DAMAGE DETECTION IN WIND TURBINE COMPONENTS UNDER ENVIRONMENTAL VARIABILITY

Maria E. Quiroz[1], Sathiskumar A. Ponnusami[1], Miguel A. Bravo-Haro[1]

[1] Department of Engineering, City St George's, University of London, London, UK.
e-mail: {maria.quiroz-roa,sathiskumar.ponnusami, miguel.bravo-haro}@city.ac.uk

Abstract

The transition to green energy requires ensuring the structural integrity of renewable energy systems throughout their lifespan. This necessitates continuous monitoring and effective maintenance strategies, where efficient damage detection plays a critical role in reducing operational costs and preventing failures.
This study presents a semi-supervised, data-driven damage detection algorithm tailored for structural health monitoring (SHM) of wind turbines. The method utilises a multivariate autoregressive model for feature extraction, principal component analysis (PCA) for data normalization, and Mahalanobis distance as a damage index. The model is trained exclusively on baseline conditions and employs K-means clustering to automate structural condition recognition. The approach is validated using two case studies: a numerical model of a monopile-supported offshore wind turbine using synthetic data and a turbine blade under varying environmental and structural conditions using experimental data. Its simplicity, transparency, and robustness make it a reliable tool for SHM and condition monitoring in complex and demanding environments.

Keywords: Structural health monitoring, damage detection, offshore wind turbines, data-driven methods, unsupervised learning.

ISSN:2623-3347
Peer-review under responsibility of the organizing committee of COMPDYN 2025.
doi: 10.7712/120125.12779.24861

1 INTRODUCTION

Structural Health Monitoring (SHM) has become a crucial tool for ensuring the reliability and safety of engineering structures across various industries, including civil, aerospace, and mechanical systems. By continuously or periodically collecting data, SHM enables the early detection of structural degradation, reducing the risk of catastrophic failures and minimising maintenance costs [1]. However, accurately identifying structural damage remains challenging due to the influence of environmental and operational (E&O) conditions, which can obscure the effects of structural changes. To address this, normalisation techniques such as regression analysis, support vector machines, artificial neural networks, and principal component analysis (PCA) have been widely employed to separate damage-related changes from environmental variations [2].

Data-driven approaches have gained significant attention due to their strong performance in SHM [3]. These methods are based on an architecture whose main components are (i) feature extraction, (ii) data normalisation and (iii) statistical classification. Semi-supervised algorithms that are trained solely on baseline data are preferable for large and complex systems where damaged data is scarce and difficult to obtain. Modal-based properties are widely used for feature extraction, such as autoregressive models and wavelet components [4]. For instance, [5] combines autoregressive models with PCA and Mahalanobis distance calculations, validated on a railway bridge. This method is further extended to online monitoring in [6]. In [7], Multiple Models with autoregressive features combined with PCA are proposed for considering uncertainty and applied to composite beams, while [8] explores the use of correlation-based features with PCA and Mahalanobis distance for evaluating an inland wind turbine blade excited by an active shaker.

Despite these advancements, SHM applications in offshore wind turbines (OWTs) remain limited. The offshore wind industry has immense potential for large-scale clean energy production, but as turbine sizes increase to maximize energy output, harsh marine environments introduce significant maintenance challenges and elevate operational costs [9]. The combination of turbulent wind conditions, dynamic loads, and complex control systems makes SHM particularly difficult for OWTs. To ensure practical applicability, online monitoring procedures must be robust to these conditions.

This study proposes a multivariate autoregressive model for feature extraction, combined with PCA for feature normalisation and Mahalanobis distance for feature, comparing new observations to the baseline condition. To eliminate human intervention in damage recognition, the fused features are then clustered using a K-means algorithm, as in [6]. The methodology is evaluated in two case studies: a numerical model of an offshore wind turbine structure and an experimentally tested wind turbine blade benchmark.

2 DAMAGE DETECTION ALGORITHM

The damage detection algorithm consists of four main steps following the collection of acceleration data from sensors installed along the structural element under monitoring: (i) feature extraction, (ii) feature modeling, (iii) feature fusion, and (iv) feature discrimination and clustering. The proposed algorithm builds upon the methods introduced in [5, 6] extending them to account for the interdependency of measured signals by employing multivariate autoregressive models (MAR) for feature extraction instead of their univariate counterparts. This approach leverages the correlation between multiple signals (i.e., multiple sensors in the network of the installed monitoring system) to describe the behavior of the system, detecting damage through variations in this correlation rather than changes in individual signal behavior.

The MAR predicts the next value in a D-dimensional time series, y_i, as a linear combination of the N previous vector values. The model is expressed as:

$$y_i = \sum_{j=1}^{N} y_{i-j} A_j + e_i \tag{1}$$

where D represents the number of sensors, N is the MAR model order, $y_i = \left[y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(D)}\right]$ is the ith sample of a D-dimensional time series, $A_j \in \mathbb{R}^{D\times D}$ are coefficient matrices capturing interdependencies between signals, and e_i is the residual error vector. The coefficients A_j, are estimated using the least-square method, minimizing the residual error across all samples.

Environmental variability, such as changing wind and wave conditions, affects the structural response, demanding data normalisation. Principal Component Analysis (PCA) addresses this by filtering environmental effects without requiring additional data. PCA transforms standardised feature space X into orthogonal principal components based on the covariance matrix C, retaining components that capture the desired variance threshold:

$$X' = XV_m \tag{3}$$

where V_m contains the eigenvectors corresponding to the largest m eigenvalues, retaining a predefined variance threshold. To normalise the features, the first components associated with the environmental variability [4] are removed. Following normalisation, the Mahalanobis distance (MD) quantifies deviations of each sensor's behavior from the baseline condition. The MD for a sensor is defined as:

$$MD_s = \sqrt{(x_s - \bar{x}) \cdot S_x^{-1} \cdot (x_s - \bar{x})^T} \tag{4}$$

Here, x_s is the vector of features for a given sensor, $\bar{x}$ and S_x are the mean vector and covariance matrix of baseline data, respectively. The MD values for all sensors are combined to compute the dissimilarity between simulations using the Euclidean distance:

$$dis_{ij} = \sqrt{\sum_{d=1}^{D} \left(MD_{s=i} - MD_{s=j}\right)^2} \tag{5}$$

The dissimilarity matrix serves as input for a K-means clustering algorithm, partitions simulations into K clusters by minimizing the intra-cluster dissimilarity:

$$W(P_K) = \frac{1}{2} \sum_{k=1}^{K} \sum_{c(i)=k} \sum_{c(j)=k} dis_{ij} \tag{6}$$

where $c(i)$ assigns simulation i to cluster k. The algorithm iteratively assigns simulations to the nearest centroid and updates the centroids until convergence. The optimal number of clusters is determined using the silhouette score, enabling the method to operate autonomously and making it suitable for real-time monitoring applications. In real-time applications, as new structural conditions emerge, the baseline conditions can be updated to improve the detection of subsequent changes.

3 CASE STUDY

Two case studies were analysed using the proposed damage detection algorithm: a monopile-supported offshore wind turbine structure and an experimentally tested blade benchmark. Remarkably, although these are very different structural systems, the proposed method can be successfully applied as discussed below.

3.1 Offshore wind turbine structure virtual model

The first case study involves a virtual model of a bottom-fixed 5-MW NREL OWT [10], illustrated in Figure 1. The model is divided into two components: (i) the supporting structure, modeled in OpenSees [11], and (ii) the turbine, which is modelled in OpenFAST [12]. This setup allows leveraging the capabilities of both software packages and are integrated through the interface force between the turbine and the tower.

Eleven virtual accelerometers, spaced at 10-meter height intervals, are positioned along the structure to measure accelerations in the primary direction of analysis (X-direction), as shown in Figure 1.

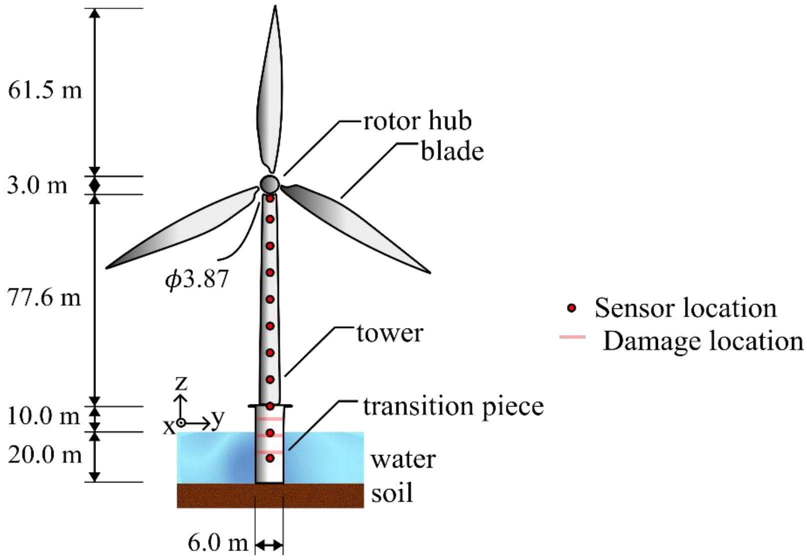

Figure 1: Illustration of the 5-MW NREL OWT model, showing its components, sensor and damage location.

A total of 109 simulations were performed, each generating 35 seconds of useful data sampled at 200 Hz. The dataset includes 91 undamaged (baseline) simulations and 18 damaged simulations. Two operational states of the turbine are considered: operating and parked. For undamaged scenarios, environmental loads consist of steady and turbulent wind conditions, along with white-noise spectrum waves, generated in OpenFAST using Gaussian copulas to model a correlation coefficient of 0.63:

- For the wind, the typical long-term wind mean speed at 90 m above the mean sea level (MSL) is modelled through a Weibull distribution, with mean 10 m/s, shape parameter of 1.14, and scale parameter of 11.28.
- For the waves, the typical long-term significant mean wave height is modelled through a Lognormal distribution, with mean 2.0 m and 0.5 coefficient of variation.

For damage scenarios, localised stiffness reductions of 5%, 10%, and 20% are introduced over 1-meter-long sections at three locations in the transition piece (see Figure 1).

Using 11 sensors and a model order of $N = 5$, the proposed damage detection algorithm extracts 55 MAR parameters per sensor. After computing MDs, these are reduced to a single parameter per sensor, resulting in 11 parameters indicative of the structural condition.

The clustering results from the 109 simulations are shown in Figure 2. Two distinct clusters are observed: the first cluster (light blue) corresponds to the 91 undamaged scenarios, while the second cluster (orange) contains the 18 damaged scenarios. These results demonstrate the algorithm's strong performance in automatically distinguishing structural condition changes. Notably, the MD values for damaged scenarios exhibit exponentially higher values than those for undamaged scenarios, underscoring the robustness of the proposed methodology for the selected sensor configuration and model order.

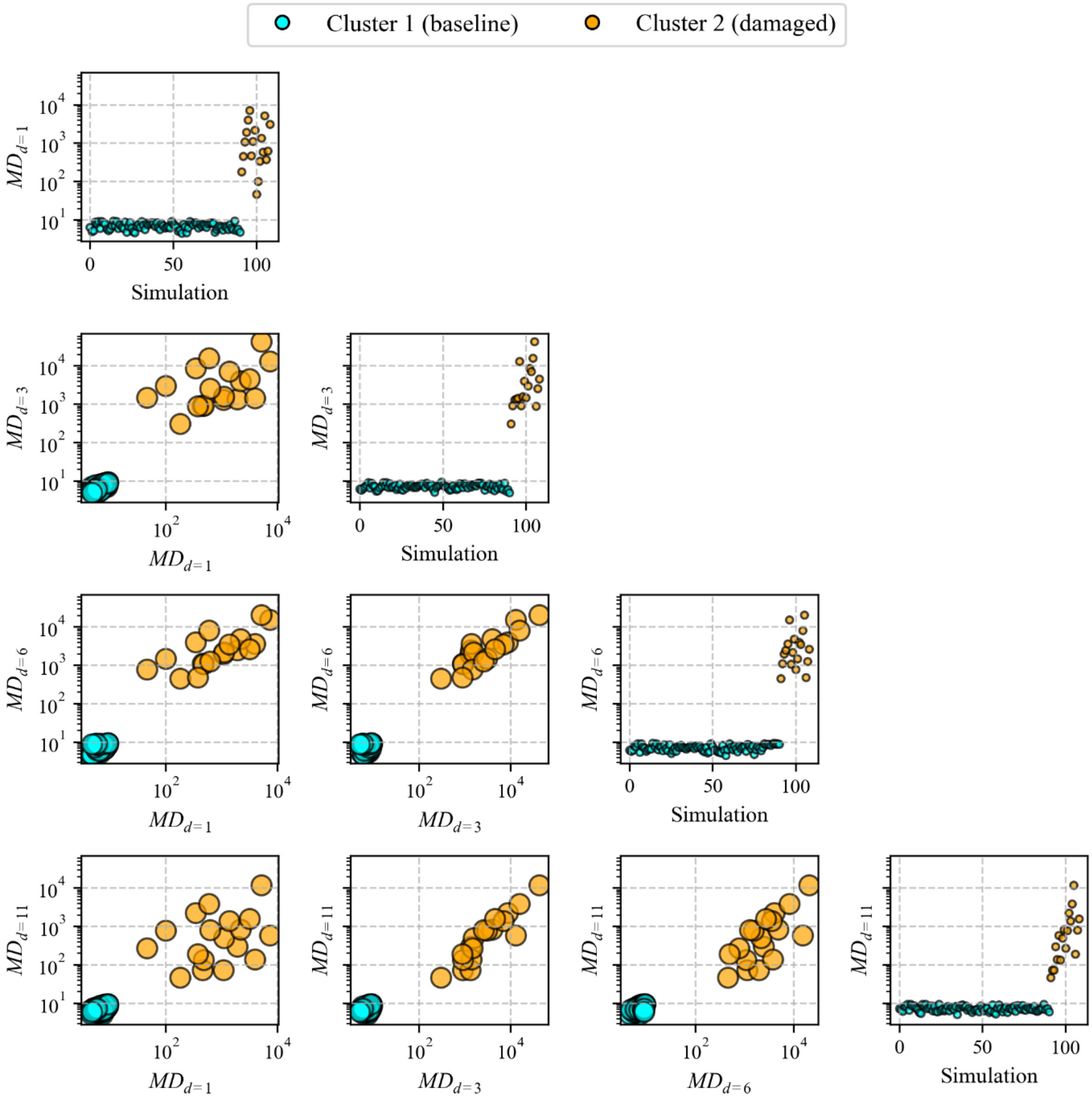

Figure 2: Results of the K-means clustering algorithm using the damage detection algorithm. Diagonal plots display MD values for individual sensors (1, 3, 6 and 11) along the simulations, while off-diagonal plots show the MD values across sensor pairs.

3.2 Offshore wind turbine blades with experimental data

The second application for the damage detection algorithm involves the experimental blade benchmark presented in [13]. This blade is part of a Windspot 3.5 kW wind turbine model, measuring 1.75 m in length with a mass of 5.0 kg. It consists of: (i) the outer shell, which is a 0.93 mm thickness double-layered composite material made of plain-weave fabric and chopped strand mat made of E-glass fibers, and (ii) the inner part, which is filled with a low-stiffness core of polyurethane foam. The blade is supported at one end by a steel frame, subjected at the same time by three concrete columns.

The blade was exposed to controlled temperature variations in a climate chamber, with temperatures ranging from -15°C to +40°C in 5°C increments, while humidity was kept constant at 60%. Dynamic excitation was applied using an electromechanical shaker, which delivered: (i) white noise signals with an effective frequency bandwidth between 0 and 400 Hz and (ii) sine sweeps with frequencies ranging from 1 to 300 Hz.

Figure 3 illustrates the sensor configuration scheme for laboratory testing along with the location of the induced cracks on the blade. A total of eight accelerometers (±10g working range) and 18 strain gauges (unidirectional and rosette types) were used. The strain gauges were configured along two axes, as indicated in Figure 4 with superscripts (1) and (2).

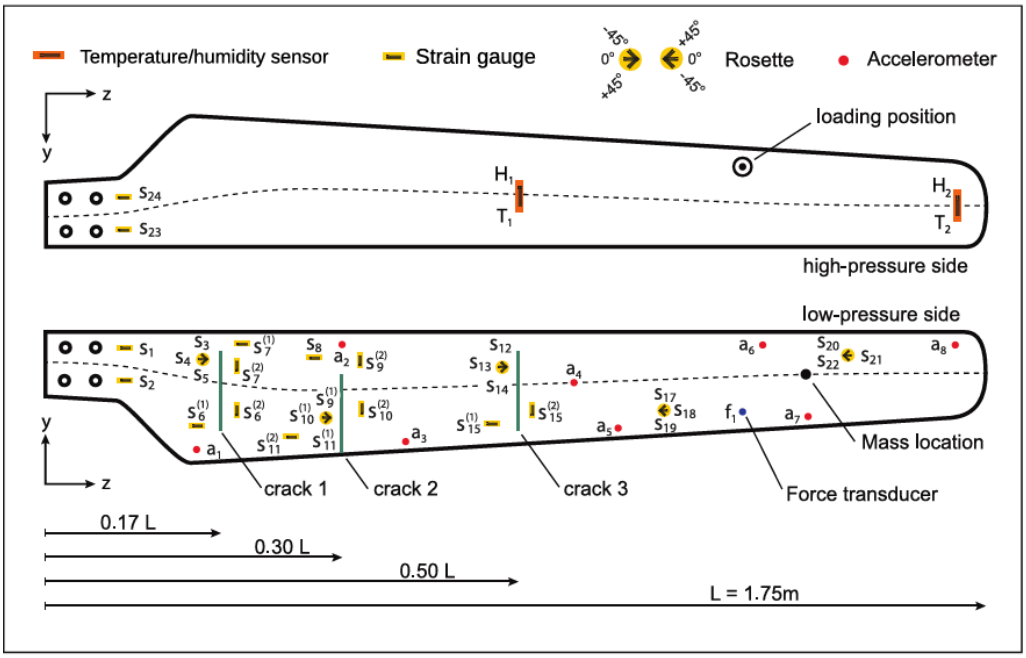

Figure 3: Experimental setup of accelerometers and strain gauges on the tested blade [13].

Progressive damage was induced into the blade in the form of cracks, in addition to the initially healthy state (denoted as R), as specified in Table 1. The cracks had uniform dimensions, with a depth of 4 mm and a width of 1.5 mm.

A total of 20 white noise tests and one sine sweep excitations were carried out in the healthy (undamaged) state, while each damage case was tested with one sine sweep and five white noise excitations.

Case	**Damage configuration**		
	Crack 1	**Crack 2**	**Crack 3**
R	-	-	-
D	5 cm	-	-
E	5 cm	5 cm	-
F	5 cm	5 cm	5 cm
G	10 cm	5 cm	5 cm
H	10 cm	10 cm	5 cm
I	10 cm	10 cm	10 cm
J	15 cm	10 cm	10 cm
K	15 cm	15 cm	10 cm
L	15 cm	15 cm	15 cm

Table 1: Damage configuration cases used for testing in [13].

Using only the strain gauge—unlike the previous analysed case, which used accelerometer data— the dissimilarity results between each experiment from the proposed algorithm are presented in Figure 4. Each damage case is compared to its preceding condition. Dark blue colours represent experiments from the previous blade condition, which is used for algorithm training and, therefore, show no dissimilarities. Other colours indicate varying levels of dissimilarity due to structural changes in the blade.

The results demonstrate that the algorithm effectively detects structural changes across all cases using the information coming from strain gauges, consistently yielding high dissimilarity values. This serves as a visual indicator of how easily the clustering algorithm can identify new structural conditions from the MD indicator—larger dissimilarities make damage detection and classification more straightforward.

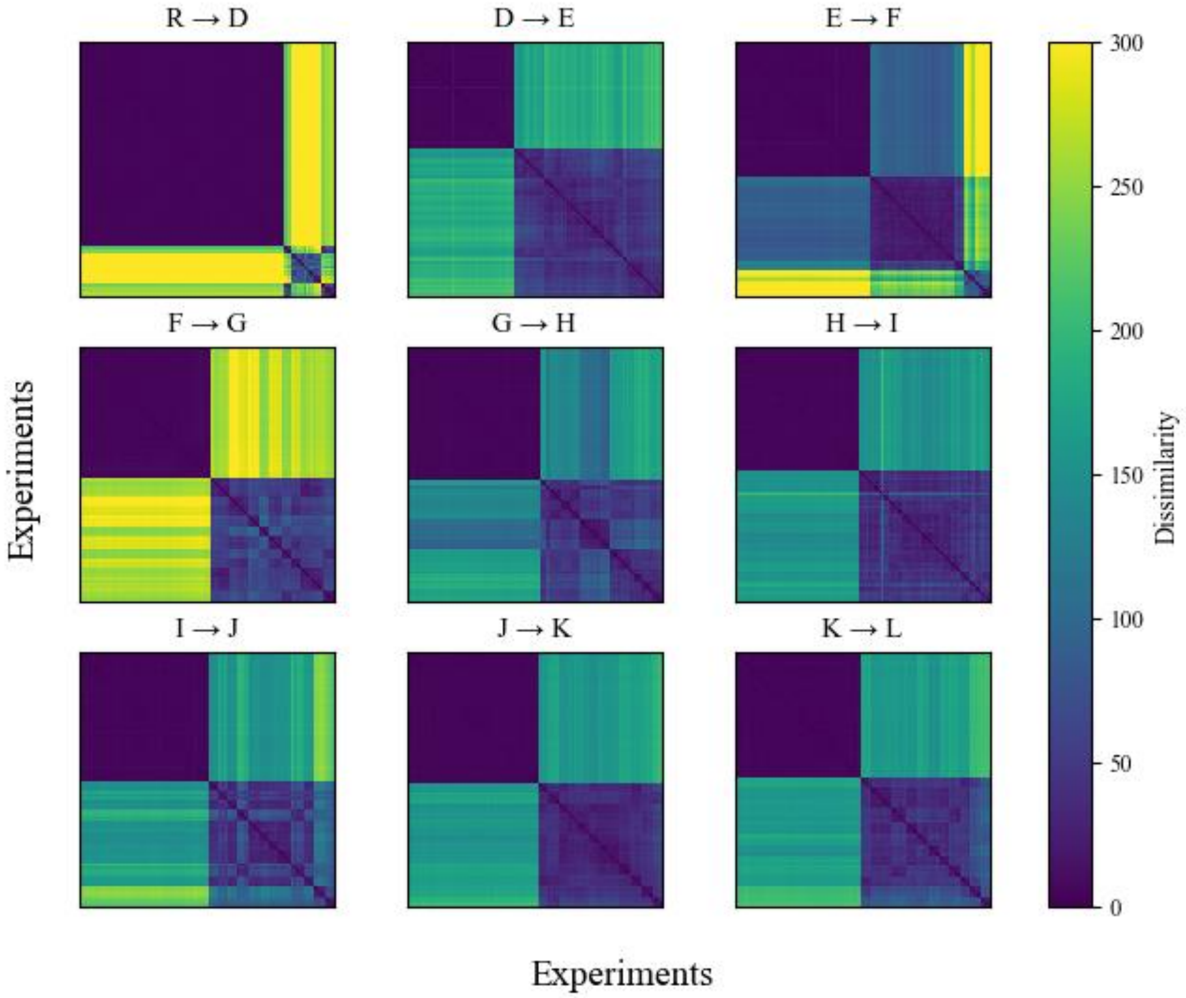

Figure 4: Progression of the dissimilarity heatmaps comparing previous and posterior conditions.

4 CONCLUSIONS

This study introduced a semi-supervised, data-driven approach for detecting damage on two kinds of components of wind turbines systems: (i) the main supporting structure, modelled virtually using a numerical model, and (ii) a blade, experimentally tested in a controlled laboratory environment. This algorithm is composed of four main steps: feature extraction using multivariate autoregressive models, feature modelling using principal component analysis, feature fusion with Mahalanobis distance, and clustering using K-means. The algorithm demonstrated high sensibility to structural changes on the monitored systems, capturing early damage based only on the outputs of the system and trained only on previous structural conditions. Its low computational demand and simplicity make it highly suitable for real-time structural health monitoring applications. Finally, the use of the acceleration time-series in the first case study and strain time-series in the second one demonstrates the flexibility of the proposed data-driven methodology for autonomous output-only damage detection.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from the City St George's, University of London through the Research at the Frontiers of Practice PhD Scholarship. Authors would also like to express their sincere gratitude to The Alan Turing Institute and the Lloyd's Register Foundation for their funding support via the Data-Centric Engineering group. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect those of the sponsors.

REFERENCES

[1] C. R. Farrar and K. Worden, 'An introduction to structural health monitoring', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 303–315, Feb. 2007, doi: 10.1098/rsta.2006.1928.

[2] Q. Han, Q. Ma, J. Xu, and M. Liu, 'Structural health monitoring research under varying temperature condition: a review', *J Civ Struct Health Monit*, vol. 11, no. 1, pp. 149–173, 2021, doi: 10.1007/s13349-020-00444-x.

[3] K. Eltouny, M. Gomaa, and X. Liang, 'Unsupervised Learning Methods for Data-Driven Vibration-Based Structural Health Monitoring: A Review', *Sensors*, vol. 23, no. 6, 2023, doi: 10.3390/s23063290.

[4] J. P. Santos, C. Crémona, A. D. Orcesi, and P. Silveira, 'Multivariate statistical analysis for early damage detection', *Eng Struct*, vol. 56, pp. 273–285, 2013, doi: https://doi.org/10.1016/j.engstruct.2013.05.022.

[5] A. Meixedo, J. Santos, D. Ribeiro, R. Calçada, and M. Todd, 'Damage detection in railway bridges using traffic-induced dynamic responses', *Eng Struct*, vol. 238, p. 112189, 2021, doi: https://doi.org/10.1016/j.engstruct.2021.112189.

[6] A. Meixedo, J. Santos, D. Ribeiro, R. Calçada, and M. D. Todd, 'Online unsupervised detection of structural changes using train–induced dynamic responses', *Mech Syst Signal Process*, vol. 165, p. 108268, 2022, doi: https://doi.org/10.1016/j.ymssp.2021.108268.

[7] K. J. Vamvoudakis-Stefanou, J. S. Sakellariou, and S. D. Fassois, 'Vibration-based damage detection for a population of nominally identical structures: Unsupervised Multiple Model (MM) statistical time series type methods', *Mech Syst Signal Process*, vol. 111, pp. 149–171, 2018, doi: https://doi.org/10.1016/j.ymssp.2018.03.054.

[8] D. Tcherniak and L. L. Mølgaard, 'Active vibration-based structural health monitoring system for wind turbine blade: Demonstration on an operating Vestas V27 wind turbine', *Struct Health Monit*, vol. 16, no. 5, pp. 536–550, 2017, doi: 10.1177/1475921717722725.

[9] T. Stehly, P. Duffy, and D. Mulas Hernando, '2022 Cost of wind energy review [slides] (No. NREL/PR-5000-84774)', 2023, *National Renewable Energy Laboratory (NREL), Golden, CO (United States)*. doi: https://doi.org/10.2172/2278805.

[10] J. Jonkman, S. Butterfield, W. Musial, and G. Scott, 'Definition of a 5-MW Reference Wind Turbine for Offshore System Development', Feb. 2009. doi: 10.2172/947422.
[11] F. McKenna, 'Open System for Earthquake Engineering Simulation', *University of California*, 2000.
[12] National Renewable Laboratory of Energy, 'OpenFAST'. Accessed: Jan. 10, 2025. [Online]. Available: https://openfast.readthedocs.io/en/v3.5.3/
[13] Y. Ou, K. E. Tatsis, V. K. Dertimanis, M. D. Spiridonakos, and E. N. Chatzi, 'Vibration-based monitoring of a small-scale wind turbine blade under varying climate conditions. Part I: An experimental benchmark', *Struct Control Health Monit*, vol. 28, no. 6, p. e2660, 2021, doi: https://doi.org/10.1002/stc.2660.