Available online at www.eccomasproceedia.org
Eccomas Proceedia SMART (2023) 1124-1135

ECCOMAS Proceedia

X ECCOMAS Thematic Conference on Smart Structures and Materials
SMART 2023
D.A. Saravanos, A. Benjeddou, N. Chrysochoidis and T. Theodosiou (Eds)

INTERPRETING ENVIRONMENTAL VARIABILITY FROM DAMAGE SENSITIVE FEATURES

Josep Font-Moré[a], Luis David Avendaño-Valencia[b], David Garcia Cava[c] and Marco A. Pérez[a]

[a]Group of Applied Mechanics and Advanced Manufacturing - IQS School of Engineering
Universitat Ramon Llull
Via Augusta, 390, 08017 Barcelona, Spain
e-mail: josepfontm@iqs.url.edu, marcoantonio.perez@iqs.url.edu

[b] Department of Mechanical and Electrical Engineering
University of Southern Denmark (SDU)
Campusvej 55, 5230 Odense, Denmark
e-mail: ldav@sdu.dk

[c] Institute for Infrastructure and Environment, School of Engineering
University of Edinburgh
Thomas Bayes Road, EH9 3FG Edinburgh, United Kingdom
e-mail: david.garcia@ed.ac.uk

Key words: Damage Detection, Structural Health Monitoring, Operational Modal Analysis, Vibration Analysis, Environmental and Operational Mitigation, Principal Component Analysis

Abstract. Mitigation of Environmental and Operational Variabilities (EOVs) remains one of the main challenges to adopt Structural Health Monitoring (SHM) technologies. Its implementation in wind turbines is one of the most challenging due to the adverse weather and operating conditions these structures have to face. This work proposes an EOV mitigation procedure based on Principal Component Analysis (PCA) which uses EOV-Sensitive Principal Components (PCs) as a surrogate of EOVs, which may be hard to measure or correctly quantify in real-life structures. EOV-Sensitive PCs are conventionally disregarded in an attempt to mitigate the effect of environmental variability. Instead, we postulate to use these variables as predictors in non-linear regression models, similar to how Environmental and Operational Parameters (EOPs) are used in explicit EOV mitigation procedures. The work results are validated under an experimental dataset of a small-scale wind turbine blade with various cracks artificially introduced. Temperature conditions are varied using a climate chamber. The proposed method outperforms the conventional-PCA based approach, implying that directly disregarding Sensitive-EOV PCs is detrimental in the decision-making within a SHM methodology. In addition, the proposed method achieves similar results to an equivalent explicit procedure, suggesting that EOV-Sensitive PCs can replace directly measured EOVs.

© 2023 The Authors. Published by Eccomas Proceedia.
Peer-review under responsibility of the organizing committee of SMART 2023.
doi: 10.7712/150123.9862.451124

1 INTRODUCTION

Maintenance is one of the key stages of any infrastructure life-cycle, expanding from its construction all the way to its end of life. Most structures are periodically examined by an inspection team, which surveys the structure's integrity and proposes repair tasks if necessary. Nonetheless, conventional maintenance tasks have two main drawbacks. First, the structure's inspection is operator-based and, therefore, prone to human errors. Second, the gaps between inspections, usually years, challenge the early detection of damage. These drawbacks led to the development of Structural Health Monitoring (SHM) [1] procedures, which have been widely researched over the last two decades. Among the different sensing technologies currently available, Vibration-based SHM (VSHM) [2] has been one of the main research areas. These methods fundamentally study the changes in the dynamic response of the structure under unexpected alterations (damage). Vibration data acquired from the system is used to extract different Damage Sensitive Features (DSFs), which attempt to give insights regarding the structure's status [3].

Although there exists a wide literature on the field of SHM [4], several challenges still compromise its widespread adoption in many applications. One of the main challenges in the field of SHM is the mitigation of Environmental and Operational Variabilities (EOV). External factors, such as temperature, traffic loading, wind or humidity, diminish the sensitivity of DSFs toward damage and ultimately reduce the accuracy of SHM systems. Several studies have reported that the global effects caused due to Environmental and Operational Parameters (EOPs) can be greater than the local changes caused by different damage scenarios present in the structure [5]. In order to tackle this problem, two types of strategies have been proposed in the literature: Explicit and Implicit procedures [6].

Explicit procedures approach EOV mitigation by fusing information from DSFs and EOPs to infer the structural integrity in different conditions. The main goal is to relate an external source, namely EOV, to an effect on DSFs. Finding the fundamental relationship between EOPs and DSFs is usually approached via deterministic or stochastic models solved using different regression strategies. On the one hand, deterministic models attempt to fit a function to the available observations. These models are simpler to optimize and can be used for predictions. Nonetheless, they are prone to overfitting and sensitive to outliers. On the other hand, stochastic models attempt to quantify the inherent uncertainty of the data in addition to the output trends. Many of these methods are described in a Bayesian framework using the prior (original belief on the model) and posterior distributions. Even though the use of more information regarding the conditions of the structure seems like an advantage at first glance, several questions arise when considering this strategy [7]. Are gradient conditions present in the structure? Which are the most representative places to measure EOPs?

In comparison, implicit procedures only rely on DSFs to assess structural integrity. Previous studies have reported that implicit strategies can outperform explicit ones [8], suggesting that information about environmental and/or operational phenomena can be detrimental if inappropriately measured. One of the most common implicit approaches is the use of projection methods, such as Factor Analysis, Principal Component Analysis (PCA) [9, 10] and Singular Spectrum Analysis (SSA) [11]. PCA is one of the most common strategies for damage detection in VSHM due to its simplicity. One of the main characteristics of this method is the uneven distribution of EOV effects in its Principal Components (PCs) [8], which has been conventionally approached by disregarding the first PCs, which are considered EOV-Sensitive. Alternatively, some studies have proposed using DSFs to inform the model about EOVs. In this sense, natural frequencies have been used as both dependent, and independent variables [12]. The premise being that lower natural frequencies will be more correlated to EOVs than damage and thus can be used as an alternative measure for EOVs.

Based on the above, this paper proposes a new implicit procedure (based solely on vibration features) to detect damage in changing environments based on Principal Component Analysis (PCA) to process an initial set of vibration features from the structure's response. The main assumption behind the PCA-based method is that the variability in the healthy state of the structure originates from EOVs. Therefore, the Principal Components (PCs) with the largest eigenvalues can be associated with the EOV influence. Conventionally, these PCs will be disregarded as a way to mitigate the effect of EOV. Nonetheless, the novelty of our approach is to utilize these so-called EOV-sensitive PCs as a surrogate of the Environmental/Operational variables driving the non-stationary behaviour in the DSFs. Hence, a regression model using EOV-sensitive PCs as predictors and remaining PCs as explained variables is proposed.

The effectiveness of this approach is assessed using an experimental vibration dataset from a small-scale wind turbine blade (WTB) located in a climate chamber. Temperature is the main driver for EOV, while various damage conditions are introduced. The proposed method is then compared with the conventional PCA-based approach and with an explicit regression-based method which uses temperature as a predictor for DSFs.

This document is organized as follows: Section 2 describes the methodology presented in this work. Section 3 briefly describes the experimental dataset used. Section 4 includes the results and discussions. Finally, the conclusions of the study are summarized in Section 5.

2 METHODOLOGY

Figure 1 illustrates the methodological approach followed in this paper. The proposed method is contrasted with two classic EOV procedures (Implicit PCA and Explicit Regression) to assess the performance achieved. All methodologies make use of the squared Mahalanobis Distance in order to obtain the damage index. Damage thresholds are defined using a percentile of the training data. The main difference among all methods is the EOV procedure used and, more specifically, how EOPs are introduced to the mitigation model. The methodology followed in this work has been divided into three main steps: Damage sensitive features extraction, EOV Procedures and Outlier Analysis. Each one of the steps are summarized next.

2.1 Damage-sensitive Features Extraction

DSFs are extracted from the spectral representation of response signals ($\mathbf{X}$), here obtained via Welch's PSD estimator. Although the dataset contains excitation-response pairs, presently, we consider the response-only case, which is more similar to a real-life situation. The responses

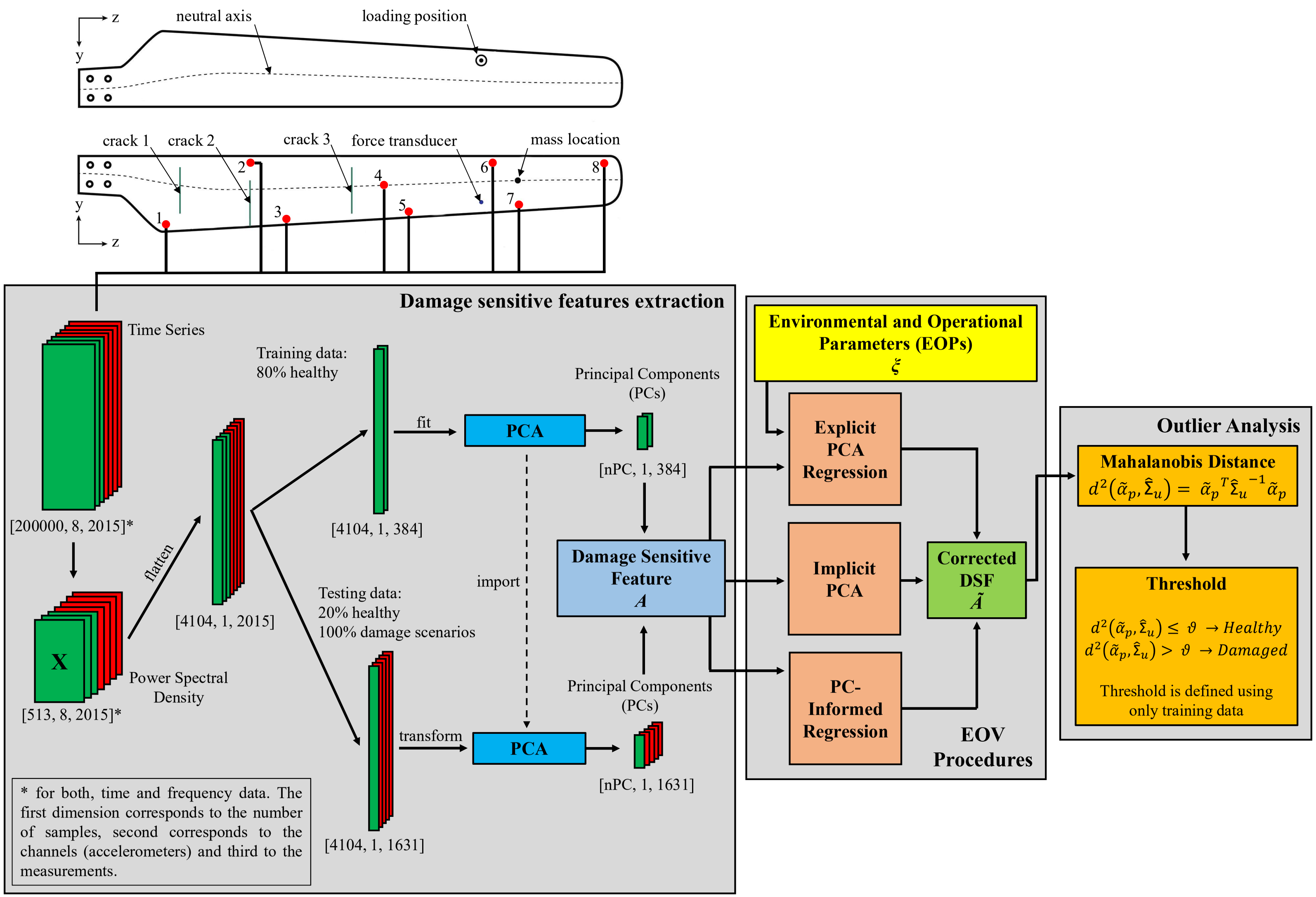

Figure 1: Methodological approach followed in this work.

for different accelerometers (s) are then merged to form a feature vector $\boldsymbol{x}_j$.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x_{1,1}} & \boldsymbol{x_{1,2}} & \cdots & \boldsymbol{x_{1,N}} \\ \boldsymbol{x_{2,1}} & \boldsymbol{x_{2,2}} & \cdots & \boldsymbol{x_{2,N}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x_{s,1}} & \boldsymbol{x_{s,2}} & \cdots & \boldsymbol{x_{s,N}} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \boldsymbol{x_1} & \boldsymbol{x_2} & \cdots & \boldsymbol{x_N} \\ | & | & & | \end{bmatrix} \tag{1}$$

where, $\boldsymbol{x}_{i,j}$ indicates the response spectrum vector from sensor i and a specific observation j, and N indicates the total number of samples.

Subsequently, the obtained PSDs are PCA-transformed to obtain the initial set of DSFs. This strategy allows to reduce the high dimensionality of the PSD while retaining the most part of the dataset variance in a compact set of Principal Components (PCs), which are used as input for the EOV procedure. In summary, the original PSD features are transformed as follows:

$$\boldsymbol{\alpha}_j = \boldsymbol{V}^T \boldsymbol{x_j} \tag{2}$$

where $\boldsymbol{\alpha}_j$ is the vector of uncorrelated variables, namely principal scores, and $\boldsymbol{V}$ is the matrix, where the columns contain the PCs. Finally, DSFs are organized in the data matrix:

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_j & \cdots & \boldsymbol{\alpha}_N \end{bmatrix} \in \mathbb{R}^{K \times N} \tag{3}$$

where, $\boldsymbol{\alpha_p}$ is the non-corrected DSF of the j-observation, K stands for the dimension of the feature vector and N for the total number of observations.

2.2 EOV Procedure

In this work we consider three different DSF post-processing procedures to control the effect of EOV [6, 13]. The first two: *Implicit PCA* and *Explicit Regression*, are used as a baseline to assess the results obtained by the proposed method *PC-Informed Regression.* Figure 2 shows each method's different steps towards EOV mitigation, and a brief description is provided next.

The rationale behind PCA is that the first PC contains the highest contribution to the overall variance of the feature vector, PC_2 contains the second highest, and so on. In that sense, the *Implicit PCA* procedure attempts to reduce the effect of EOV by disregarding the set of PCs with largest eigenvalues, as these are considered to be most influenced by EOV. Resulting from this procedure is the cropped PC feature matrix, as follows:

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \tilde{\boldsymbol{\alpha}}_1 & \tilde{\boldsymbol{\alpha}}_2 & \cdots & \tilde{\boldsymbol{\alpha}}_j & \cdots & \tilde{\boldsymbol{\alpha}}_N \end{bmatrix} \in \mathbb{R}^{\tilde{K} \times N} \tag{4}$$

where $\tilde{K} < K$ is the dimension of the reduced PC feature vector, and $\tilde{\boldsymbol{\alpha}}_j$ the vector of $\tilde{K}$ EOV-insensitive PC features.

Both, *Explicit* and *PC-Informed* regression methods use non-linear regression to adjust the variation of the DSFs to EOV. This is done by finding the best fitting function for the data in the least squares sense. The main difference between the two methods is that in *Explicit regression*, measured EOPs are used as predictors for the DSFs. On the other hand, in *PC-informed regression*, the most EOV-sensitive PCs are used as predictors instead.

In both methods, a part of the training data (Healthy State) is used to fit the polynomial functions, while the rest is used to calculate the validation Mean Squared Error (MSE). In addition, the polynomial order is presently limited to 5 to control the model complexity. Figure 3 shows the regression models obtained using only training data.

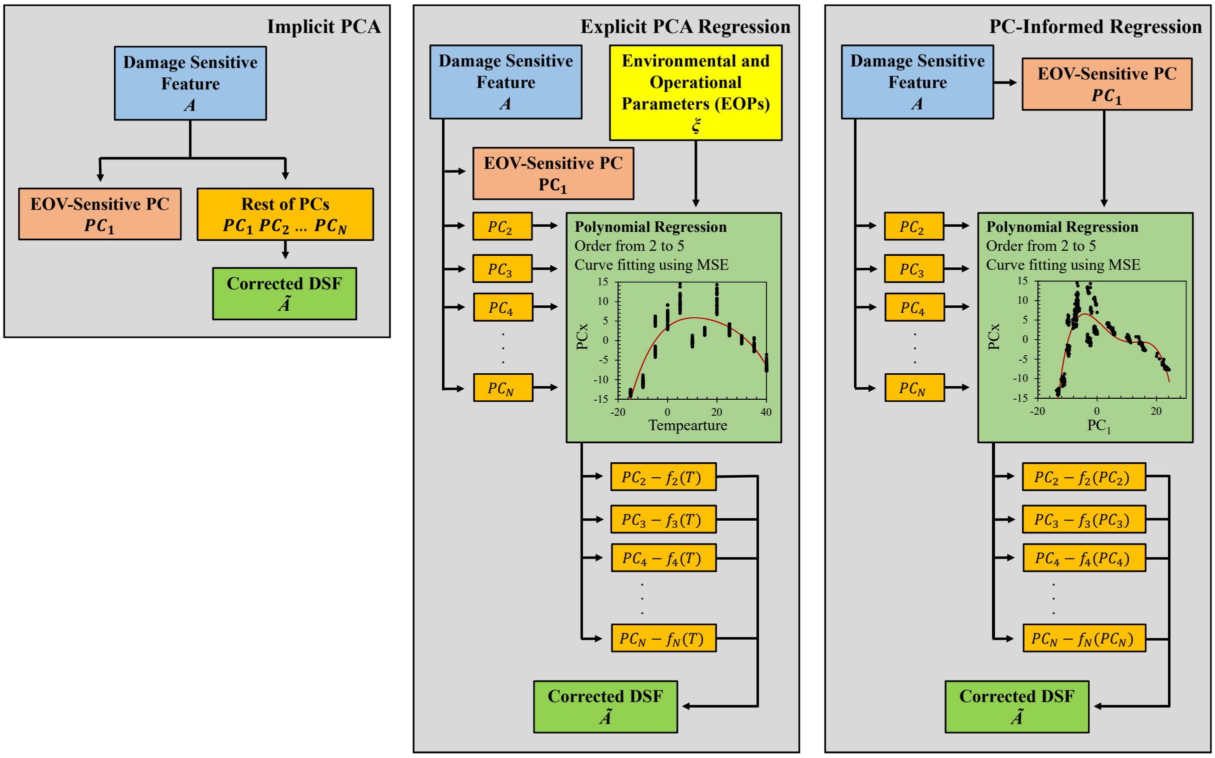

Figure 2: Schemes for EOV Procedures [13] used in this work.

2.3 Outlier Analysis

Outlier analysis is usually applied for damage assessment after mitigating environmental effects. Outlier analysis defines boundaries based on healthy measurements (training data) to detect whether a new observation corresponds to a healthy or damaged condition. Here, a damage index is calculated from the corrected DSFs with the squared Mahalanobis distance (MD), as follows:

$$d^2(\tilde{\boldsymbol{\alpha}}_j, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}) = \tilde{\boldsymbol{\alpha}}_j^T \, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}^{-1} \, \tilde{\boldsymbol{\alpha}}_j \tag{5}$$

where $\tilde{\boldsymbol{\alpha}}_j$ is the corrected DSF of the j-observation and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}$ is the covariance of the training data. A simple damage diagnosis can be established by comparing the above-defined damage index with a threshold $\vartheta > 0$, which is defined as a percentile of the training data. Each observation can then be classified as follows:

$$H_1 : d^2(\tilde{\boldsymbol{\alpha}_p}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}) \leq \vartheta \rightarrow \text{Healthy} \tag{6a}$$

$$H_2 : d^2(\tilde{\boldsymbol{\alpha}_p}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}) > \vartheta \rightarrow \text{Damaged} \tag{6b}$$

2.4 Model Evaluation

Performance evaluation of each one of the considered methods is done with the F1 score. This performance measure is well suited for imbalanced datasets like the one considered in this work. First, the confusion matrix is built based on the predicted labels and true labels for each

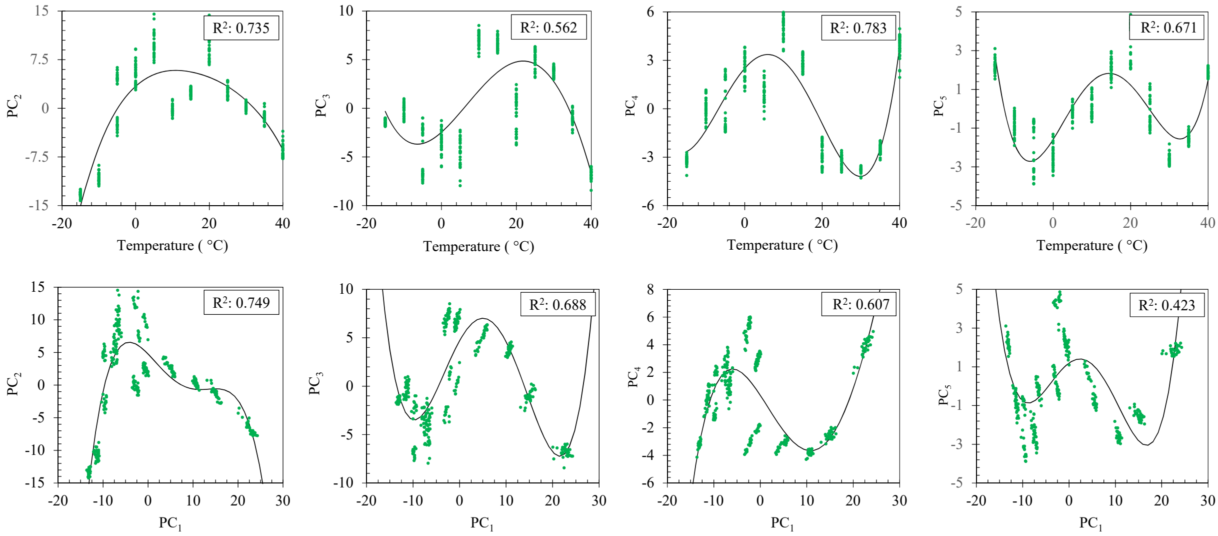

Figure 3: Polynomial regression fit for first Principal Components using Explicit Regression Method (first row) and PC-Informed regression (second row).

damage scenario. The F1-Score is defined as follows:

$$F1\ Score = \frac{2TP}{2TP + FN + FP} \tag{7}$$

where each measurement is classified into one of four categories: True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN).

2.5 Mutual information between Temperature and PCs

Mutual Information (MI) is used to quantify the degree of statistical dependence between PCs and Temperature. This metric allows to assess which variables share the most information objectively. Mutual Information is a non-negative metric that yields values between 0, meaning variables are statistically independent and $+\infty$. The function *mutual_info_classif* from the Python library *Scikit-Learn* was used to estimate MI for each Principal Component. Several algorithms exist to estimate MI, depending on the nature of vectors: discrete or continuous. Presently, a continuous approach based on k-nearest neighbour distances [14] is used as follows:

$$I(\boldsymbol{PC_X}, \boldsymbol{T}) = \psi(k) + \psi(N) - [\psi(n_{PC_X} + 1) + \psi(n_T + 1)] \tag{8}$$

where $\psi(x)$ is the digamma function, k is the number of neighbours to search for at each point (predefined as 3), N is the number of observations and n_{PC_X} and n_T are the number of neighbours within a specified radius r for each of the vectors.

The higher the mutual information is, the more information one variable holds about the other. For the sake of clarity, Mutual Information is usually normalized between 0 and 1 to ease the comparison with commonly used correlation methods.

As shown in Figure 4a, the influence of temperature on PCs is not equally distributed. Instead, PC_1 contains more information about temperature than the rest. Mutual Information

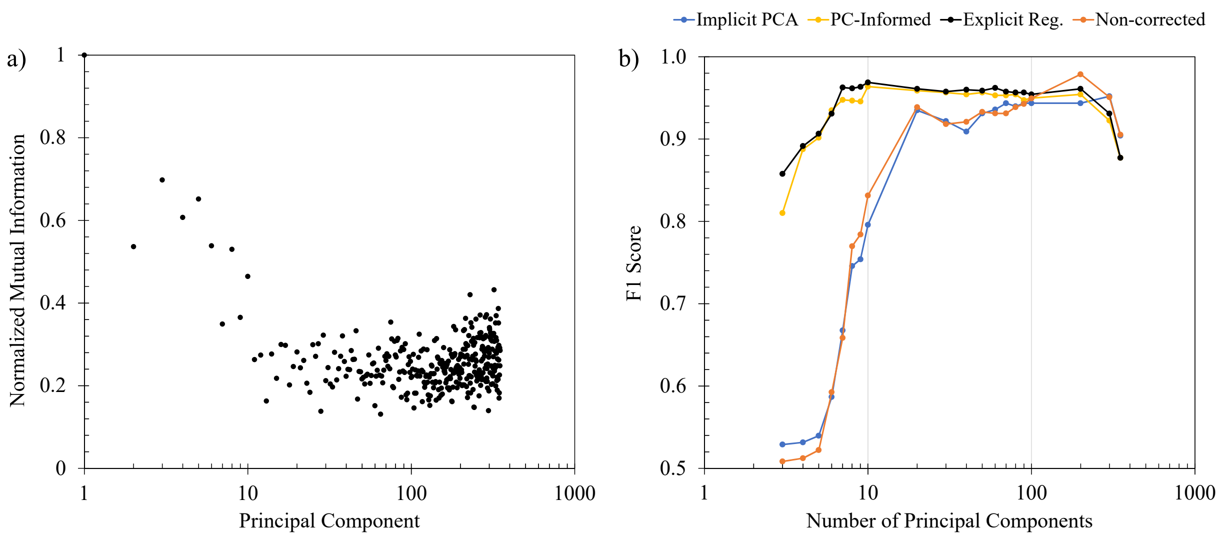

Figure 4: a) Normalized Mutual Information between Principal Components and temperature. b) Performance for each method based on different sets of Principal Components.

follows a downward trend, with higher PCs sharing less information with EOPs. As mentioned above, PC_1 holds the most information on the temperature amongst the PCs, which justifies its use as a surrogate to construct a regression model for EOV compensation. This is precisely used in the proposed PC-informed approach.

3 EXPERIMENTAL DATA

The data for this research was taken from a small-scale WTB tested under different temperature conditions (-15 to 40 °C) using a climate chamber. The experimental tests were carried out by the Chair of Structural Mechanics and Monitoring of ETH Zürich. The dataset comprises vibration responses measured from a small-scale wind turbine blade in its healthy state and under two different damage typologies: ice accumulation, emulated through mass addition, and stiffness reduction, implemented as cuts across the blade. The total mass of the blade is around 5kg as reported by [15] while the added mass represents at maximum a 3% increase approximately.

A brief explanation of the experimental tests is given in this work, especially focusing on vibration analysis. Nonetheless, a full description of the experimental testing can be found in [15]. The turbine blade was instrumented using 8 accelerometers and multiple strain gauges in different layouts. Although the experimental testing included two different excitation forces: sine sweep and white noise signal, only the latter has been used due to its higher resemblance to real conditions for these types of structures. Each observation had a duration of 120s and a bandwidth between 0 and 400 Hz. Each damage scenario tested in the WTB is depicted in Table 1. Figure 5 shows PSD estimators for different environmental conditions.

Table 1: List of experimental cases tested in the experimental campaign [15]. Cracks 1,2 and 3 correspond to different locations along the blade. Blade has a total mass of 5kg.

Case label	Description		
R	Healthy state		
A	Added mass 1 x 44g (0.88% mass increase)		
B	Added mass 2 x 44g (1.76% mass increase)		
C	Added mass 3 x 44g (2.64% mass increase)		
D	Crack 1: l_1 = 5 cm		
E	Crack 1: l_1 = 5 cm	Crack 2: l_2 = 5 cm	
F	Crack 1: l_1 = 5 cm	Crack 2: l_2 = 5 cm	Crack 3: l_3 = 5 cm
G	Crack 1: l_1 = 10 cm	Crack 2: l_2 = 5 cm	Crack 3: l_3 = 5 cm
H	Crack 1: l_1 = 10 cm	Crack 2: l_2 = 10 cm	Crack 3: l_3 = 5 cm
I	Crack 1: l_1 = 10 cm	Crack 2: l_2 = 10 cm	Crack 3: l_3 = 10 cm
J	Crack 1: l_1 = 15 cm	Crack 2: l_2 = 10 cm	Crack 3: l_3 = 10 cm
K	Crack 1: l_1 = 15 cm	Crack 2: l_2 = 15 cm	Crack 3: l_3 = 10 cm

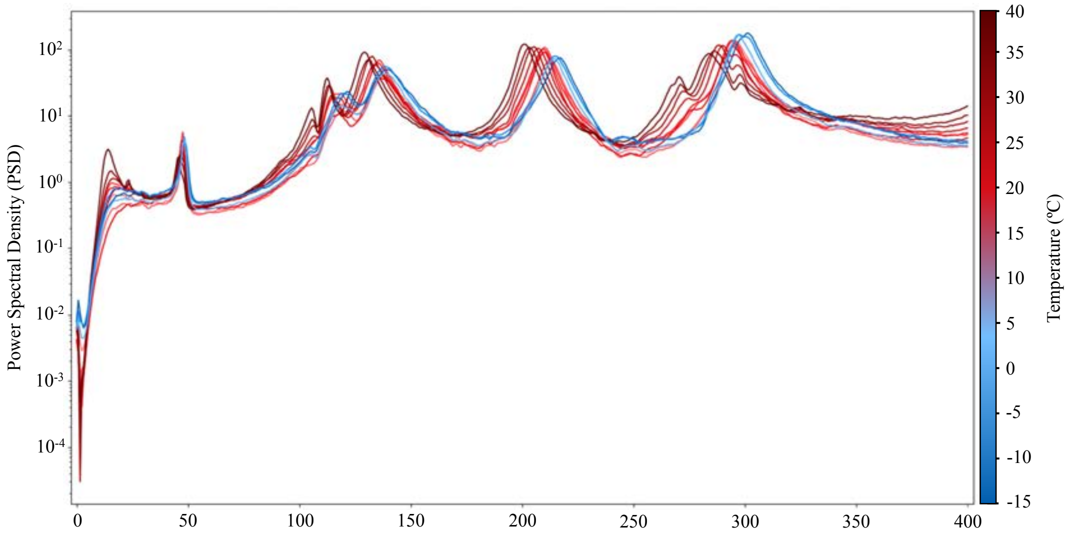

Figure 5: Average Power Spectral Density estimators for healthy state under different temperature conditions.

4 RESULTS AND DISCUSSION

4.1 Performance evaluation of the proposed method

The performance of all methods is measured using the F1 Score as mentioned in Section 2.4. Although not considering all variables affecting a method's performance, this work focuses on how the number of PCs influences classification results. In Figure 4b, the performance of methods is tested using a wide range of Principal Components. The non-corrected DSF is used as a baseline to see if EOV Procedures improve or diminish damage detection capabilities.

Both, Explicit Regression and PC-Informed Regression have similar performance. Initially, these methods take advantage of added PCs, as can be seen by the increase in performance until reaching PC_{10}. From there, using more PCs does not substantially improve performance until reaching PC_{200}, after which the performance drops considerably. Additional PCs reach the noise subspace and therefore add uncertainty to the DSFs. As for the implicit method, it underperforms compared to the other two strategies used in this work.

Damage Indexes for each method can be seen in Figure 6. Implicit metholodogy (Figure 6a) shows unnoticed damages for almost all cases. This becomes prevalent in low-level stiffness damages (D-F). Explicit Regression (Figure 6b) has almost no unnoticed damages, in part due to the lower dispersion of its results. The proposed method achieves similar results to the Explicit Regression, as seen in Figure 6c. Damage scenarios A,B and C correspond to increasing mass additions correctly identified by the method and assigned the expected severity. As for stiffness-based damage scenarios corresponding to cases D to L. The method correctly assigns the expected severity, with L having the highest damage index due to the extreme damage conditions introduced. Finally, the results for non-corrected DSF are given in Figure 6d.

5 CONCLUSIONS

This study provides an alternative EOV mitigation procedure, named PC-Informed Regression, that uses sensitivity towards EOVs to model the interdependencies with the DSF. The regression model is used to mitigate the effects of temperature in DSFs.

The results show that the proposed method achieves similar results to its explicit counterpart without considering data from environmental parameters. Therefore, suggesting that the Environmental and Operational Parameters (EOPs) can be deduced solely from Principal Components, avoiding the need to measure EOPs directly. Furthermore, the method correctly assesses the damage severity as seen by the upward trend in groups: A-C and D-L, respectively. This strategy could be beneficial when EOPs are not trivial to measure, especially in large structures where Environmental and Operational Parameters may not be homogeneous.

Future works should consider testing this EOV Procedure under the combined influence of multiple EOPs or large structures with temperature gradients.

References

[1] C. R. Farrar and K. Worden, *Structural health monitoring: a machine learning perspective.* John Wiley & Sons, 2012.

[2] R. Lado-Roigé and M. A. Pérez, "Stb-vmm: Swin transformer based video motion magnification," *Knowl. Based Syst.*, p. 110493, 2023.

[3] J. Font-Moré, G. Reyes-Carmenaty, R. Lado-Roigé, and M. A. Pérez, "Performance analysis of vibration-based damage indicators under low-modal information structures," *Mech. Syst. Signal Process.*, vol. 190, p. 110166, 2023.

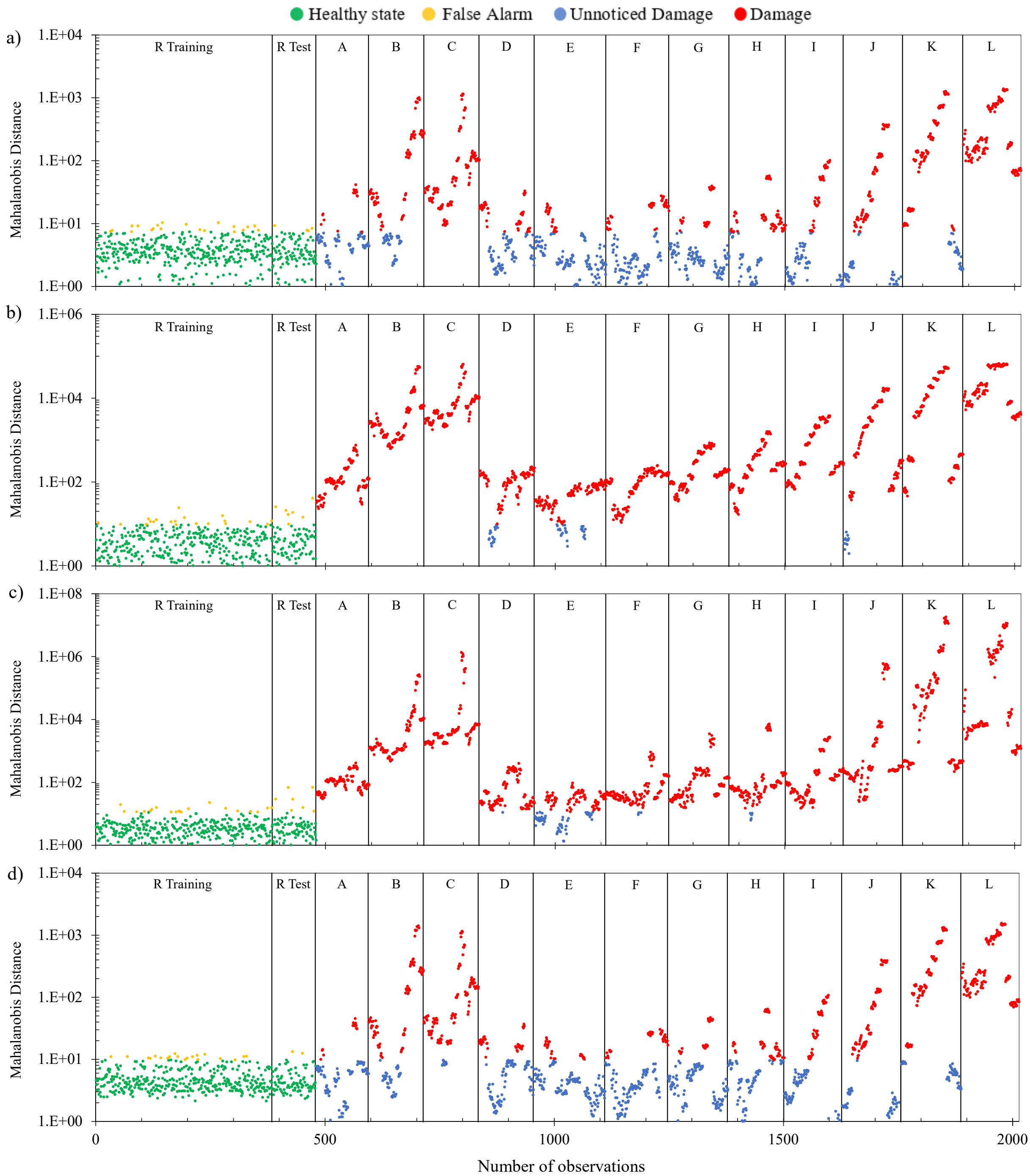

Figure 6: Damage detection control chart for the wind turbine blade. a) Implicit PCA b) Explicit Regression c) PC-Informed Regression d) Non-corrected DSF

[4] E. Figueiredo and J. Brownjohn, "Three decades of statistical pattern recognition paradigm for shm of bridges," *Struct. Health Monit.*, vol. 21, no. 6, pp. 3018–3054, 2022.

[5] B. Peeters and G. De Roeck, "One-year monitoring of the z24-bridge: environmental effects versus damage events," *Earthquake engineering & structural dynamics*, vol. 30, no. 2, pp. 149–171, 2001.

[6] D. García Cava, L. D. Avendaño-Valencia, A. Movsessian, C. Roberts, and D. Tcherniak, *On Explicit and Implicit Procedures to Mitigate Environmental and Operational Variabilities in Data-Driven Structural Health Monitoring.* Cham: Springer International Publishing, 2022, pp. 309–330.

[7] J. Kullaa, "Eliminating environmental or operational influences in structural health monitoring using the missing data analysis," *J. Intell. Mater. Syst. Struct.*, vol. 20, no. 11, pp. 1381–1390, 2009.

[8] B. Ali Qadri, L. Avendano-Valencia, J.-O. Hooghoudt, D. Tcherniak, and M. Ulriksen, "Removal of environmental and operational effects in damage detection: A comparative study with an operating wind turbine," 12 2020.

[9] A.-M. Yan, G. Kerschen, P. De Boe, and J.-C. Golinval, "Structural damage diagnosis under varying environmental conditions—part i: A linear analysis," *Mech. Syst. Signal Process.*, vol. 19, no. 4, pp. 847–864, 2005.

[10] C. Roberts, D. Garcia, and D. Tcherniak, "A comparative study on data manipulation in pca-based structural health monitoring systems for removing environmental and operational variations," pp. 182–198, 2020.

[11] I. Trendafilova, "Singular spectrum analysis for the investigation of structural vibrations," *Engineering Structures*, vol. 242, p. 112531, 2021.

[12] W. Soo Lon Wah, Y.-T. Chen, and J. S. Owen, "A regression-based damage detection method for structures subjected to changing environmental and operational conditions," *Eng. Struct.*, vol. 228, p. 111462, 2021.

[13] J. Font-Moré, "Eovmitigation," 3 2023. [Online]. Available: https://github.com/josepfontm/EOVmitigation

[14] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, p. 066138, Jun 2004.

[15] 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.