A Robust PCA-based Framework for Long-Term Condition Monitoring of Civil Infrastructures

M. Mousavi, A. H. Gandomi

Faculty of Engineering and Information Technology

University of Technology Sydney, Ultimo, NSW 2007, Australia

ABSTRACT

This paper proposes an output-only method for condition monitoring of civil infrastructures through studying a couple of lowest structural natural frequency signals. The main challenge in this sort of problem is to mitigate the effect of the Environmental and Operational Variations (EOV) on the structural natural frequencies to avoid misinterpretation of these effects as damage. To this end, a robust Principal Component Analysis (PCA)-based approach is proposed that uses a couple of lowest structural natural frequency signals obtained from vibration data over a long period of time. First, the proposed method utilizes a truncated transformation matrix of the robust local PCA of a portion of the dataset corresponding to the healthy state of the structure to remove the EOV effects by mapping the dataset to a new space. The difference between the mapped signals and the original signals is deemed to minimize the effect of the EOV. As such, extracting the mapped data from the original data, termed error signals, will remove the EOV effects and can be further used for damage detection. To this end, the Mahalanobis distances of the errors in the test set from the distribution of the errors in the baseline data are used for condition monitoring through constructing a Hotelling (T^2) control chart. The proposed PCA-based method does not apply the covariance matrix and the mean vector of the entire dataset, but instead the Minimum Covariance Determinant (MCD) algorithm, in its fast mode (FastMCD), is employed to obtain a robust covariance matrix and mean vector of the dataset. It is shown through solving the benchmark problem of the Z24 bridge that the proposed method can drastically increase the accuracy of the damage detection compared with the case when the normal PCA is used.

Keywords: SHM, Robust PCA, EOV, MCD, Hotteling charts

INTRODUCTION

Long-term condition monitoring of civil infrastructures can be classified based on the type of the employed data into two main categories: (1) output-only methods, and (2) input-output methods [1]. The inputs are usually either the applied forces or ambient data such as temperature and humidity, whereas the outputs are structural responses either in time or frequency domain such as acceleration, natural frequencies or mode shapes. The output-only methods have been given a great deal of attention lately, since they do not require multi-type sensors deployment on the structure under study to measure multi-type data. There have been different types of output-only methods proposed by researchers during the past decade. Some of the output-only methods for structural condition monitoring include: cointegration-based methods [2], principal component analysis-based methods [3], clustering methods [4], and outlier analysis-based methods [1].

Static and dynamic regression models complemented by a Principal Components Analysis (PCA) were employed for condition monitoring of the Infante D. Henrique Bridge using control charts [5]. There have been also some attempts to design and apply novel software to be used for long-term condition monitoring of structures [6]. Clustering, as a popular unsupervised technique, has been widely used for damage identification of structures. Some of the well-known clustering methods for structural damage detection include: k-means, k-medoids, Gaussian mixture model and fuzzy clustering [7, 4]. Likewise to the supervised methods, EOV can adversely affect the performance of unsupervised methods as well.

This paper aims to develop a supervised condition monitoring algorithm which is robust to the EOV effects. The data used to this end are a couple of lowest natural frequencies of structure, identified over a long period of time. Knowing the EOV can compromise the effectiveness of condition monitoring methods, the main challenge is to develop a technique that can deal with the unwanted effects of EOV on the structural natural frequencies. This point is addressed through exploiting an advanced signal decomposition method, termed Variational Mode Decomposition (VMD) algorithm coupled with a robust PCA algorithm. To validate the performance capability of the proposed method, the condition monitoring problem of the Z24 bridge benchmark problem is studied. The superiority of the proposed method is demonstrated through comparison against the case where the normal PCA algorithm is used. The Novelties of this study can be thus summarised as follows:

1. A novel robust PCA-based structural condition monitoring framework is proposed. The proposed method outperforms a traditional version of itself.
2. The VMD algorithm is employed for removing seasonal patterns in structural natural frequency signals.
3. The FastMCD algorithm can produce slightly different values for the mean and covariance matrix of the dataset at each run. This will make the results of condition monitoring vary from one run of the computer program to another. Therefore, it is proposed to run the algorithm 100 times and the average and median of the results be taken for structural condition monitoring.
4. The proposed method was successfully tested on the Z24 bridge benchmark problem.

BACKGROUND

The proposed method uses the Principal Component Analysis (PCA) of the structural natural frequency signals, identified over a long period of time, for condition monitoring of the structure. The proposed method is constructed based on a similar method presented in [5]. However, in this study, the robust scatter and location of the frequency signals are identified through employing a robust PCA algorithm and further used for damage identification. This strategy will be shown to be able to address the challenges imposed by the EOV effects on the structural condition monitoring problem. The proposed method is a baseline-dependent method, meaning that it requires baseline data from the healthy state of structure. Therefore, the proposed method has two main stages. The first stage considers constructing a baseline and obtaining a transformation matrix (to minimise the effect of the EOV on frequencies) based on data obtained from the healthy state of the structure. The obtained transformation matrix from the first stage is then used for conducting condition monitoring using data available from a secondary state of the structure.

It is known that the EOV effects mark mainly two different patterns on structural response data. These are: (1) a short-term seasonal pattern stemming from the short-term fluctuations of temperature or other EOV effects, and (2) a long-term pattern caused by the long-term fluctuations of such effects. Condition monitoring algorithms are adversely affected by the change of the variance of heteroscedastic data stemming from the seasonal effects, making the procedure of condition monitoring a complex task. Hence, the first stage of condition monitoring algorithms is usually aimed to remove complex seasonal patterns from structural response (natural frequencies in this paper) [8, 9, 2, 1]. To this end, an advanced signal decomposition algorithm, termed Variational Mode Decomposition (VMD) [10], is used in this study. The VMD algorithm is generally used to decompose a signal $S(t)$ into its constructive oscillatory modes, termed Intrinsic Mode Functions (IMF). Each IMF is narrow-band and can be thus characterised by its center frequency ω. The algorithm of the VMD solves the following constraint variational optimisation problem:

$$\min_{\{u_k\}\ \&\ \{\omega_k\}} \sum_k \left\| \partial_t \left(\delta(t) + \frac{j}{\pi t} \ * \ u_k(t) \right) e^{-j\omega_k t} \right\|_2^2 ; \ \ s.t. \ \ S(t) = \sum_k u_k(t) \tag{1}$$

in which $*$ is the convolution operator, j is the imaginary unit, u_k and ω_k are respectively the k^{th} IMF and its center frequency, $\delta(t)$ represents dirac distribution, and $||.||_2$ represnts L^2-norm.

In this study, VMD is employed for two main reasons [1, 2, 9]: (1) to denoise the natural frequency signals, and (2) to remove the seasonal patterns in the signals. In order to solve the constraint optimisation problem of (1) a Lagrangian multiplier and a quadratic penalty term are added to the equation. This makes the VMD a parametric decomposition algorithm. The specified values of the VMD parameters, for the purpose of this study, are listed in Table 1. For further information about the ways of specifying VMD parameters the readers are referred to [1, 2, 9].

The proposed method uses a robust PCA technique to obtain a transformation matrix out of a set of frequency signals corresponding to the healthy state of the structure under study. The obtained transformation matrix is then used in the second stage to map the frequency signals pertaining to an unknown state of the structure to a new space which can be further used to

Table 1: Specified VMD parameters.

Parameters	Description	Specified values
K	Number of IMFs	2
α	Denoising factor	100
τ	Time interval	0
ϵ	Convergence threshold	10^{-5}
$init$	Center frequency initialiser	0
DC	Boolean parameter	0

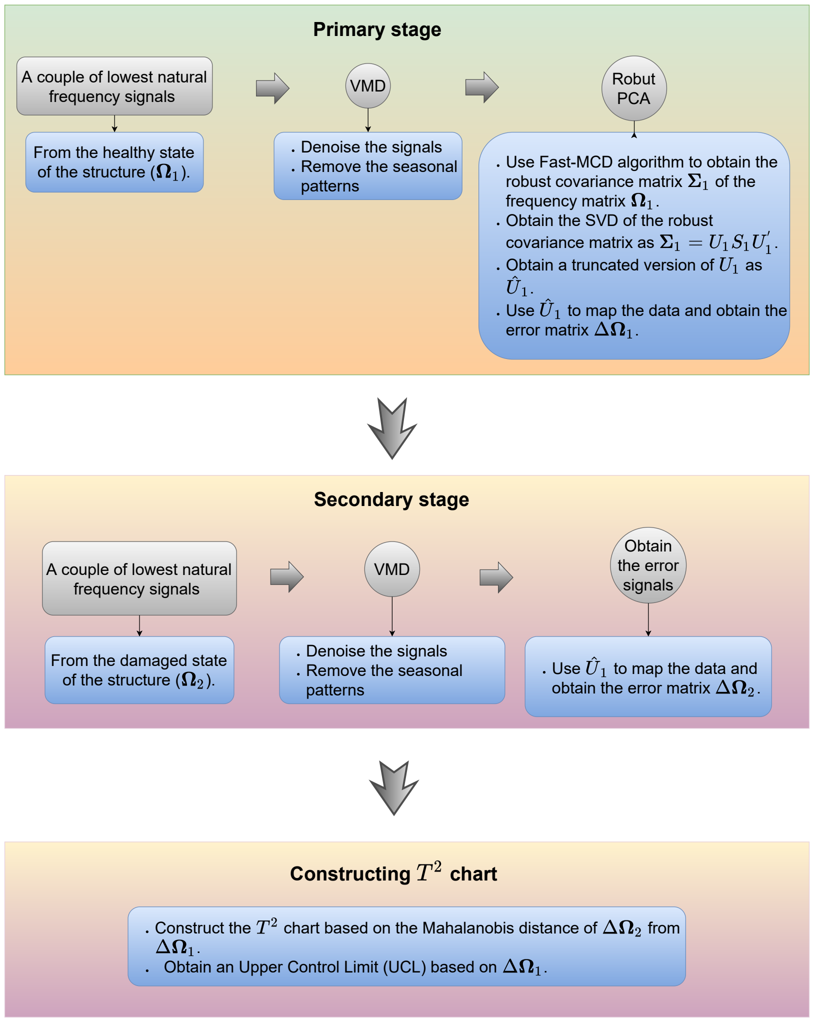


Figure 1: The scheme of the proposed method.

construct a damage sensitive feature (DSF). To this end, the fast Minimum Covariance Determinant (FastMCD) algorithm [11] is employed to derive the robust covariance matrix and mean vector pertaining to the matrix of all observations.

Figure 1 shows the steps of the proposed condition monitoring framework. It is assumed that some data (a couple of lowest natural frequencies) from the healthy state of the structure are available as baseline in the first step. As such, the first step is outlined as follows:

1. Identify a couple of lowest natural frequencies of the healthy structure from the vibration signals at some instants over a long period of time and set them in the matrix $\mathbf{\Omega}_1$. To this end, the stochastic subspace identification method is usually employed [12].

2. Employ the VMD algorithm to denoise the signals and remove the seasonal patterns from each of the columns of $\mathbf{\Omega}_1$.

3. Employ the FastMCD algorithm to obtain the robust covariance matrix of $\mathbf{\Omega}_1$ as Σ_1.
4. Obtain the Singular Value Decomposition (SVD) of Σ_1 as $\mathbf{\Sigma}_1 = \mathbf{U}_1\mathbf{S}_1\mathbf{U}'_1$, where $\mathbf{S}$ is the diagonal matrix of eigenvalues, $\mathbf{U}$ is the matrix of corresponding eigenvectors, and $'$ denotes the transpose of a matrix.
5. Obtain a truncated form of the transormation matrix $\mathbf{U}_1$, shown as $\hat{\mathbf{U}}_1$, through eliminating the columns corresponding to the principal components of higher orders. Note that in this paper, the first two columns corresponding to the PC_1 and PC_2 were retained.

The second step of the proposed method is outlined as follows:

1. Identify the very number of lowest natural frequencies of the structure, likewise to the first stage, from the recorded vibration data and set them in the matrix $\mathbf{\Omega}_2$.
2. Employ the VMD algorithm to denoise and remove the seasonal patterns of the columns of $\mathbf{\Omega}_2$.
3. Employ the truncated $\hat{\mathbf{U}}_1$, obtained from the first stage, to map the natural frequency signals of the second stage as follows:
$$\hat{\mathbf{\Omega}}_2 = \mathbf{\Omega}_2\ \hat{\mathbf{U}}_1\ \hat{\mathbf{U}}_1' \tag{2}$$
Likewise, the same procedure is followed for the first stage as follows:
$$\hat{\mathbf{\Omega}}_1 = \mathbf{\Omega}_1\ \hat{\mathbf{U}}_1\ \hat{\mathbf{U}}_1' \tag{3}$$
4. Consider the error associated with this mapping as a damage sensitive feature (DSF), as follows:
$$\Delta\mathbf{\Omega}_2 = |\hat{\mathbf{\Omega}}_2 - \mathbf{\Omega}_2| \tag{4}$$
Likewise,
$$\Delta\mathbf{\Omega}_1 = |\hat{\mathbf{\Omega}}_1 - \mathbf{\Omega}_1| \tag{5}$$

where $|.|$ is the absolute value operator.

Finally, a new signal needs to be constructed out of the mapped frequency matrix for damage identification. This is done through using the concept of Shewhart (T^2) charts as follows:

The Phase II Shewhart (T^2) chart is proposed to be employed for obtaining a signal out of the mapped frequency signals which is suitable for condition monitoring of structures [13, 14]. The T^2 chart, constructed for each individual observation in the second stage (vector $\Delta\mathbf{\Omega}_2$), is the second power of the Mahalanobis distance of that observation from the average point of the dataset corresponding to the first stage (vector $\Delta\mathbf{\Omega}_1$) as follows:

$$T^2 = (\Delta\mathbf{\Omega}_2 - \mu_1)'\ \mathbf{\Sigma}_1^{-1}\ (\Delta\mathbf{\Omega}_2 - \mu_1) \tag{6}$$

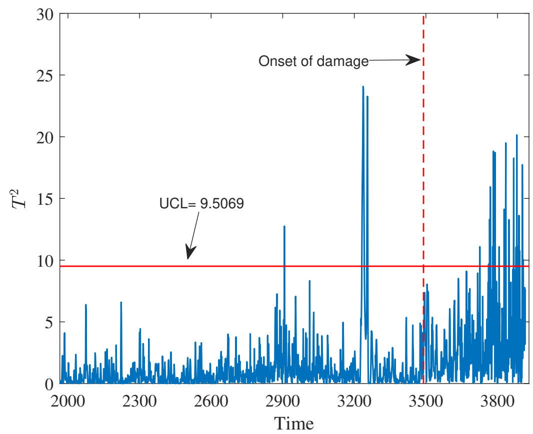
where μ_1 and $\mathbf{\Sigma}_1$ are respectively the mean vector and covariance matrix of $\Delta\mathbf{\Omega}_1$ obtained through the FastMCD algorithm. An Upper Control Limit (UCL) for the constructed chart can be obtained based on information corresponding to the first state of the structure as follows:

$$\mathrm{UCL} = \frac{p(m+1)(m-1)}{m^2 - mp} \times F_{\alpha;\, p,m-p} \tag{7}$$

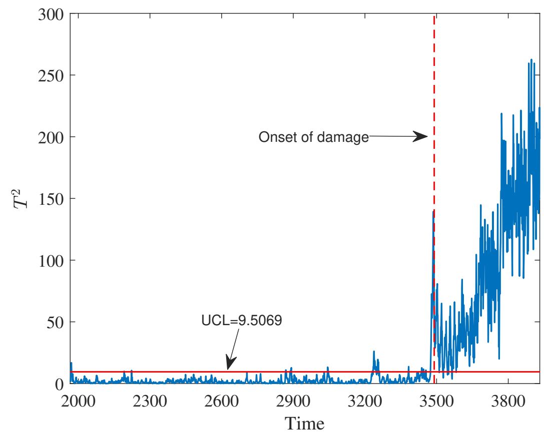
where p is the number of variables–identified natural frequencies at each time instant, $F_{\alpha;\, p,m-p}$ represents the value of the F-distribution with p and $m-p$ degrees of freedom, and α is the confidence level. To get a 95 % confidence for the UCL, α was set to 0.05, m is the number of observations in the first stage which corresponds to the healthy state of the structure.

ANALYSIS

In this section, the problem of condition monitoring of the Z24 bridge is studied. To this end, the four lowest natural frequency signals of the structure are analyzed using the proposed method for condition monitoring of the structure. The benchmark problem of the Z24 bridge has been used in many applications such as those aimed at removing the influence of the EOV effects on structural modal data [15, 16, 17]. The corresponding data of the structural natural frequencies can be found upon request from "`https://bwk.kuleuven.be/bwm/z24`".



(a) Normal PCA



(b) Robust PCA

Figure 2: Damage detection of the Z24 bridge using (a) normal, and (b) robust PCA.

First, the results of the application of the normal PCA are presented in Figure 2a. As is evident from the figure, the results of the condition monitoring is not satisfactory, as the numbers of false positive/negative cases are plentiful. Next, the robust PCA was employed in the proposed condition monitoring framework. To this end, the entire dataset was divided into two parts, where the first 50% portion of the dataset was used for obtaining the truncated matrix $\hat{\mathbf{U}}_1$, as discussed in the previous section. Note that the first 50% dataset thus corresponds to the healthy state of the structure. The remaining 50% portion of the dataset thus include information about damage. First, the effect of the EOV on the natural frequency signals was eliminated using the proposed PCA-based approach. Finally, Eq. 6 was used to obtain the T^2 chart of the second portion of the dataset, for which the UCL was obtained via Eq. 7. Note that the FastMCD algorithm can produce marginally different results in different attempt which can affect the final damage identification results. Here, a typical obtained results is shown in Figure 2b, though later on we propose a strategy to deal with this problem. As is evident from the results, the proposed method is far more successful in condition monitoring of the Z24 bridge. In order to quantify the results, the F_1 score was employed which is defined as follows:

$$F_1 = \frac{tp}{tp + \frac{1}{2}(fn + fp)} \tag{8}$$

where tp, fn, and fp denote respectively the number of true positive, false negative, and false positive cases in the obtained results. The F_1 score was respectively calculated for the normal and robust PCA-based methods results as 0.0996 and 0.9257, indicating the far better performance of the proposed condition monitoring framework. Note that in both cases VMD was employed for removing the seasonal patterns from the dataset. Reiterated, the only downside of the proposed strategy is that the results can vary in different runs of the FastMCD algorithm, making the final decision of the condition monitoring challenging. To confront this problem, it is recommended that the algorithm be ran multiple times and the average value as well as the median of the final results be taken. Here, the algorithm was ran 100 times. The mean and the median of the final results are presented in Figure 3.

Next, the F_1 scores as for the mean and median of the results were calculated. This value was obtained respectively for the mean and median of the results as 0.7915 and 0.8069. This truly again demonstrates the far better performance of the proposed strategy compared with the case when the normal PCA was employed (with $F_1 = 0.0996$) for condition monitoring of the Z24 bridge.

Note that although some false positive/negative cases can be identified in the final results of the both Figures 3a and 3b, one can ignore the short term violations of the threshold as they cannot be referred to damage/undamage cases when the results are not persistent. This is mainly due to the fact that damage is believed to have a stationary effect on the results and, therefore, its presence cannot appear in the final results as a short term violation of the threshold [5] (likewise to its absence). As such, one may devise a better strategy for quantifying the performance of the proposed condition monitoring strategy than the F_1 score. Since all such fp and fn cases were ruled in calculation of the F_1 score, it is actually a worst case evaluation criteria.

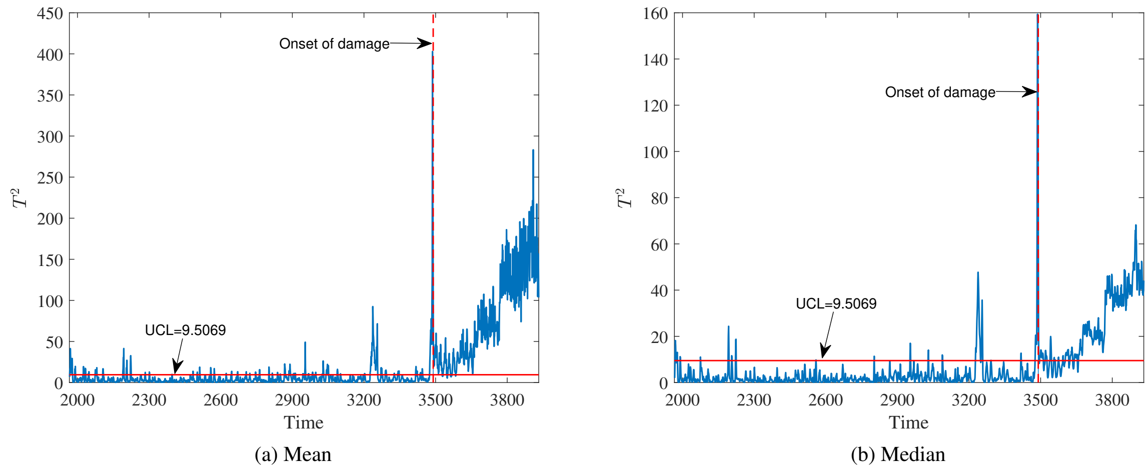


Figure 3: Damage detection of the Z24 bridge using the proposed robust PCA-method. The algorithm was ran 100 times and the mean ((a)) and median ((b)) of the final results were taken..

CONCLUSION

A robust PCA-based condition monitoring method for identification of anomalous observations of the natural frequencies that can be referred to damage has been proposed. The proposed method utilizes the FastMCD algorithm to obtain the robust covariance matrix of the healthy state of structure to obtain a transformation matrix to be further used for transforming data to the original space. Once the data from the secondary state of the structure was transformed to the new space, the error associated with this transformation was considered as a DSF. The Mahalanobis distances of the errors pertaining to the secondary state of the structure from the primary state (healthy sate) of the structure (T^2 chart) were monitored for condition monitoring of the structure. Since the results of the FastMCD algorithm can vary from one run of the algorithm to another, the FastMCD algorithm was ran 100 times and the average and median of the results were considered for condition monitoring of the structure. The results demonstrate the far better performance of the proposed method compared with the case when a normal PCA was employed in the calculations. The authors aim to improve the performance of the proposed condition monitoring strategy through the application of other methods for robust identification of the scatter and location of dataset.

Acknowledgement

The KU Leuven Structural Mechanics Section is acknowledged as the source of the data for Z24 bridge.

REFERENCES

[1] M. Mousavi and A. H. Gandomi, "Structural health monitoring under environmental and operational variations using mcd prediction error," *Journal of Sound and Vibration*, vol. 512, p. 116370, 2021.

[2] M. Mousavi and A. H. Gandomi, "Prediction error of johansen cointegration residuals for structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 160, p. 107847, 2021.

[3] 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," in *Proceedings of the 13th International Conference on Damage Assessment of Structures*, pp. 182–198, Springer, 2020.

[4] H. Sarmadi, A. Entezami, M. Salar, and C. De Michele, "Bridge health monitoring in environmental variability by new clustering and threshold estimation methods," *Journal of Civil Structural Health Monitoring*, pp. 1–16, 2021.

[5] F. Magalhães, Á. Cunha, and E. Caetano, "Vibration based structural health monitoring of an arch bridge: from automated oma to damage detection," *Mechanical Systems and Signal Processing*, vol. 28, pp. 212–228, 2012.

[6] E. García-Macías and F. Ubertini, "Mova/moss: Two integrated software solutions for comprehensive structural health monitoring of structures," *Mechanical Systems and Signal Processing*, vol. 143, p. 106830, 2020.

[7] R. de Almeida Cardoso, A. Cury, and F. Barbosa, "Automated real-time damage detection strategy using raw dynamic measurements," *Engineering Structures*, vol. 196, p. 109364, 2019.

[8] H. Shi, K. Worden, and E. J. Cross, "A cointegration approach for heteroscedastic data based on a time series decomposition: an application to structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 120, pp. 16–31, 2019.

[9] M. Mousavi and A. H. Gandomi, "Deep learning for structural health monitoring under environmental and operational variations," in *Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, Civil Infrastructure, and Transportation XV*, vol. 11592, p. 115920H, International Society for Optics and Photonics, 2021.

[10] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[11] M. Hubert, M. Debruyne, and P. J. Rousseeuw, "Minimum covariance determinant and extensions," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 10, no. 3, p. e1421, 2018.

[12] A. Cancelli, S. Laflamme, A. Alipour, S. Sritharan, and F. Ubertini, "Vibration-based damage localization and quantification in a pretensioned concrete girder using stochastic subspace identification and particle swarm model updating," *Structural Health Monitoring*, vol. 19, no. 2, pp. 587–605, 2020.

[13] R. Thomas, *Statistical Methods For Quality Improvement*. John Wiley & Sons, 2nd Ed., 2000.

[14] D. C. Montgomery, *Introduction to statistical quality control, Chapter 10*. John Wiley & Sons, 4th ed., 2001.

[15] 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.

[16] E. Reynders, G. Wursten, and G. De Roeck, "Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification," *Structural Health Monitoring*, vol. 13, no. 1, pp. 82–93, 2014.

[17] R. Langone, E. Reynders, S. Mehrkanoon, and J. A. Suykens, "Automated structural health monitoring based on adaptive kernel spectral clustering," *Mechanical Systems and Signal Processing*, vol. 90, pp. 64–78, 2017.