Structural Health Monitoring under Environmental and Operational Variations Using MCD Prediction Error

Mohsen Mousavi[a], Amir H. Gandomi [*a]

[a]*Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia*

Abstract

This paper proposes a novel technique that aims at detecting the effect of damage on structural frequency signals as "bad" outliers. To this end, a procedure is developed based on the Variational Mode Decomposition (VMD), Minimum Covariance Determinant (MCD), and Recurrent Neural Network (RNN) with Bi-directional Long-Short Term Memory (BiLSTM) cells. The VMD is first used in a pre-processing stage to denoise the signals and remove the seasonal patterns in them. Then, the proposed method seeks to learn the rules behind calculation of the Mahalanobis distances of the points from their distribution, using the parameters obtained from the MCD algorithm, through training an RNN on signals obtained from the inferior state of the structure (healthy state). It will be shown that, since the rule behind the effect of damage on the Mahalanobis distances has not been learnt by the trained RNN, the prediction errors of these values will increase significantly as soon as damage occurs using the data obtained from the posterior state of the structure (including damage). The performance of the proposed method is first tested on a numerical example and further validated through solving an experimental example of the Z24 bridge. Moreover, the proposed method is compared against a PCA-based method. The results demonstrate the superiority of the proposed method in long-term condition monitoring of civil infrastructures. The proposed method is an output-only condition monitoring method that requires only a couple of lowest structural natural frequency signals measured over a long-term monitoring of the structure. Therefore, it is recommended for cases when the measurements from the EOV are not available. Also the proposed method can be used along with other output-only or input-out methods to either improve or confirm the validity of their results.

Keywords: `Structural health monitoring, Minimum co-variance determinant, C-step algorithm, Variational mode decomposition, Recurrent neural netwrok`

Email address: `gandomi@uts.edu.au` (Amir H. Gandomi [*])

1. Introduction

The effect of the operational and environmental variations (EOV) on the mode shape and natural frequencies of structures is well understood [1]. As such, the feasibility of using natural frequency shifts for structural damage identification has been demonstrated in the literature [2]. These effects will usually cause false-positive or false-negative outcomes in the structural condition monitoring using structural modal data. Temperature is known to be the main meteorological parameter that can modify the structural modal properties [3]. As such, there is a strong long-term correlation between the temperature and variation of the natural frequencies [4, 5]. Therefore, the first step of structural condition monitoring is dedicated to obtaining the structural modal data from measured structural vibration data such as acceleration responses. To this end, the stochastic subspace identification method can be used in the first step [6].

Generally, the methods for long-term condition monitoring of civil infrastructures can be divided into two main categories which are: (1) input–output methods, and (2) output-only methods. The main assumption behind input–output methods is that measurements form both the EOV (input) and structural responses (output) are available. Therefore, these methods basically aim at establishing a map between these two datasets to monitor the prediction errors for damage. Input–output methods have been widely used for condition monitoring of civil infrastructures under EOV effects. For instance, in a study conducted by Peeters et. al, the authors established a map between a couple of lowest natural frequencies of the Z24 bridge and the environmental variables via ARX models [7]. The predicted natural frequencies were then monitored for raising an alarm as soon as damage occurs in the structure. Zhou et. all trained three different back-propagation neural networks (BPNN) based on taking mean temperatures, effective temperatures, and principal components (PCs) of temperatures as input to each model in order to further establish a map between the structural natural frequencies and environmental temperatures [8]. The authors used a $770\ h$ recorded dataset from the instrumented Ting Kau Bridge to study the effectiveness of their methods. In order to enhance the generalisability of the models, an early stopping approach was taken in the training process of the RNN models. The results of this study revealed that the developed BPNN model based on the input mean temperatures brings about satisfactory results. In another study, Kromanis et. al used the support vector regression (SVR) method to establish regression models for prediction of thermal responses of a bridge given the distributed temperature measurements as input [9]. The errors obtained from the comparison between the predicted strain data from the developed SVR models and measurements made on the simulated damaged bridge girder are then introduced as damage sensitive features. In a similar study, Reilly et. al explored the relationship between the tem-

perature measured on structure and the resultant strain and displacement through monitoring of the variation of the Coefficient of the Thermal Expansion (CTE) for the structural condition monitoring [10]. In a recent study, Ceravolo et. al sought the correlation between 10 years measurements from environmental data and static sensors regarding the Sanctuary of Vicoforte, a monumental Italian church, through time series analysis for condition monitoring [11].

Cointegration (CI), a data normalisation technique, is an output-only method that has been used for removing the EOV effects on the structural responses [12, 13, 14, 15]. CI is a technique adapted from econometrics that seeks to find a linear/non-linear combination of the feature signals (natural frequencies) which is stationary. Therefore, the resulting signal is deemed to be clear of any non-stationary effects stemming from EOV. Knowing that the effect of damage on the structural response is stationary, one can use the CI residuals to monitor structures for damage. There have been several techniques developed by researchers based on the concept of the CI to mitigate the effect of the temperature variation in condition monitoring of structures, such as damage detection of: aluminum plates using lamb waves [16], composite plates and composite sandwich panels using the non-linear vibro-acoustic modulation technique [14], steel plates using electro-mechanical impedance responses [15], a cable-stayed bridge using cable forces [17], and gearboxes using recorded vibration signals [18]. Although CI seems a promising technique for obtaining a stationary representation of frequency signals which does not encompass any effect of the EOV, one may still experience some technical challenges using it for SHM. For instance, while the measured responses might show a non-stationary behavior in a couple of days, they can be stationary over a long period of time (e.g. a couple of months) [13]. Therefore, using a fraction of the data recorded over a long period of time for training CI may contradict the basic unit root assumption of such a technique. Notwithstanding, it has been argued that the philosophical question of whether a unit root assumption is valid can simply be ignored from the engineering application point of view [13]. There have been, however, other output-only methods used to mitigate the effect of the EOV on vibration signals that also deserve some attentions such as principal component analysis [19, 20], time series models [21, 22], and outlier analysis [23, 24, 25]. For example, Magalhaes et. al [26] tested alternative static and dynamic regression models complemented by a principal components analysis for damage identification using control charts. State-of-the-art sensing technologies are also capable of monitoring structures at the presence of EOV. As an example, Garcia Macias and Ulbertini [27] developed two novel software solutions for long-term condition monitoring of structures.

A quick comparison between input–output and output-only methods reveals that the former require multiple types of sensors to be deployed on different locations on structure to measure

the EOV and the structural responses data. In contrast, the output-only methods rely only on the structural response data for condition monitoring of the structure. Since these methods do not require any information from the EOV, they are less demanding in terms of the types of sensors to be used for performing multi-type measurements. However, developing such methods require an extensive study of the structural responses which is the main objective of this paper.

Consider a couple of lowest natural frequencies of a structure measured at each time instant over a long period of time. Reiterated, it is known that the natural frequencies are modified by the EOV. One can work out the multivariate location μ and scatter of such data using a highly robust estimator termed as Minimum Covariance Determinant (MCD) algorithm [28]. Next, the distance of each point from the distribution of the data can be measured by the Mahalanobis distance measure. We propose to develop a procedure that can learn the underlying rule for obtaining these distances for the healthy structure, i.e. the distance of the multivariate feature vector (natural frequencies) from its representative distribution. Now consider the case when the natural frequencies of the structure are modified by damage. Since this phenomenon has not been seen by the proposed procedure before, therefore, one may expect that any changes due to the damage will appear as outliers. However, these outliers might not be easily detected due to the EOV effects. Let us classify the detected outliers as "good" (undamaged case) and "bad" (damaged case) outliers. Therefore, some rules are required to be set to distinguish between the good and bad outliers. According to Hawkins an outlier "is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism" [29]. We adapt the Hawkins's definition of outliers as "bad" outliers in our own proposed strategy. Now, the main question is how to distinguish between bad and good outliers.

We propose to train a Recurrent Neural Network (RNN) with Bi-directional Long-Short-Term Memory (BiLSTM) cells that can learn the underlying rule of determining the Mahalanobis distance of each point from its distribution when the structure is undamaged. Note that, in this stage, the effect of EOV on the data still exists meaning that there are still some good outliers in the calculated Mahalanobis distances. Next, the trained RNN is used to estimate the Mahalanobis distances of the data obtained from a secondary state of the structure from their distribution. Using the trained RNN alone to work out the Mahalanobis distances, one still may not be able to detect the "bad" outliers clearly. To do so, the Mahalanobis distances of the data in the second stage are also calculated from their true distribution. Now, one needs only to monitor the prediction errors by subtracting the predicted from the real values of the Mahalanobis distances. According to Hawkins, while the underlying rule of calculating the Mahalanobis distances at the presence of EOV is learnt, the trained RNN is not familiar with

the secondary effect of damage on the Mahalanobis distances. As such, One may expect that the prediction errors increase significantly from the average error. To this end, we propose the data to be treated in a pre-processing stage as follows.

Generally, EOV effects can be divided into two parts: (1) the short-run stationary seasonal pattern, (2) the nonstationary long-run pattern.[1] As such, damage detection algorithms can suffer from the change of the variance of heteroscedastic data stemming from the former. Therefore, as a signal pre-processing stage, it is recommended that the complex seasonal pattern to be removed from the frequency signals [30]. Therefore, we propose an advanced signal decomposition technique termed as Variational Mode Decomposition (VMD) to be used to this end [31]. As such, VMD is used for both denoising and decomposing the frequency signals. The mode corresponding to the seasonal pattern will be, therefore, omitted from the signals. This will further facilitate the process of training the RNN to learn the underlying rule of calculating Mahalanobis distances. Therefore, the organisation of this paper follows:

In section 2, the basic theory of MCD algorithm is discussed in two parts, namely 1) basic definitions and properties (Section 2.1), and 2) C-step algorithm which is used for computation of MCD parameters (Section 2.2). In Section 3, the proposed damage detection algorithm is presented. As such, the scheme of the proposed method is presented in Section 3.1 and a threshold is obtained for the proposed damage sensitive feature in Section 3.2. Next, the basic theory of the VMD algorithm is discussed, in Section 3.3. In section 3.4, the basic theory of RNN and LSTM/BiLSTM cells is explained. In Section 3.5, a PCA based method, proposed in [26], is presented that is further used for comparing its results with the results obtained from the proposed method in both of the numerical and experimental sections. In order to illustrate the procedure of the proposed algorithm and to walk the readers through the steps, a numerical example is studied in Section 4. In Section 5, the applicability of the proposed method to real world problems is examined through condition monitoring of the Z24 bridge. Finally, Section 6 presents some conclusion remarks drawn from the obtained results and some idea for future work is presented.

2. Minimum co-variance determinant

2.1. Basic definition and properties

Multivariate location and scatter setting is presented as an $n \times p$ data matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)'$ where $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})$ represents the ith observation. As such, n and p indicate respectively

---

[1]Note that the second part might also show stationary trend if the data are recorded over a sufficiently long period of time [13].

the number of observations and features. It is assumed that the observations can be represented by an elliptically symmetric and unimodal distribution with unknown parameters μ and $\mathbf{\Sigma}$, where μ and $\mathbf{\Sigma}$ are respectively a vector with p components and a positive definite $p \times p$ matrix. Therefore, according to the definition, there exists a strictly decreasing real function g such that

$$f(\mathbf{x}) = \frac{1}{\sqrt{|\mathbf{\Sigma}|}} \, g\left(d^2\left(\mathbf{x}, \mu, \mathbf{\Sigma}\right)\right) \tag{1}$$

where the *statistical distance*[2] $d(\mathbf{x}, \mu, \mathbf{\Sigma})$ is calculated as

$$d\left(\mathbf{x}, \mu, \mathbf{\Sigma}\right) = \sqrt{(\mathbf{x} - \mu)'\mathbf{\Sigma}^{-1}(\mathbf{x} - \mu)} \tag{2}$$

As such, for n observations in $\mathbb{R}^p$ the raw MCD algorithm with the tuning parameter $n/2 \leq h \leq n$ obtains the pair $(\hat{\mu}_0, \hat{\mathbf{\Sigma}}_0)$ so that [32],

1. the point $\hat{\mu}_0$ indicates the mean of h observations for which the determinant of the co-variance matrix is minimum among all the other possible number of selections.

2. $\hat{\mathbf{\Sigma}}_0$ is the corresponding co-variance matrix (scatter matrix) multiplied by a consistency factor c_0.

3. The consistency factor c_0, to make the estimator consistent at the normal model, is obtained as $\alpha/F_{\chi^2_{p+2}}(q_\alpha)$ where $\alpha = \lim_{n \to \infty} h(n)/n$ and q_α is α-quantile of the χ^2_p distribution [33]. However, for making the estimator unbiased for small datasets, Pison et al. proposed that a finite-sample correction factor to be incorporated in equations [34].

4. h has to be chosen in a way that the determinant of the co-variance matrix be non-zero, i.e. $h > 2p$. This yields $n \geq 2p$, i.e. the number of the observations has to be at least twofold the number of dimensions. To avoid excessive noise, it is recommended that $n \geq 5p$, i.e. to have at least five observations per dimension. In higher dimensions, one can use minimum regularised covariance determinant as proposed in [35].

5. The MCD estimator is *affine equivariant*. In other words, for any non-singular matrix $\mathbf{A}_{p \times p}$ and vector $\mathbf{b}_{p \times 1}$ the followings hold,

$$\hat{\mu}_{\mathrm{MCD}}\left(\mathbf{X}\mathbf{A}' + \mathbf{1}_n\mathbf{b}'\right) = \hat{\mu}_{\mathrm{MCD}}\left(\mathbf{X}\right)\mathbf{A}' + \mathbf{b} \tag{3}$$

$$\hat{\mathbf{\Sigma}}_{\mathrm{MCD}}\left(\mathbf{X}\mathbf{A}' + \mathbf{1}_n\mathbf{b}'\right) = \mathbf{A}\hat{\mathbf{\Sigma}}_{\mathrm{MCD}}\left(\mathbf{X}\right)\mathbf{A}' \tag{4}$$

in which $\mathbf{1}_n$ is a column vector of n ones, and $\mathbf{X}$ represents $n \times p$ matrix of observations.

---

[2]Mahalanobis distance

Generally, for n observations, there are $\binom{n}{h}$ subsets of size h. Therefore, the process of finding a subset with minimum covariance determinant becomes time consuming. The alternative, is to use the FastMCD algorithm [36]. The FastMCD algorithm relies upon a procedure termed as C-step[3] which is based on the following theorem,

Theorem 1: *Let's $\mathbf{H}_{1h \times p}$ be a subset of observation matrix $\mathbf{X}_{n \times p}$ with $\hat{\mu}_1$ and $\hat{\boldsymbol{\Sigma}}_1$ as its empirical mean and covariance matrices. If $\left|\hat{\boldsymbol{\Sigma}}_1\right| \neq 0$, then the relative distances is obtained as $d_1(i) := d\left(\mathbf{x}_i, \hat{\mu}_1, \hat{\boldsymbol{\Sigma}}_1\right)$ for $i = 1, \ldots, n$. Reorder the obtained distances as $(d_1)_{1:n} \leq (d_1)_{2:n} \leq \cdots \leq (d_1)_{n:n}$ and choose $\mathbf{H}_2$ such that $\{d_1(i); i \in \mathbf{H}_2\} := \{(d_1)_{1:n}, \ldots, (d_1)_{h:n}\}$. Obtaining $\hat{\mu}_2$ and $\hat{\boldsymbol{\Sigma}}_2$ based on $\mathbf{H}_2$, then we have $\left|\hat{\boldsymbol{\Sigma}}_2\right| \leq \left|\hat{\boldsymbol{\Sigma}}_1\right|$ with equality if and only if $\hat{\mu}_2 = \hat{\mu}_1$ and $\hat{\boldsymbol{\Sigma}}_2 = \hat{\boldsymbol{\Sigma}}_1$.*

Following propositions can be directly concluded from Theorem 1,

1. The C-step yields a new h-subset with lower co-variance determinant if $\left|\hat{\boldsymbol{\Sigma}}_1\right| > 0$.
2. $\hat{\boldsymbol{\Sigma}}_2$ is more concentrated than $\hat{\boldsymbol{\Sigma}}_1$, i.e. has a lower determinant.
3. In case $\hat{\boldsymbol{\Sigma}}_1 = 0$, the minimal objective value is already achieved.
4. C-step can be iterated until the criteria $\left|\hat{\boldsymbol{\Sigma}}_{\text{new}}\right| = \left|\hat{\boldsymbol{\Sigma}}_{\text{old}}\right|$ is met.
5. Since the number of h-subsets is finite the C-step algorithm converges eventually. In practice, the convergence is achieved quickly.
6. Despite the previous point, there is no guarantee that the final value $\left|\hat{\boldsymbol{\Sigma}}_{\text{new}}\right|$ is the global minimum of the MCD objective function. This enforces to take different choices of $\hat{\mathbf{H}}_1$ and apply C-step to each one. The one that bears the minimum determinant may be taken as the solution.

Figure 1 depicts the flowchart of the C-step algorithm.

Running two C-steps on initial subsets usually leads to the choices that have rather small determinants. As such, the 10 subsets with lowest determinants are usually selected and further C-steps are run on all of them. Although this procedure is very fast for small sample size n, the computation time can increase when n is large. In such a case, FastMCD partitions the dataset to avoid doing all repetitive calculations on the entire dataset.

3. The proposed damage detection method

3.1. The proposed condition monitoring scheme

Figure 2 shows the flowchart of the proposed damage detection strategy. First, a portion of the frequency signals, corresponding to the healthy state of the structure is decomposed into

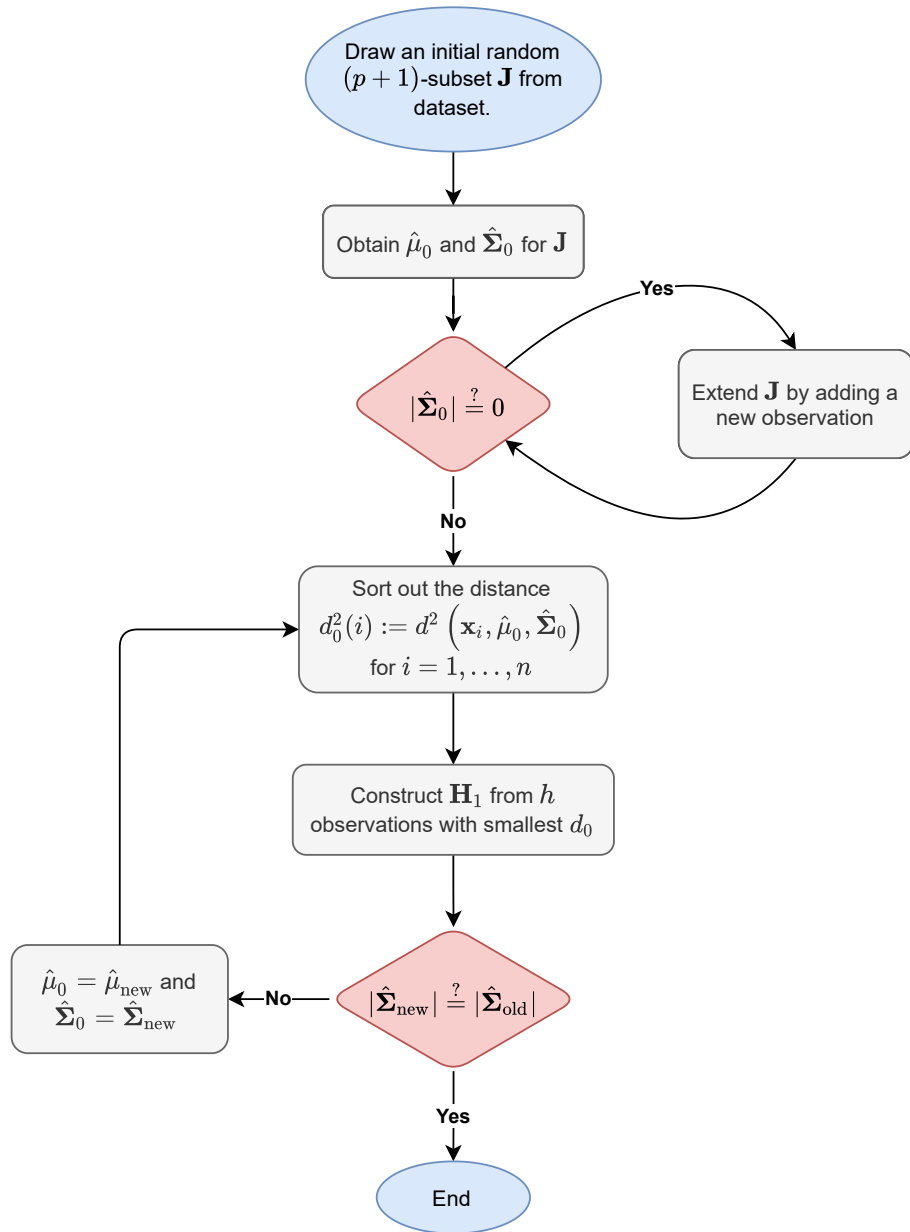---

[3]C stands for concentration

Figure 1: Flowchart of C-step algorithm.

its constituent parts using VMD. Accordingly, VMD is used for denoising and removing the seasonal pattern in the signals.

The output of the VMD is then fed into the FastMCD algorithm to obtain the unknown parameters μ_1 and Σ_1. These parameters are further used in (2) to obtain the Mahalanobis distance corresponding to each data-point of VMD output from the location μ_1.

Next, a Recurrent Neural Network (RNN) with Bi-directional Long-Short Term Memory (BiLSTM) cells is trained to learn the underlying rule of obtaining the Mahalanobis distances of the data obtained from the healthy structure from their distribution. As such, the output of the VMD and the obtained Mahalanobis distances are used respectively as features and targets to train the RNN.

Once the underlying rule is learned, the trained network is used to obtain the Mahalanobis distances of the data obtained from the secondary state of the structure from its distribution. To that end, the frequency signals from the secondary state of the structure are decomposed using VMD for denoising and removing the seasonal pattern, similar to the first stage. The VMD results are used in FastMCD algorithmm to obtain μ_2 and Σ_2 corresponding to the secondary state of the structure which are further used to obtain the Mahalanobis distances of the VMD outcomes in this stage from their true distribution. Next, the outputs of the VMD are also used as features in the trained RNN from the first stage to obtain the predicted Mahalanobis distances which are further compared against those obtained from the FastMCD algorithm. We hypothesise that the prediction errors will increase significantly as damage occurs. In this paper, the data from the both states are mixed and unique μ and Σ are obtained after the application of the VMD to work out the Mahalanobis distances of the whole dataset as target values

3.2. Setting an upper bond threshold

The obtained error signal can be considered as an R-chart constructed based on the difference between the obtained true Mahalanobis distances and the predicted ones. Therefore, one can set a threshold for the chart to obtain a confidence interval for the errors, so an alarm be raised as soon as the threshold is violated. To this end, first, the entire dataset is divided into three subsets which are: (1) the training set (50% of the entire dataset), (2) validation set (25% of the entire dataset), and (3) the test set (25% of the entire dataset).[4] Reiterated, the obtained plot of the errors is an R-chart for which one can obtain an upper control limit (UCL). Following

---

[4]Note that the validation set needs also to be from the healthy state of the structure. Regarding the examples of this paper, damage only appears in the test set.
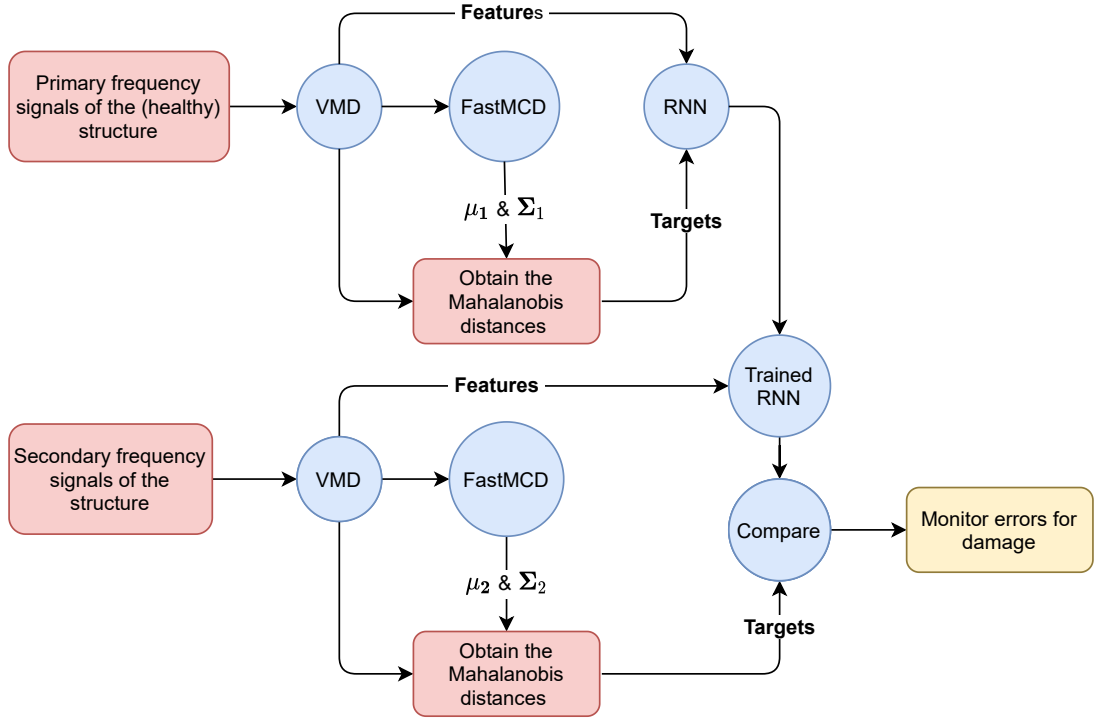
Figure 2: Flowchart of the proposed damage detection strategy. In this paper, unique μ and $\mathbf{\Sigma}$ are obtained from the mixed dataset after the application of the VMD to be further used to obtain the target values.

equation is therefore used to this end,

$$\text{UCL} = D_4 \, \delta \bar{d}_{\mathrm{v}} \tag{5}$$

where $\delta \bar{d}_{\mathrm{v}}$ is the mean value of the errors regarding the obtained Mahalanobis distances of the data-points in the validation set and $D_4 = 3.267$, a constant selected from a reference table, depends on the number of variables, i.e. 2 in here [37].

3.3. Variational Mode Decomposition (VMD)

VMD is a parametric signal decomposition algorithm which may be used to decompose a non-linear non-stationary signal into k oscillatory modes known as Intrinsic Mode Functions (IMF) [38] where each IMF is characterised by its center frequency ω. VMD can be also used for denoising the signals, concurrently. It solves the following optimisation problem,

$$\min_{\mathbf{u}_k \,\&\, \omega_k} \sum_k \left\| \partial_t \left(\delta(t) + \frac{\mathrm{j}}{\pi t} * u_k(t) \right) e^{-\mathrm{j}\omega_k t} \right\|^2 \tag{6}$$

where $*$ and j represent the convolution operator and the imaginary unit, respectively. $\mathbf{u_k}$ is the matrix of all IMFs where their corresponding center frequencies are listed in the vector ω_k. The solution to the minimisation problem of (6) is the saddle point of the augmented Lagrangian in a sequence of iterative sub-optimisations called alternate direction method of multipliers

(ADMM). This makes the VMD a parametric decomposition algorithm and, therefore, there are some parameters that need to be specified prior to the decomposition. Reiterated, there are two main reasons for using VMD to decompose the frequency signals in this paper which are:

1. Denoising the frequency signals.

2. Removing the IMF corresponding to the seasonal variations in the frequency signals.

To target the above goals, the parameters of VMD have been specified as follows,

1. $k = 2$: to set the number of IMFs into which the original signals are to be decomposed, i.e. 2.

2. α: to set the quadratic penalty term which is a denoising factor the value of which is selected 100 for the numerical examples and 10 for the experimental example of this paper. Note that the larger value of α admits less noise into the decomposition process.

3. $\tau = 0$: to choose the denoising factor α to be in effect. Any non-zero value of τ, say 0.1, will enforce the perfect reconstruction, i.e. denoising does not happen and, therefore, the value of α becomes irrelevant.

4. $\epsilon = 10^{-7}$: to set the tolerance parameter which controls the convergence of the algorithm.

5. $init = 0$: to initialise the center frequency of IMFs ω at zero. The alternatives are: 1 which indicates uniform initialisation, and 2 representing random initialisation. It has been argued, however, that the way of initialising IMF center frequencies has a little effect on the decomposition results [39].

6. DC=1: to determine the first IMF to be kept at zero center frequency. The alternative is to set DC at 0 which indicates the center frequency of the first IMF can be non-zero. Setting DC=1 will preserve the non-stationary behaviour of the first mode which is preferred in this work. However, it has been noted that selecting DC=0 has little effects on the results.

3.4. Training a Recurrent Neural Network

Using Long-Short Term Memory (LSTM) cells in Recurrent Neural Networks (RNN) have been widely used as an effective technique for the forecasting of time series [40, 41]. Therefore, an RNN with LSTM cells is trained to learn the rule behind calculation of the Mahalanobis distances of the data-points, obtained from the healthy structure, from their distribution. To this end, 50% of the VMD decomposition outcomes is used as the training features. The corresponding Mahalanobis distances obtained from the FastMCD algorithm are used as training targets. The remaining 50% of the frequency signals is used for damage detection as discussed in Section 3.

To keep the paper self-contained, a brief background theory of LSTM cells is explained in the paragraphs to follow. Figure 3 depicts a LSTM unit. There are three gates in the LSTM unit
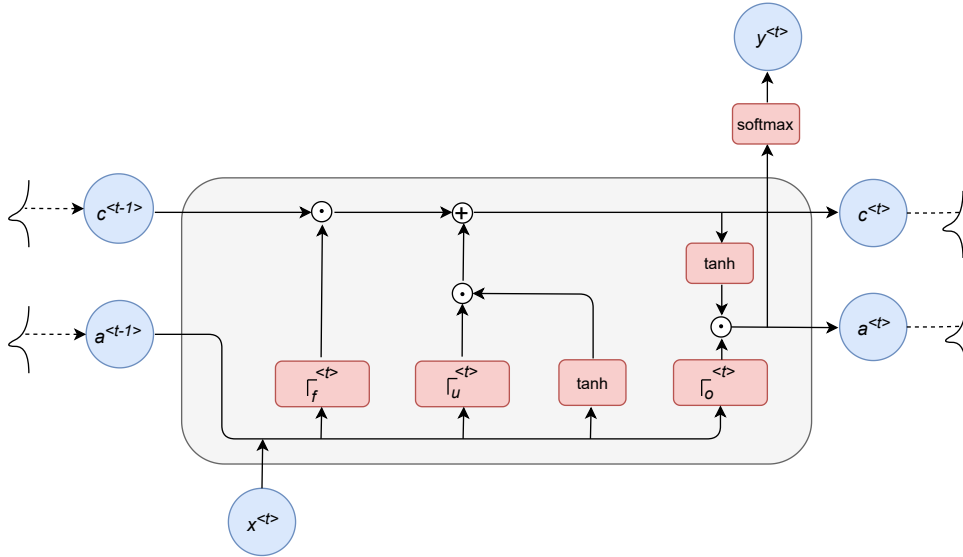
Figure 3: Visualisation of an LSTM unit. $y^{<t>}$ is the final output of the unit at time t which is computed by the softmax activation function.

including update, forget, and output gates. There are also three cells, namely input, memory, and update cells. As such, the candidate value $\tilde{c}^{<t>}$ for updating the memory cell at time t is calculated using the output value of the unit at time $t-1$, $a^{<t-1>}$, and the input value at time t, $x^{<t>}$, using the following equation,

$$\tilde{c}^{<t>} = \tanh\left(\mathbf{W}_c\left[a^{<t-1>}, x^{<t>}\right] + \mathbf{b}_c\right) \tag{7}$$

in which tanh(.) indicates the hyperbolic tangent activation function. $\mathbf{W}_c$ and $\mathbf{b}_c$ represent respectively the matrix of parameters and the biased vector of the memory cell. The value of the memory cell $c^{<t>}$ is then updated using the candidate value $\tilde{c}^{<t>}$ and the cell's previous value $c^{<t-1>}$ as follows,

$$c^{<t>} = \Gamma_u \odot \tilde{c}^{<t>} + \Gamma_f \odot c^{<t-1>} \tag{8}$$

where

$$\Gamma_u = \sigma\left(\mathbf{W}_u\left[a^{<t-1>}, x^{<t>}\right] + \mathbf{b}_u\right) \tag{9}$$

and

$$\Gamma_f = \sigma\left(\mathbf{W}_f\left[a^{<t-1>}, x^{<t>}\right] + \mathbf{b}_f\right) \tag{10}$$

where $\odot$ indicates element-wise product of two vectors or matrices, Γ_u and Γ_f are the values of the update and forget gates, and σ (.) is the sigmoid activation function. $\mathbf{W}_u$ and $\mathbf{b}_u$ indicate the matrix of parameters and the bias vector corresponding to the update gate, and $\mathbf{W}_f$ and $\mathbf{b}_f$ are their counterparts corresponding to the forget gate, respectively.

Finally, the output value of the LSTM unit at time t is calculated as

$$a^{<t>} = \Gamma_o \odot \tanh\left(c^{<t>}\right) \tag{11}$$

12

where

$$\Gamma_o = \sigma \left(\mathbf{W}_o \left[a^{<t-1>}, x^{<t>} \right] + \mathbf{b}_o \right) \tag{12}$$

in which Γ_o is the value of the output gate, and $\mathbf{W}_o$ and $\mathbf{b}_o$ are respectively the matrix of parameters and bias vector corresponding to the output gate.

A bidirectional LSTM (BiLSTM) layer is designed to learn bidirectional long-term dependencies between time steps of time series through additional training where the input data are traversed twice, i.e. from (1) left-to-right, and (2) right-to-left) [42]. These dependencies can facilitate learning from the complete time series at each instant in some problems. It has been noted that incorporating a BiLSTM layer into the architecture of the RNN leads to the betterment of the results regarding the problems of this paper.

3.5. A method for comparison

3.5.1. A PCA-based method

A method presented in [26] is adapted in here to compare its results with the ones obtained from the proposed method of this paper. Note that the original method is more detailed as the main approach is based on regression models complemented by the presented PCA-based method in here. However, since developing such regression models requires measurements from EOV data (such as temperature variations), it is in fact an input-output method. As such, we only use the PCA-based approach to make it output-only method to be further used for comparison with the method of this paper. The main reason to select this method for comparison lies in the fact that presented PCA-based method uses the covariance matrix of the dataset obtained through a normal procedure other than the FastMCD algorithm.

The method as such is based on the principal component analysis of the matrix of features (frequencies) $\mathbf{X}$. To this end, $\mathbf{X}$ is divided into the baseline $\mathbf{X}_b$ and the test set $\mathbf{X}_t$. The steps of the algorithm of the method follows,

1. Obtain the covariance matrix of the matrix of the features corresponding to the baseline $\mathbf{X}_b$ (healthy state of the structure) as $\mathbf{\Sigma}_b$. Note that the covariance matrix $\mathbf{\Sigma}_b$ in here is not obtained from the FastMCD algorithm and, therefore, corresponds to the whole feature space regarding the baseline.

2. Obtain the Singular Value Decomposition (SVD) of $\mathbf{\Sigma}_b$ as $\mathbf{\Sigma}_b = \mathbf{U}_b \, \mathbf{S}_b \, \mathbf{U}'_b$.

3. Take the transformation matrix $\mathbf{T}_b = \mathbf{U}_b$ and remove the columns of $\mathbf{U}_b$ that show the least variability in the dataset. This is done by looking at the variances listed at the main diagonal of $\mathbf{S}_b$ to work out how many columns should be retained in this stage. The outcome of this stage is the truncated transformation matrix $\hat{\mathbf{U}}_b$.

4. The truncated transformation matrix $\hat{\mathbf{U}}_b$ obtained from the baseline is then used to map the dataset corresponding to the test set $\mathbf{X}_t$ into a new space as,

$$\hat{\mathbf{X}}_t = \mathbf{X}_t \; \hat{\mathbf{U}}_b \; \hat{\mathbf{U}}_b' \tag{13}$$

5. Work out the matrix of errors $\epsilon_t = |\mathbf{X}_t - \hat{\mathbf{X}}_t|$ that does not include the EOV effect. $|.|$ takes the absolute value of all entries of a matrix.

3.5.2. Multivariate T^2 control charts

The Phase II Shewhart (T^2) chart with individual observations is used for performing condition monitoring of structure after removing the EOV effects using the PCA-based method [37, 43]. The T^2 chart, constructed for each individual observation in the test set, is the second power of the Mahalanobis distance of the observation from the average point of the validation set, corresponding to the healthy state of the structure, as follows,

$$T^2 = \left(\mathbf{X} - \bar{\mathbf{X}}_v\right)' \; \mathbf{\Sigma}_v^{-1} \; \left(\mathbf{X} - \bar{\mathbf{X}}_v\right) \tag{14}$$

where $\bar{\mathbf{X}}_v$ and $\mathbf{\Sigma}_v$ are respectively the average value and covariance matrix of the validation set. The UCL is obtained through the following formula,

$$\text{UCL} = \frac{p(m+1)(m-1)}{m^2 - mp} \times F_{p,m-p}(1 - \alpha/2) \tag{15}$$

Where p denotes the number of variables. $F_{p,m-p}(1-\alpha/2)$ denotes the value of the F-distribution with degrees of freedom p and $m - p$ evaluated at the point $1 - \alpha/2$. α is set at 0.05 to get 95 % confidence interval for the UCL. The procedure of specifying the UCL is based on the validation set with, m observations, which corresponds to the healthy state of the structure regarding the examples of this paper.

A numerical example is first studied in the next section using the proposed method for demonstration purpose. Next, the proposed strategy is tested on an experimental example of the Z24 bridge to verify the effectiveness of the proposed strategy in solving real-world problems. Moreover, the results of the proposed method are compared against the presented PCA-based method in both of the numerical and experimental sections.

4. Illustrative numerical example

The numerical example studied in this paper is a 4-dof spring-mass system as depicted in Figure 4. A similar numerical example was also studied by others (e.g. see [44]). The weight of the masses is 2 kg each and the stiffness of the springs, varying by temperature, is obtained as
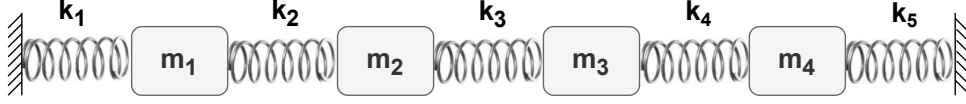
Figure 4: A 4-dof spring-mass system.

follows,

$$k_i = \begin{cases} -0.11 \times T + 4, & \text{if } T < 0 \\ -0.03 \times T + 4, & \text{if } T \geq 0 \end{cases} \tag{16}$$

for $i = 1, 2, 4, 5$, and

$$k_3 = \begin{cases} -0.11 \times T + 5, & \text{if } T < 0 \\ -0.2 \times T + 5, & \text{if } T \geq 0 \end{cases} \tag{17}$$

in kN/m where the different behaviour of k_3 introduces some nonlinear effects to the vibration modes. The right-left plot of Figure 5 shows a -15 ^oC shifted 10000 hourly recorded temperature signal in Basel-Switzerland from June 2019 to July 2020 which is used to calculate the stiffness of the springs hourly in the same period of time [45]. Note that the reason for shifting temperatures is to provide a wider range of negative and positive temperatures.

In order to simulate damage, the stiffness of k_3 is reduced by 10% at 8000th record and is deemed to endure until the end of the records, i.e. 10000th record.

The generalised eigenvalue problem of $\left|[\mathbf{M}] - \lambda^2[\mathbf{K}]\right| = 0$ is used to work out the natural frequencies (λ) of the system where $[\mathbf{K}]$ and $[\mathbf{M}]$ are the generalised stiffness and mass matrices of the system.

It is assumed that the natural frequencies of the system are contaminated by 10% noise. The following equation is used to contaminate the obtained frequency signals by noise, [46],

$$\hat{\delta} = \delta + \frac{\kappa}{100} n_{\text{noise}} \ \sigma(\delta), \tag{18}$$

in which δ and $\hat{\delta}$ represent respectively the noise-free and noisy calculated frequency times series, $\sigma(\delta)$ denotes the standard deviation of δ, κ is the noise level ($= 10$), and n_{noise} represents a random independent variables vector of the same length as δ sampled from a standard normal distribution.

The left-right plot of Figure 5 shows the obtained four natural frequency time series contaminated by 10% noise. The next step is to denoise the frequency signals as well as to remove any seasonal patterns in them. To this end, VMD is used.

The four natural frequency time series of the spring-mass system have been decomposed into two modes: a DC (IMF_1) and a seasonal mode (IMF_2). Reiterated, in order to deal with noise in the frequency signals, the quadratic penalty term was set at $\alpha = 100$. The decomposition
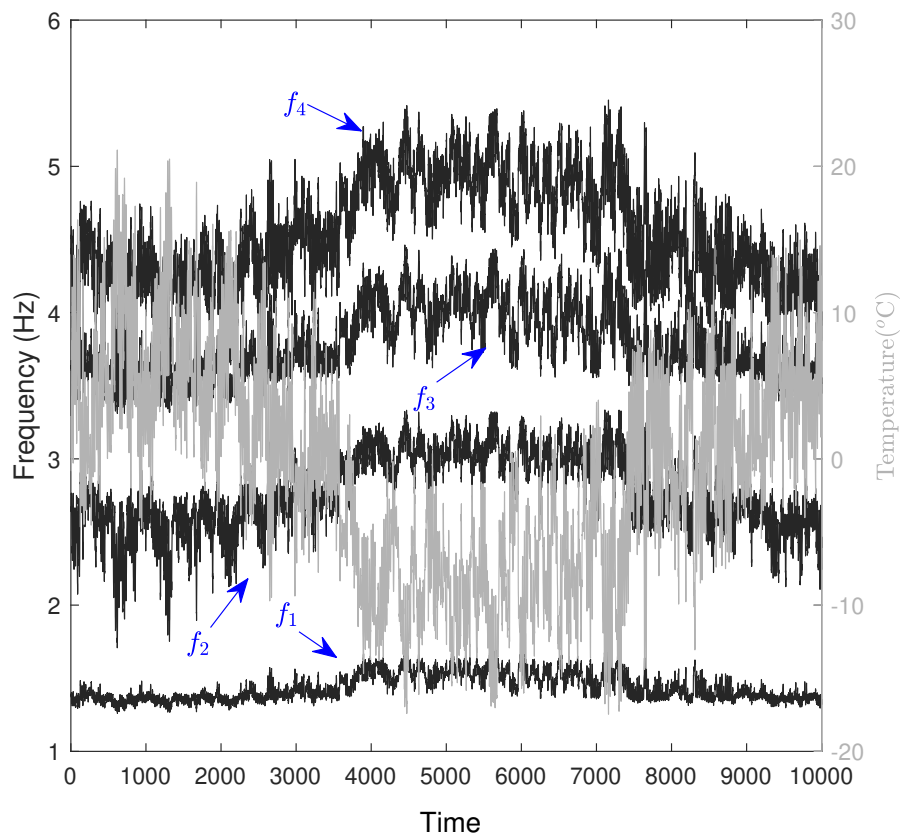
15

Figure 5: Hourly recorded temperature of Basel-Switzerland shifted $-15\ ^oC$ (right axis), and calculated natural frequencies of the spring-mass system for the time period from June 2019 to July 2020 (left axis).

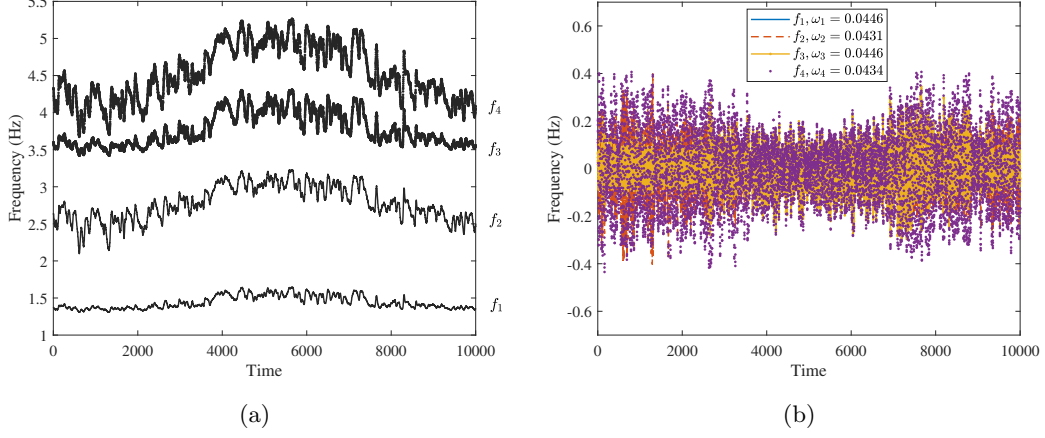Figure 6: Decomposed natural frequencies of the spring-mass system along with their center frequencies; (a) IMF_1, and (b) IMF_2. The first IMF in all cases is kept at DC (zero center frequency).

results along with the IMFs' center frequencies are shown in Figure 6. As such, a zero center frequency is obtained for all the first IMFs. The average center frequency of the second IMFs is about 0.0439 cycles per hour, or, equivalently, almost a full cycle per 24 hours. This is perfectly in line with the fact that the temperature variations between day and night has a cyclic pattern throughout the year. Therefore, the cyclic pattern of the second IMF represents only the temperature variations between day and night and does not include any information about damage. As a result, we propose that only the first IMFs to be used for the purpose of the damage detection.

The plots of Figure 7 show the distribution of the frequency signals (main diagonal entries) and their mutual co-distribution (off-diagonal entries) before (Figure 7a) and after (Figure 7b) the application of the VMD algorithm to the signals. It can be noted that denosing and removing the second IMFs using VMD can significantly decrease the presence of noise in the scatter plots of the co-distribution plots (off-diagonal plots). This will assist in training an RNN that can better capture interaction between the IMF_1 signals. Moreover, the Skewness and Kurtosis of the original signals and their corresponding IMF_1s are listed in Table 1. As such, it can be seen that both of the Skewness and Kurtosis decrease in all cases after the application of the VMD.

Next the outcomes of VMD are divided into two parts including 5000 data-points in each part.[5] The FastMCD algorithm is then used to work out the parameters corresponding to the distribution of the data in each part. These parameters are further used to work out the Mahalanobis distances of the data-points, in each part, from their true distribution. Note that

---

[5]Note that the second part itself is divided into two sets of validation and test set where the validation results are used for obtaining a UCL for the plot of obtained errors
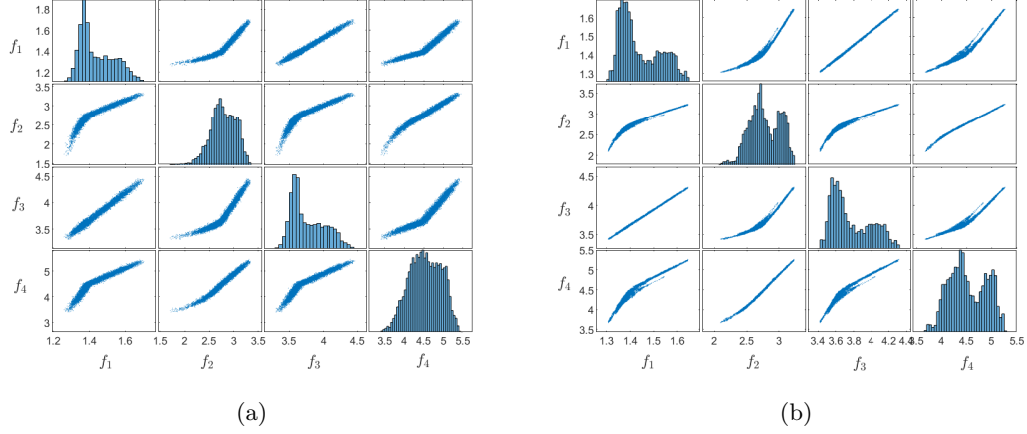
Figure 7: The mutual co-distribution of (a) the original frequency signals, and (b) IMF$_1$ of the frequency signals, regarding the spring-mass system.

Table 1: The skewness and Kurtosis of the original frequency signals and their IMF$_1$ signals of the spring-mass system after application of the VMD algorithm.

	Skewness			
Signal	f_1	f_2	f_3	f_4
Original	0.6239	-0.3950	0.6213	-0.1252
IMF$_1$	0.6005	-0.0851	0.5943	0.0819
	Kurtosis			
Original	2.2681	3.0926	2.2695	2.2711
IMF$_1$	2.0681	2.2500	2.0705	1.9465

the second dataset includes the effects of the damage as damage occurs at the 8000th record. Next, an RNN with BiLSTM cells is trained to learn the rule of obtaining Mahalanobis distances of the data-points from their distribution regarding the first part of the VMD outcomes which is further used to obtain the Mahalanobis distances of the second part of the VMD outcomes from their expected distribution.

A multivariate stacked RNN with BiLSTM units is used in this section with the following architecture:

1. a sequence input layer taking four features as inputs.

2. a BiLSTM layer with 40 units.

3. a fully connected layer with 50 units.

4. a fully connected layer with one output unit.

The detailed further settings for the developed RNN follows,

1. Adam optimisation is chosen as the optimisation algorithm [47].

2. The learning rate is set initially at 0.005 and decreased by a factor of 0.2 at every 200 epochs.

3. The number of maximum epochs of 1000 is chosen in settings.

4. A gradient threshold of 1 is considered to avoid exploding gradients effect.

The RNN is used to train on the 50% of the IMF_1 of the frequency signals corresponding to the healthy system to learn the Mahalanobis distances of these points from their true distribution obtained from the FastMCD algorithm (Figure 8a). Once the RNN has been trained, the outcomes of the VMD corresponding to the remaining 50% of the frequency signals are used in the trained RNN to predict their Mahalanobis distances from their expected distribution. The very frequency signals are also used to obtain their Mahalanobis distances from their true distribution. The absolute values of the prediction errors are shown in Figure 8b. It can be seen that the prediction error deviates significantly from the specified UCL line as soon as the damage occurs at the time 8000, showing the proposed method is successful in raising an early alarm. Moreover, it can be seen from the figure that the value of the error remains high so long as the damage presents in the system. Moreover, the obtained UCL line can clearly differentiate the damage cases from undamaged events, although some false negative/positive observations are evident from the figure. Note that these false negative/positive observations can be simply ignored as they appear for a short period of time at each time instant.

Next, the result of the proposed method is compared against the method presented in Section 3.5. Figure 9 shows the absolute error associated with remapping the IMF_1 of the first, second, third, and fourth natural frequency signals after performing the dimensionality reduction. The best results were obtained by retaining the two first PCs, i.e. PC_1 and PC_1, which is equivalent to preserving 99.92% of the variance in the dataset. Next, the T^2 control is constructed as depicted in Figure 10. A preliminary comparison between the obtained results of Figures 10 and 8b demonstrates the superiority of the proposed method of this paper. Note that further quantitative comparison of the methods would have been meaningful using some accuracy measures, such as F1 score, if the results were marginally different.
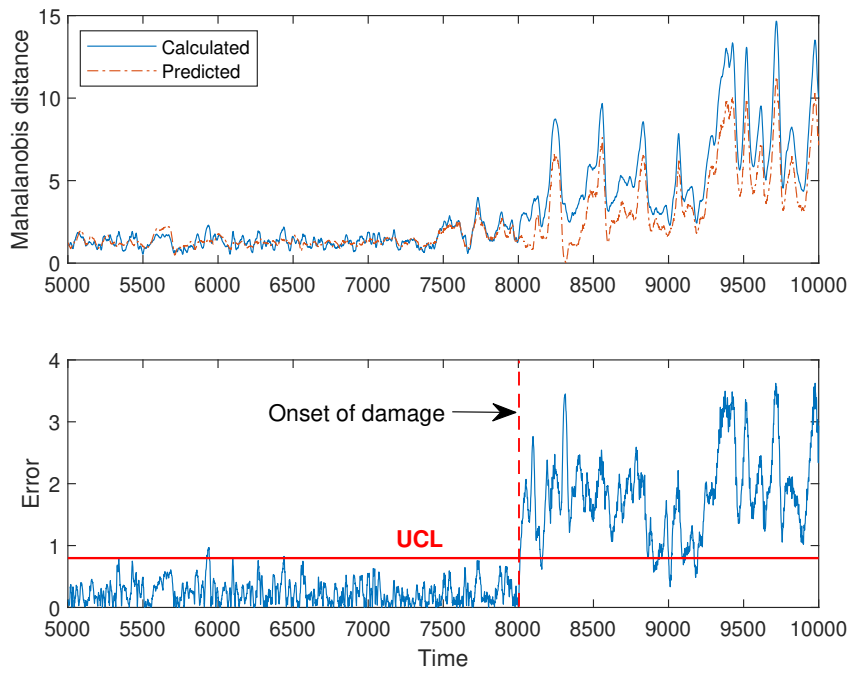
In the next section, an experimental problem of the Z24 bridge is solved using the proposed method.

5. Experimental validation: Z24 bridge

The Z24 Bridge is a well-known benchmark problem among the SHM community which is used in different problems including:
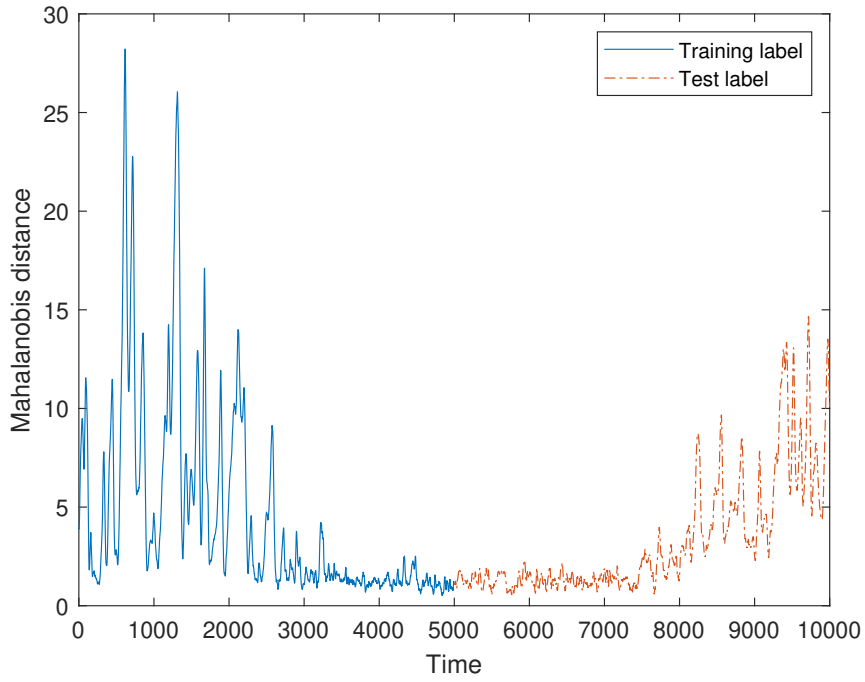
(a)

(b)

Figure 8: Mahalanobis distances regarding the spring-mass system; (a) calculated training and test labels, (b) prediction results on the test set.
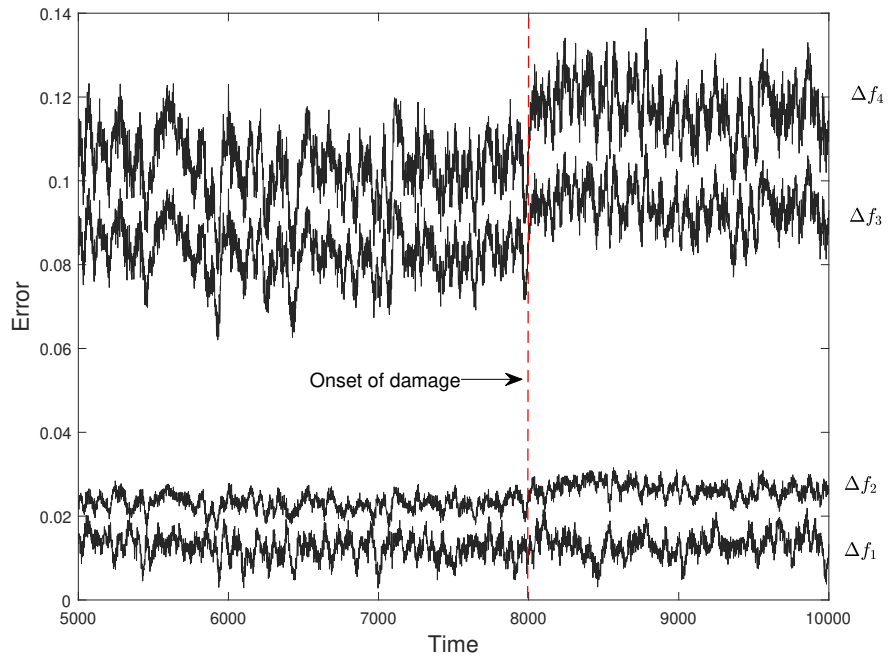
Figure 9: The clean natural frequency signals of the spring-mass system after removing EOV effects using the PCA-based strategy.
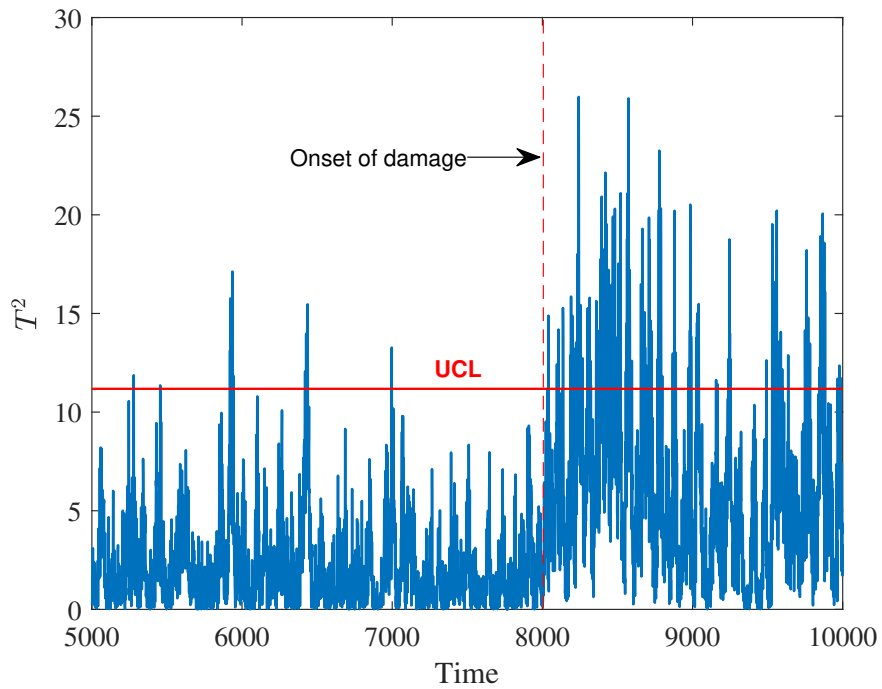

Figure 10: The absolute value of the T^2 control chart obtained from the clean natural frequency signals using PCA regarding the spring-mass system.
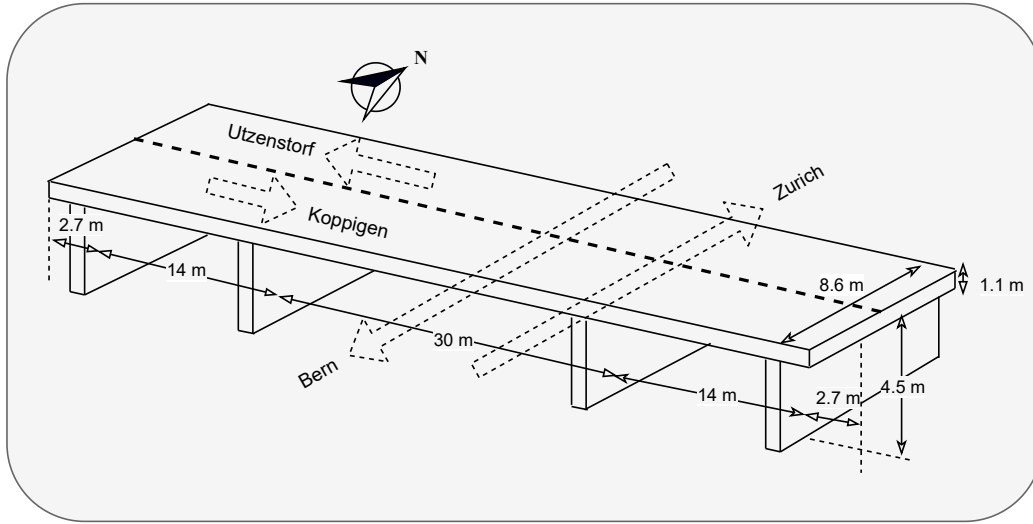
Figure 11: Z24 bridge geometry and location.

1. System identification methods for modal analysis [48, 49, 50].
2. Removing the influence of EOV on the structural dynamic characteristics [7, 51, 52].
3. Vibration-based damage identification methods [53, 54, 55].

One aim of establishing such a benchmark problem was to investigate the feasibility of vibration-based SHM of full-scale civil engineering structures through long-term tests and progressive failure assessment of a representative structure, i.e. the Z24 Bridge. The design and geometrical properties of the bridge follow,

1. It is a classical post-tensioned concrete two-cell box-girder bridge.
2. The bridge has a main span of 30 m and two side spans of 14 m as shown in Figure 11.

The monitoring campaign established one year before the bridge was demolished. During this period, several damage scenarios were implemented to facilitate a long-term continuous monitoring of the bridge prior to its total dismantlement. Since temperature is known to have a key influence on the dynamics of civil engineering infrastructures, the air and bridge's thermal state was monitored during this period in full details.

The number of 16 accelerameters were deployed on the bridge at different locations and directions to monitor the bridge dynamics. The number of 10 scans of environmental data, sampled at 48 sensors, and 8 averages of 8192 acceleration samples, taken at 16 sensors, were collected in every hour. However, there were some missing data in the original dataset. Therefore, as a data pre-processing stage, the points corresponding to time instants when the missing data occur are all removed in this paper following the same approach taken in [44]. As such, the number of 3932 data-points remained in the dataset.
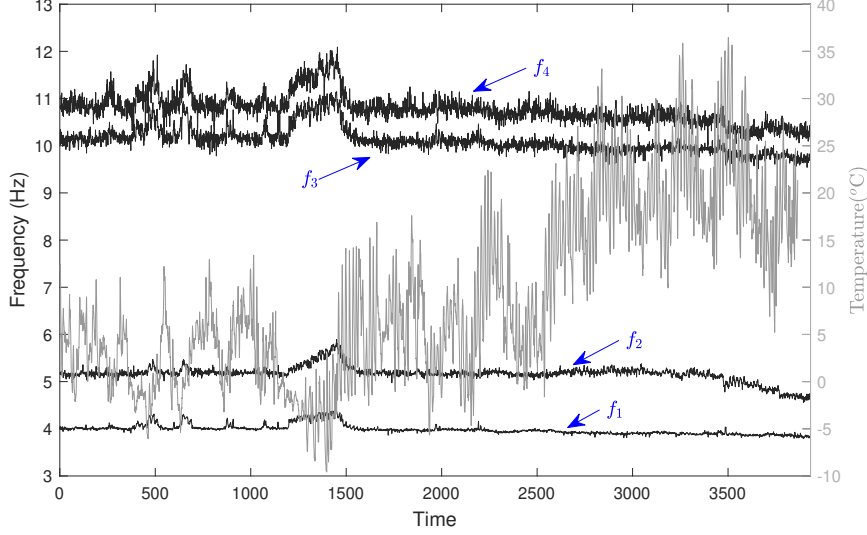
Figure 12: The recorded air temperature,and the obtained natural frequencies of the Z24 bridge during the same period of time.

Table 2: The skewness and Kurtosis of the original frequency signals and their IMF$_1$ signals of the Z24 bridge after application of the VMD algorithm.

	Skewness			
Signal	f_1	f_2	f_3	f_4
Original	1.4685	-0.3287	1.2732	0.8966
IMF$_1$	1.4834	-0.3572	1.3587	0.9520
	Kurtosis			
Original	5.3162	6.7237	5.0035	4.4209
IMF$_1$	5.3078	6.9393	5.1957	4.6265

The left-right plot of Figure 12 shows the air temperature recorded during this period. The right-left plots of the figure show the first four natural frequencies of the bridge obtained from the measured vibration signals during the same period of time. A period of very cold temperatures (roughly between the record 1200 to 1500) induced nonlinear relationships among natural frequencies and temperature. More details about the Z24 bridge can be found in [56, 57].

The natural frequency time series are decomposed into two IMFs using VMD to remove the IMF corresponding to the seasonal effects (second IMF). To this end, the same parameters, similar to those in Section 4, are used in the VMD settings except for the quadratic penalty term α which is set at 10 in here. The reason for this is that the frequency time series of the Z24 bridge are less noisy compared to the numerical example of the Section 4. The obtained decomposition results along with their center frequencies are shown in Figure 13. Note that
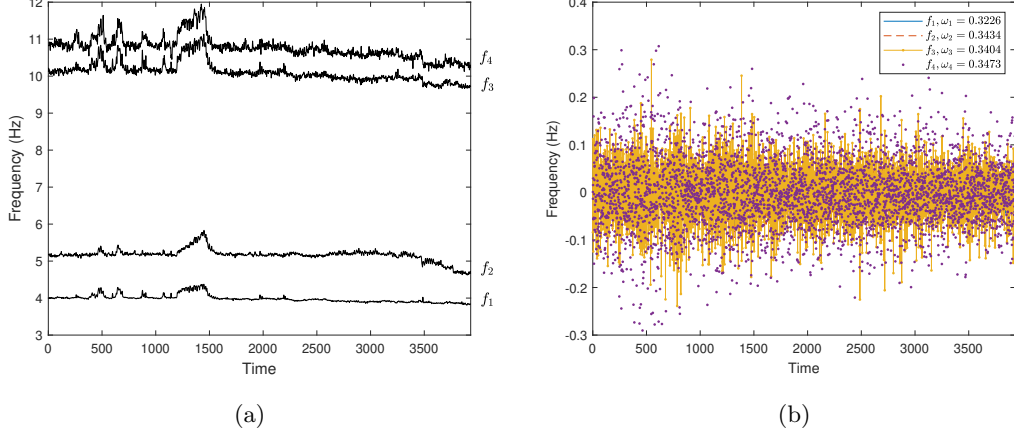
23

Figure 13: The decomposed natural frequencies of the Z24 bridge along with their center frequencies; (a) IMF$_1$, and (b) IMF$_2$. The first IMF in all cases is kept at DC (zero center frequency).
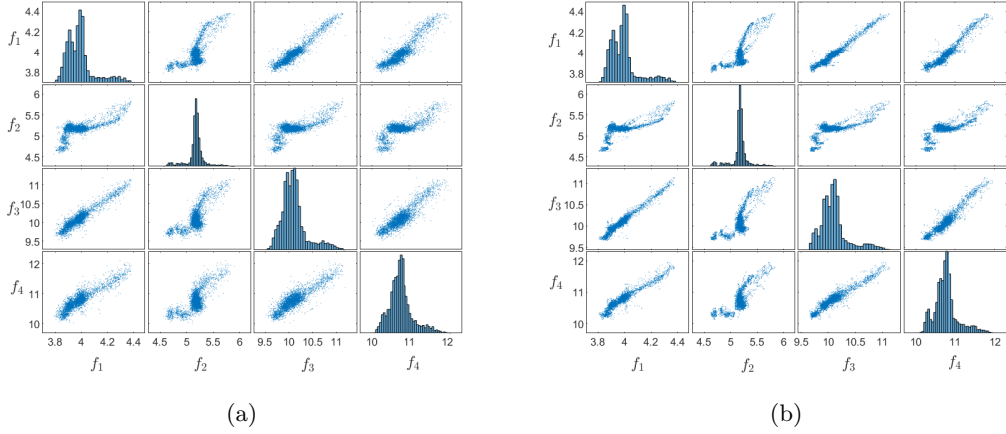

Figure 14: The mutual co-distribution of (a) the original frequency signals, and (b) IMF$_1$ of the frequency signals, regarding the Z24 bridge.

the first IMFs are kept at DC (zero center frequency) in here as well.

The plots of Figure 14 show the mutual co-distribution of the frequency signals (Figure 14a) and their obtained IMF$_1$ signals (Figure 14b). As such, one can see that VMD reduces the noise content of the co-distribution scatter plots of the off-diagonal entries. In order to understand how VMD affects the distribution of the frequency signals the skewness and kurtosis of the main diagonal plots, before and after the application of the VMD, are presented in Table 2. As such, unlike the numerical example, the skewness of the plots are slightly worsened after application of the VMD. Moreover, although the value of the kurtosis has slightly decreased for the IMF$_1$ of the first an second frequencies, a bigger value is obtained for the IMF$_1$ signals corresponding to the third and fourth frequency signals. Reiterated, note that the main rationales behind using VMD are to remove any seasonal patterns from the frequency signals as well as denoising the

signals which will further assist in training process of an RNN. As such, further investigation of the effect of VMD on the distribution of the dataset can be a subject of future work.

Next, a 50% of the data (first IMFs of the decomposed frequency signals) corresponding to the healthy state of the structure is used for training. To this end, reiterated, the FastMCD algorithm is first applied to the whole obtained IMF_1 signals to find the true distribution of the data. Then, the Mahalanobis distance of each data-point from its distribution is used as the target in an RNN. The obtained targets along with the features are then used to train the RNN. The architecture of the RNN used in this section is the same as the one used in Section 4 except for the number of units in the fully connected hidden layer which is 20 in here. The settings were mostly remained unchanged, except a regularisation term was considered to avoid over-fitting on the training set. The main reason is that the number of the elements in the training set in this section (1966), is less than the one used in the example of Section 4 (5000). A small regularisation parameter of 1×10^{-4} was, therefore, selected to this end. Moreover, the number of maximum epochs was set at 1200 where the learning rate was set to drop at every 300 epochs with the dropping rate 0.2.

Once the RNN is trained, the remaining 50% of the IMF_1 signals is used as features in the trained RNN to predict the Mahalanobis distances of the data-points from their expected distribution. The same data are used to obtain their Mahalanobis distances from their true distribution using the parameters obtained from the whole IMF_1 signal. Next, the prediction errors are calculated and used for monitoring of the health condition of the bridge.

Figure 15a shows the obtained Mhalanobis distances for the entire dataset while Figure 15b shows the prediction results obtained for the IMF_1 of the frequency signals corresponding to the test set using the trained RNN. As can be seen from the figure, the prediction error increases significantly as soon as damage occurs in the system. Moreover, the obtained UCL from the validation set, the third quarter of the dataset, can clearly differentiate the damage cases from undamaged events, although some false positive observations are evident from the figure. Note that these false positives can be ignored as they appear for a short period of time at each time instant.

Next, the damage detection of the Z24 bridge using the method presented in Section 3.5 is conducted. Figure 16 shows the obtained results of the purified IMF_1s of the frequency signals after removing the EOV effects using PCA. Likewise to the numerical section, the two columns of the transformation matrix $\mathbf{T}_b$ corresponding to the PC_1 and PC_2 are only retained. This accounts for retaining 97.7% of the variance in the dataset. Next, the obtained frequency signals are used to construct the T^2 control chart of Figure 17. It is obvious that the specified UCL line
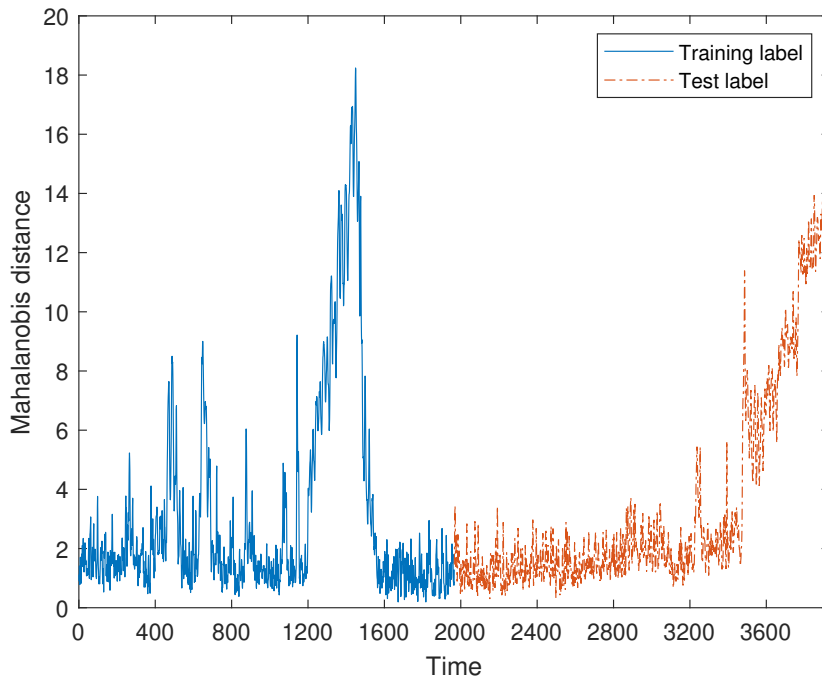
(a)
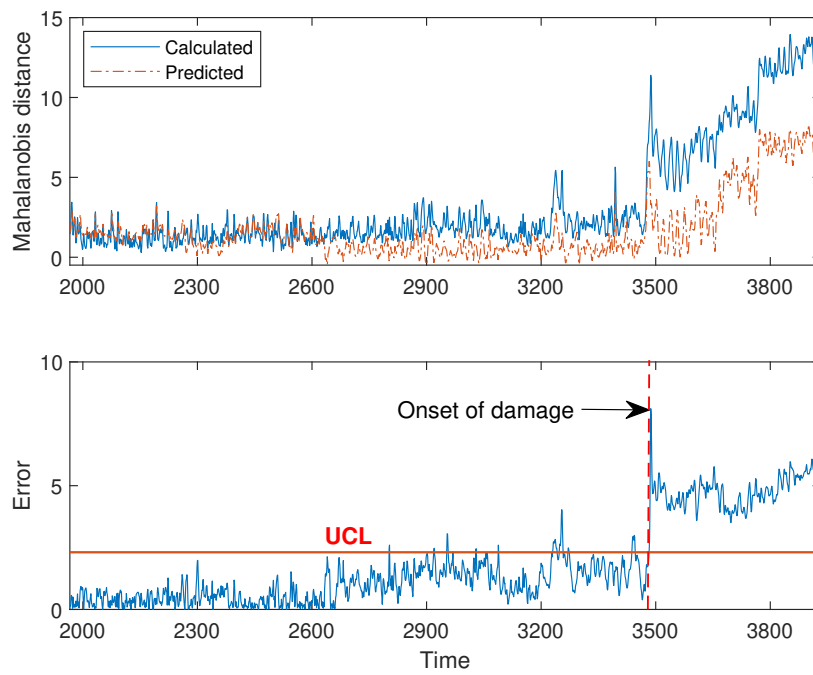
(b)

Figure 15: Mahalanobis distances regarding the Z24 bridge; (a) calculated training and test labels, (b) prediction results on the test set.
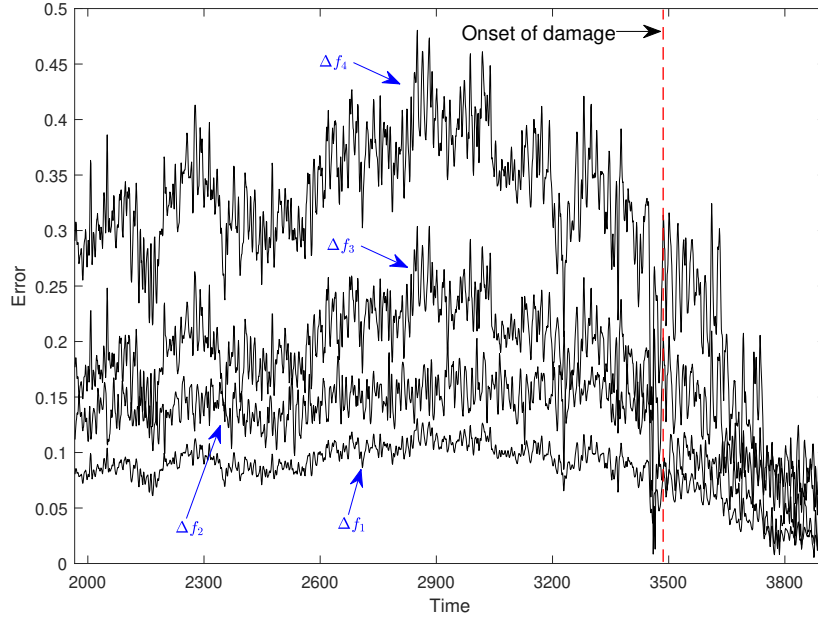
Figure 16: The clean natural frequency signals of the Z24 bridge after removing EOV effects using the PCA-based strategy.

cannot differentiate the damaged and undamaged cases properly as the number of false negative cases is plentiful. A preliminary comparison of the results of Figures 17 and 15b demonstrates the superiority of the proposed method.

6. Conclusions and future work

A systematic output-only condition monitoring of civil infrastructures has been developed in this paper. The proposed method uses VMD as a pre-processing algorithm, FastMCD to obtain the location and scatter of the inferior and posterior states of the structure, and an RNN with a BiLSTM layer to learn the rule behind calculation of the Mahalanobis distances of the data-points from their distribution. The proposed method has been tested through solving a numerical example of a spring-mass system as well as an experimental example of the Z24 bridge. The results show that the proposed method can be successfully used to raise an early alarm as soon as the structure undergoes any changes due to damage. Moreover, the proposed method was compared against a PCA-based method in both of the numerical and experimental examples of this paper. The presented PCA-based method, adapted from [26], uses the covariance matrix of the entire dataset, unlike the FastMCD algorithm that seeks to identify a robust covariance matrix of a subset of the dataset. We showed that the proposed method is more successful in condition monitoring of structures under influence of the EOV effects.

In the present paper, the FastMCD algorithm was used to work out the robust parameters
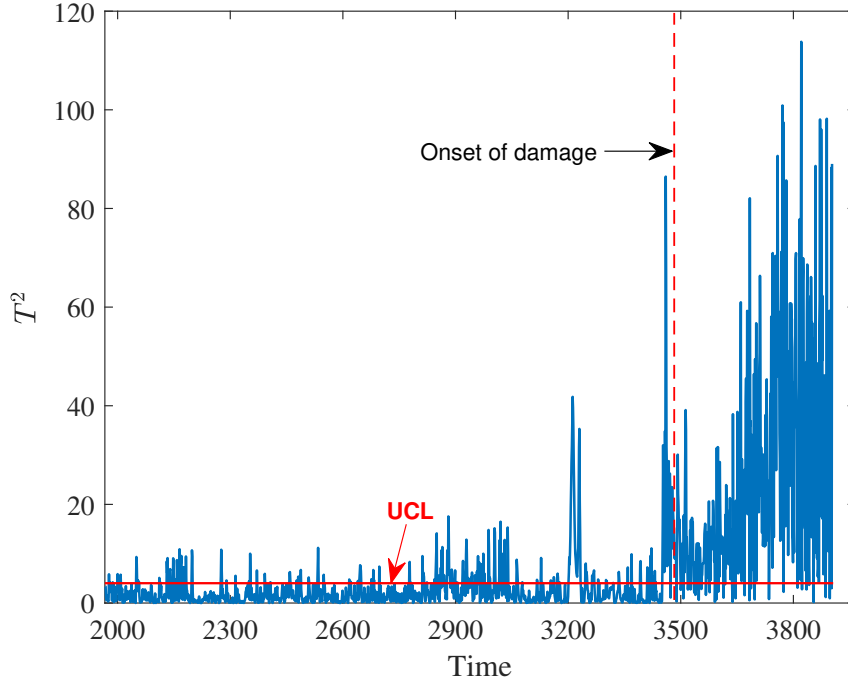
27

Figure 17: The absolute value of T^2 control chart obtained from the clean natural frequency signals using PCA regarding the Z24 bridge.

of the distribution of the mixed data of the both inferior and posterior states of the structure. These parameters were further used to obtaine Mahalanobis distances of the whole dataset from its distribution. These Mahalanobis distances were further split into two parts. The first part was used as training targets in an RNN. The second part was, however, used as a metric to be used to compare against the predicted Mahalanobis distances obtained from the trained RNN. While the entire dataset from both states was fed into the FastMCD algorithm and since the distribution of the dataset corresponding to the both of the numerical and experimental examples were shown to be skewed, one may devise a procedure in which a piecewise calculation of the location and scatter is used to obtain piecewise Mahalanobis distances. This may further improve the accuracy of the obtained results and is considered as a subject of the authors' future work.

Acknowledgement

28

References

[1] M. Regni, D. Arezzo, S. Carbonari, F. Gara, D. Zonta, Effect of environmental conditions on the modal response of a 10-story reinforced concrete tower, Shock and Vibration 2018 (2018).

[2] C. Gentile, M. Guidobaldi, A. Saisi, One-year dynamic monitoring of a historic tower: damage detection under changing environment, Meccanica 51 (11) (2016) 2873–2889.

[3] C. Liu, J. T. DeWolf, Effect of temperature on modal variability of a curved concrete bridge under ambient loads, Journal of structural engineering 133 (12) (2007) 1742–1751.

[4] G. Feltrin, Temperature and damage effects on modal parameters of a reinforced concrete bridge, in: Proc., 4th Int. Conf. on Structural Dynamics, Eurodyn 2002, Balkema, 2002, pp. 373–378.

[5] C. Rainieri, F. Magalhaes, D. Gargaro, G. Fabbrocino, A. Cunha, Predicting the variability of natural frequencies and its causes by second-order blind identification, Structural Health Monitoring 18 (2) (2019) 486–507.

[6] A. Cancelli, S. Laflamme, A. Alipour, S. Sritharan, 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 19 (2) (2020) 587–605.

[7] B. Peeters, G. De Roeck, One-year monitoring of the z24-bridge: environmental effects versus damage events, Earthquake engineering & structural dynamics 30 (2) (2001) 149–171.

[8] H. Zhou, Y. Ni, J. Ko, Constructing input to neural networks for modeling temperature-caused modal variability: mean temperatures, effective temperatures, and principal components of temperatures, Engineering Structures 32 (6) (2010) 1747–1759.

[9] R. Kromanis, P. Kripakaran, Support vector regression for anomaly detection from measurement histories, Advanced Engineering Informatics 27 (4) (2013) 486–495.

[10] J. Reilly, H. Abdel-Jaber, M. Yarnold, B. Glisic, Evaluating the coefficient of thermal expansion using time periods of minimal thermal gradient for a temperature driven structural health monitoring, in: Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, and Civil Infrastructure 2017, Vol. 10169, International Society for Optics and Photonics, 2017, p. 1016929.

[11] R. Ceravolo, G. Coletta, G. Miraglia, F. Palma, Statistical correlation between environmental time series and data from long-term monitoring of buildings, Mechanical Systems and Signal Processing 152 (2021) 107460.

[12] E. J. Cross, K. Worden, Q. Chen, Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 467 (2133) (2011) 2712–2732.

[13] E. J. Cross, K. Worden, Cointegration and why it works for shm, in: Journal of Physics: Conference Series, Vol. 382, IOP Publishing, 2012, p. 012046.

[14] P. B. Dao, A. Klepka, Ł. Pieczonka, F. Aymerich, W. J. Staszewski, Impact damage detection in smart composites using nonlinear acoustics—cointegration analysis for removal of undesired load effect, Smart Materials and Structures 26 (3) (2017) 035012.

[15] X. Li, W. Qu, L. Xiao, Y. Lu, Removal of temperature effect in impedance-based damage detection using the cointegration method, Journal of Intelligent Material Systems and Structures 30 (15) (2019) 2189–2197.

[16] P. B. Dao, W. J. Staszewski, Data normalisation for lamb wave–based damage detection using cointegration: A case study with single-and multiple-temperature trends, Journal of intelligent material systems and structures 25 (7) (2014) 845–857.

[17] E. S. Tomé, M. Pimentel, J. Figueiras, Damage detection under environmental and operational effects using cointegration analysis–application to experimental data from a cable-stayed bridge, Mechanical Systems and Signal Processing 135 (2020) 106386.

[18] A. Michalak, J. Wodecki, A. Wyłomańska, R. Zimroz, Application of cointegration to vibration signal for local damage detection in gearboxes, Applied Acoustics 144 (2019) 4–10.

[19] C. Rainieri, E. Reynders, Pre-image reconstruction for compensation of environmental effects in structural health monitoring by kernel pca, in: Life Cycle Analysis and Assessment in Civil Engineering: Towards an Integrated Vision: Proceedings of the Sixth International Symposium on Life-Cycle Civil Engineering (IALCCE 2018), 28-31 October 2018, Ghent, Belgium, Vol. 5, CRC Press, 2018, p. 267.

[20] C. Roberts, D. Garcia, 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, Springer, 2020, pp. 182–198.

[21] A. Liu, L. Wang, L. Bornn, C. Farrar, Robust structural health monitoring under environmental and operational uncertainty with switching state-space autoregressive models, Structural Health Monitoring 18 (2) (2019) 435–453.

[22] A. Entezami, H. Sarmadi, M. Salar, A. Behkamal, A. Arslan, C. De Michele, A novel structural feature extraction method via time series modelling and machine learning techniques for early damage detection in civil and architectural buildings, in: International Conference on Emerging Technologies In Architectural Design (ICETAD2019), 2019.

[23] K. Worden, G. Manson, N. R. Fieller, Damage detection using outlier analysis, Journal of Sound and vibration 229 (3) (2000) 647–667.

[24] N. Dervilis, E. Cross, R. Barthorpe, K. Worden, Robust methods of inclusive outlier analysis for structural health monitoring, Journal of Sound and Vibration 333 (20) (2014) 5181–5195.

[25] M. D. Ulriksen, D. Tcherniak, L. Damkilde, Damage detection in an operating vestas v27 wind turbine blade by use of outlier analysis, in: 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings, IEEE, 2015, pp. 50–55.

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

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

[28] N. Dervilis, K. Worden, E. Cross, On robust regression analysis as a means of exploring environmental and operational conditions for shm data, Journal of Sound and Vibration 347 (2015) 279–296.

[29] D. M. Hawkins, Identification of outliers, Vol. 11, Springer, 1980.

[30] H. Shi, K. Worden, 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 120 (2019) 16–31.

[31] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE transactions on signal processing 62 (3) (2013) 531–544.

[32] M. Hubert, M. Debruyne, P. J. Rousseeuw, Minimum covariance determinant and extensions, Wiley Interdisciplinary Reviews: Computational Statistics 10 (3) (2018) e1421.

[33] C. Croux, G. Haesbroeck, Influence function and efficiency of the minimum covariance determinant scatter matrix estimator, Journal of Multivariate Analysis 71 (2) (1999) 161–190.

[34] G. Pison, S. Van Aelst, G. Willems, Small sample corrections for lts and mcd, Metrika 55 (1-2) (2002) 111–123.

[35] K. Boudt, P. J. Rousseeuw, S. Vanduffel, T. Verdonck, The minimum regularized covariance determinant estimator, Statistics and Computing 30 (1) (2020) 113–128.

[36] P. J. Rousseeuw, K. V. Driessen, A fast algorithm for the minimum covariance determinant estimator, Technometrics 41 (3) (1999) 212–223.

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

[38] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Transactions on Signal Processing 62 (3) (2014) 531–544.

[39] Y. Wang, R. Markert, J. Xiang, W. Zheng, Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system, Mechanical Systems and Signal Processing 60 (2015) 243–251.

[40] Z. Zhao, W. Chen, X. Wu, P. C. Chen, J. Liu, LSTM network: a deep learning approach for short-term traffic forecast, IET Intelligent Transport Systems 11 (2) (2017) 68–75.

[41] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, Y. Zhang, Short-term residential load forecasting based on LSTM recurrent neural network, IEEE Transactions on Smart Grid 10 (1) (2017) 841–851.

[42] S. Siami-Namini, N. Tavakoli, A. S. Namin, The performance of lstm and bilstm in forecasting time series, in: 2019 IEEE International Conference on Big Data (Big Data), IEEE, 2019, pp. 3285–3292.

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

[44] H. Shi, K. Worden, E. J. Cross, A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring, Mechanical Systems and Signal Processing 103 (2018) 381–397.

[45] Weather history download basel (Retrieved August 18, 2020).
URL https://www.meteoblue.com/en/weather/archive/export/basel_switzerland_2661604

[46] W.-Y. He, W.-X. Ren, S. Zhu, Damage detection of beam structures using quasi-static moving load induced displacement response, Engineering Structures 145 (2017) 70–82.

[47] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[48] B. Peeters, C. Ventura, Comparative study of modal analysis techniques for bridge dynamic characteristics, Mechanical Systems and Signal Processing 17 (5) (2003) 965–988.

[49] E. Reynders, G. De Roeck, Reference-based combined deterministic–stochastic subspace identification for experimental and operational modal analysis, Mechanical Systems and Signal Processing 22 (3) (2008) 617–637.

[50] E. Reynders, J. Houbrechts, G. De Roeck, Fully automated (operational) modal analysis, Mechanical Systems and Signal Processing 29 (2012) 228–250.

[51] E. Reynders, G. Wursten, G. De Roeck, Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification, Structural Health Monitoring 13 (1) (2014) 82–93.

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

[53] J. Maeck, B. Peeters, G. De Roeck, Damage identification on the z24 bridge using vibration monitoring, Smart materials and structures 10 (3) (2001) 512.

[54] A. Teughels, G. De Roeck, Structural damage identification of the highway bridge z24 by fe model updating, Journal of Sound and Vibration 278 (3) (2004) 589–610.

[55] E. Reynders, G. De Roeck, A local flexibility method for vibration-based damage localization and quantification, Journal of sound and vibration 329 (12) (2010) 2367–2383.

[56] J. Maeck, G. De Roeck, Description of z24 benchmark, Mechanical Systems and Signal Processing 17 (1) (2003) 127–131.

[57] E. Reynders, G. De Roeck, Vibration-based damage identification: the z24 benchmark (2014).