Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 188 (2017) 248 – 255
Procedia Engineering
www.elsevier.com/locate/procedia

6th Asia Pacific Workshop on Structural Health Monitoring, 6th APWSHM

Damage Detection of Structures Subject to Nonlinear Effects of Changing Environmental Conditions

William Soo Lon Wah[a]*, Yung-Tsang Chen[a], Gethin Wyn Roberts[a] and Ahmed Elamin[a]

[a]*Department of Civil Engineering, The University of Nottingham, 199 Taikang East Road, Ningbo 315100, People's Republic of China.*

Abstract

Damage detection of civil structures has been carried out by mainly analysing the vibration properties of the structures which change when damages occur. However, these properties are also affected by the changing environmental conditions the structures are face with, and these conditions usually produce nonlinear effects on the vibration properties. Hence, a method is proposed in this paper to analyse structures subjected to nonlinear effects of environmental conditions. The method first applies Principal Component Analysis (PCA) on a bank of damage sensitivity features, followed by applying Gaussian Mixture Model on the obtained first principal component scores to cluster the data into several linear regions. By creating a baseline for each linear region using two extreme and opposite environmental conditions, and adding new measurements to the baseline one at a time followed by applying PCA, damage detection can be achieved. The method is validated on a numerical truss structure model and on the Z24 Bridge. The results demonstrate the ability of the method to analyse structures under nonlinear environmental effects.

Peer-review under responsibility of the organizing committee of the 6th APWSHM

Keywords: Principal Component Analysis; Gaussian Mixture Model; Environmental conditions; Temperature; Damage detection; Nonlinear

1. Introduction

Bridges and buildings are in constant degradation due to the severe environmental and operational conditions they are face with, therefore, during the past decades, researchers have been devising damage detection methods to analyse the structures so that early warning of structural damages can be achieved. Those developed methods are mainly based

* Corresponding author. Tel.: +86 13136373521.
E-mail address: William.SOO@nottingham.edu.cn

Peer-review under responsibility of the organizing committee of the 6th APWSHM
doi:10.1016/j.proeng.2017.04.481

on analysing the vibration properties (e.g. natural frequencies) of the structures, which change when damages occur. However, these vibration properties are not only affected by the occurrence of damages, but they are also affected by the changing environmental conditions (e.g. temperature), which limits the use of most developed damaged detection methods that did not account for these effects [1].

Thus, to tackle the problem of the changing environmental conditions, several approaches have been proposed in the literature. For example, one popular approach is to make use of multivariate statistical tools (e.g. Principal Component Analysis (PCA)) to extract features that are sensitive to damages but less sensitive to the effects of the changing environmental conditions [2, 3]. However, in those methods, a linear relationship between the environmental conditions and the damage sensitivity features was assumed. Yet, it is well known that the changing environmental conditions usually have nonlinear effects on the damage sensitivity features. Hence, in this instance, Yan et al. [4] proposed to first cluster the damage sensitivity features data set into several linear data sets, followed by applying PCA to each linear data set for damage detection. Sohn et al. [5] proposed to use an auto-associative neural network to perform nonlinear PCA (NLPCA). They proposed to train the NLPCA using the auto-associative neural network to extract the dependency of the damage sensitivity features on the unmeasured environmental conditions. Damage is indicated when the prediction error increases. Reynders et al. [6] proposed to use Kernel PCA to create nonlinear output-only model of the undamaged structure to be used as baseline. New measurements can then be compared to the model and any growth in prediction error can then be attributed to damage.

Although different techniques have been proposed to tackle the problem of the nonlinear effects produced by the environmental conditions, most of the methods have the drawback that the baseline of the undamaged structure needs to be created using damage sensitivity features obtained from a wide range of environmental conditions. This has the disadvantage that the sensitivity of the system to detect small damages may be reduced. Thus, in this paper, a damage detection method is proposed to tackle the nonlinear effects produced by environmental conditions on damage sensitivity features for damage detection. The proposed method introduces the concept of using two extreme and opposite environmental conditions as baseline so as to remove the need to use features from a wide range of environmental conditions as baseline. The method uses PCA and Gaussian Mixture Model (GMM) to process the damage sensitivity features and to linearize the system through clustering, respectively. The proposed method is tested using a numerical truss structure model and validated using measurements from the Z24 Bridge. The results obtained demonstrate the ability of the proposed method for damage detection under nonlinear effects of changing environmental conditions.

2. Methodology

The proposed damage detection method uses GMM to linearize the system and PCA for data processing. An introduction on those two mathematical tools is first given, followed by a description of the proposed method.

2.1. Principal Component Analysis

PCA is a multivariate statistical tool used to reduce the dimensions of a data set and to highlight the similarities and differences in the data set. It forms new variables called principal components to characterize the variances in the data set using linear combinations of each of the original variables [7]. Damage sensitivity features (e.g. natural frequencies) can be processed using PCA to extract the main factors driving the variances in the data set to distinguish the different factors affecting the data set. These factors may be due to the changing environmental conditions as well as to damage of structural components. A brief description on the principal behind PCA is given below.

Let $\boldsymbol{Z}$ denotes a $n \times q$ data set of damage sensitivity features collected from q observations. For each observation, n numbers of features are collected. To perform PCA on the features data set, mean centering of the original data set $\boldsymbol{Z}$ is first required to give a matrix $\boldsymbol{S}$. This is achieved by subtracting the mean of each row in the data set to each measurement in that row. PCA transforms the data set $\boldsymbol{S}$ into a new $m \times q$ data set $\boldsymbol{Y}$ with smaller dimensions which characterizes most of the variances in the original data set. The relationship between the newly formed data set $\boldsymbol{Y}$ and the data set $\boldsymbol{S}$ can be represented using a loading matrix $\boldsymbol{T}$ which has dimensions $m \times n$ as follows.

$$\boldsymbol{Y} = \boldsymbol{TS} \tag{1}$$

$\boldsymbol{Y}$, the newly formed data set, is termed the score matrix. It combines scores each observation obtains for different factors affecting the data set into a matrix. The factors here, called principal components, may represent the changing environmental conditions as well as damage of structural components. The principal components are arranged in descending order in the matrix $\boldsymbol{Y}$ with the first principal component representing the factor(s) creating the largest variance in the data set, while the last principal component creating the least variance. The score of each observation for each principal component can be thought of as a coordinate along the principal component axis representing the location of each observation along each axis. PCA makes the first principal component to represent the factor(s) creating the greatest variance by rotating the cloud of data in such a way to minimize the distance of each point in the cloud to the first principal component axis while assuring that the axis goes through the zero centroid. It is for this reason that PCA requires mean centering of the original data set prior to application.

The loading matrix $\boldsymbol{T}$ contains coefficients which are used to compute the score matrix through linear combinations of the variables in the data set $\boldsymbol{S}$. The rows of the loading matrix correspond to the eigenvectors of the covariance matrix of $\boldsymbol{S}$ and they can be obtained by decomposing the matrix $\boldsymbol{S}$ using singular value decomposition and use that decomposition to construct the covariance matrix of $\boldsymbol{S}$ as follows.

$$\frac{1}{q-1}\boldsymbol{S}\boldsymbol{S}^T = \boldsymbol{U}\frac{\boldsymbol{A}^2}{q-1}\boldsymbol{U}^T \tag{2}$$

Where $\boldsymbol{U}$ is an orthonormal matrix ($\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{I}$) whose columns represent the eigenvectors of the covariance matrix of $\boldsymbol{S}$ (hence $\boldsymbol{T} = \boldsymbol{U}^T$), and $\boldsymbol{A}$ is a diagonal matrix with the diagonal terms representing the singular values arranged in descending order. Each singular value indicates how much variance its corresponding eigenvector explains in the original data set.

By applying Equation (1), the score matrix $\boldsymbol{Y}$ can be generated. Usually, to reduce the dimensions of the original data set, only the first m rows (eigenvectors) of the loading matrix $\boldsymbol{T}$ are used to construct the score matrix. However, in this study, all the rows are used since PCA is used here to extract the similarities and differences in the original data set, rather than reducing the dimensions of the original data set. Analysing only the first few rows of the score matrix (first few principal components), damage detection can be performed.

2.2. Gaussian Mixture Model

GMM is often used for data clustering and is used here to linearize nonlinear systems produced by nonlinear effects of changing environmental conditions on damage sensitivity features. A brief description on GMM is given below.

Let $\boldsymbol{X}$ be a multivariate data set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ composed of nonlinear data (e.g. damage sensitivity features) captured under nonlinear effects of environmental conditions from N observations. Nonlinear data are not normally distributed, hence they cannot be modelled as a single Gaussian distribution. Therefore, a mixture of Gaussian components whose distribution can be written as a linear superposition of K Gaussian densities can be adopted to model the data [8].

$$p(\boldsymbol{x}) = \sum_{i=1}^{K} \pi_i \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{3}$$

Each Gaussian component of the mixture given as $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_i\boldsymbol{\Sigma}_i)$ has its own mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$. The parameters π_i in Equation (3) are called the mixing coefficients and they ranged between 0 and 1 ($0 \leq \pi_i \leq 1$) and summed up to one. The goal is to maximize the likelihood function given in Equation (4) with respect to the parameters ($\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$ and π_i) so as to determine the component each data point $\boldsymbol{x}_j$ belongs to.

$$\ln p(\boldsymbol{X}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \pi) = \sum_{j=1}^{N} \ln\left\{\sum_{i=1}^{K} \pi_i \mathcal{N}(\boldsymbol{x}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)\right\} \tag{4}$$

However, these parameters are unknown since it is not known which data points belong to which component. Thus, these unknown parameters can be estimated using the Expectation-Maximization (EM) algorithm. The EM algorithm is an iterative process which is composed of two steps, namely the expectation (E) step and the maximization (M) step. In the E step, the parameters (initial guess at the beginning) are held fixed and the posterior probability of the component i given the data point $\boldsymbol{x}_j$ is evaluated. Then, in the M step, the parameters are re-estimated using the posterior probability calculated in the E step. The log-likelihood given in Equation (4) is then evaluated. Convergence of either the parameters or the log likelihood is checked, and if the criteria is not satisfied, the process will iterate using the up-to-data values until the criteria is met. For more information about the EM algorithm, see Bishop [8].

2.3. Proposed damage detection method

It is well known that the changing environmental conditions usually have nonlinear effects on the damage sensitivity features of structures. Therefore, in this paper, a damage detection method is proposed to tackle the nonlinear effects. It is proposed to first cluster the damage sensitivity features data set into several linear data sets using GMM. To facilitate clustering, it is proposed to first apply PCA on the damage sensitivity features data set and to cluster the first principal component scores obtained for all the observations arranged in ascending order. This will allow the observations to be ranked according to their environmental conditions, hence the clusters will represent the different environmental conditions (e.g. observations above 0 °C and observations below 0 °C). It is then proposed to create the baseline of the undamaged structure for each linear region so as to represent the different conditions (e.g. above and below 0 °C) that may happen to the structure. To create the baseline of each linear region, it is proposed to use two extreme and opposite environmental conditions so as to force the first principal component to represent the environmental effects. This is achieved by selecting the first few and the last few observations of each linear region. Adding each new measurement to be monitored one at a time to each baseline and using PCA for data processing, damage detection can be achieved by analysing the first few principal components. Damage will be indicated to have occurred when the monitored observation lies outside all the baselines (e.g. above and below 0 °C) in the principal component plots. It should be noted that only the principal components that have the baseline observations obtained from the two extreme and opposite cases to be separated from each other that should be analysed since two extreme and opposite cases have been used for the baseline so as to represent two different conditions obtained from the same factor(s) affecting the data set. The procedures to follow for damage detection is given below in Fig. 1.

3. Case study

3.1. Two-dimensional truss structure

A two-dimensional truss structure model (Fig. 2(a)) is used to test the proposed method for different undamaged and damaged cases subjected to varying temperature condition given in Table 1. The structure consists of 30 elements and is made up of steel material with Young's modulus of 200 GPa, density of 7850 kg/m^3 and cross sectional area of 0.001 m^2. The Young's modulus of the material is made to be temperature dependent with reference temperature taken at 20 °C. As was mentioned by Kullaa [9], when the temperature the structures are faced with falls below 0 °C, the vibration properties of structures often change abruptly. Therefore, to simulate the sudden change in behaviour of the vibration properties, a bilinear relationship is assumed between the Young's modulus of the material and the temperature condition as shown in Fig. 2(b). Damage is assumed to be represented by a reduction in stiffness of the members. It is assumed that the first four natural frequencies of the undamaged structure obtained at temperatures of -20 °C to 40 °C with 1 °C interval are available from which the baseline of the undamaged structure can be constructed.

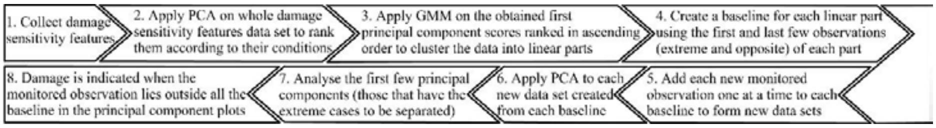

Fig. 1. Procedures to follow for damage detection.

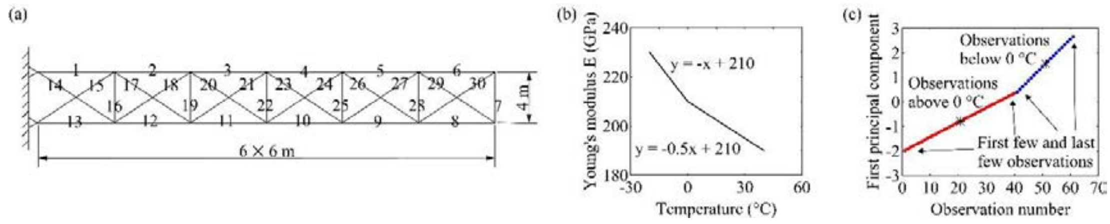

Fig. 2. (a) Two-dimensional truss structure; (b) Graph of Young's modulus versus temperature; (c) Data clustering of damage sensitivity features.

Table 1. Descriptions of the undamaged and damaged cases.

Case number	Temperature (°C)	Element number	Damaged extent (%)
Undamaged case 1	-15	-	-
Undamaged case 2	20	-	-
Damaged case 1	-10	8	5
Damaged case 2	35	2 and 26	30 and 20

PCA is first applied to the natural frequencies used as damage sensitivity features and the first principal component scores of all the observations are arranged in ascending order so as to rank the observations according to their temperature condition. GMM is then applied to the ranked first principal component scores for clustering and the results obtained are given in Fig. 2(c). It can be seen that two linear regions exist in the data set which represent the behaviour of the structure at temperatures below (blue) and above (red) 0 °C. Thus, two baselines are created using two extreme and opposite temperature conditions. For each baseline, ten observations are used, with five observations taken from the first five observations of the linear region and the other five observations obtained from the last five observations. This allows the two baselines to cover all possible temperature conditions the structure may encounter.

Each case to be analysed is then added one at a time to the two baselines to form new data sets and PCA is applied. From the results of undamaged case 1 (Fig. 3(a) & (b)), it can be seen that the monitored observation (the cross) lies between the two extreme cases used as baseline (ten dots) in the first principal component plot of the cold temperature baseline (Fig. 3(a)). This indicates that the temperature the structure is faced with lies somewhere between the two extreme cases adopted in the cold temperature baseline. In the hot temperature baseline (Fig. 3(b)), it is logical to see the monitored observation outside the baseline since the temperature condition is outside the range of the temperature considered in this baseline. Hence, only the principal components of the cold temperature baseline need to be analysed. In the second principal component plot of the cold temperature baseline, since all the baseline observations are mixed together, this component is not analysed since this component does not represent a consequent factor(s) affecting the data set. Thus, it is concluded that the structure is undamaged and the only consequent factor(s) affecting the data set is the temperature deviation from the baseline. For the undamaged case 2 (Fig. 3(c) & (d)), it is also concluded that no damage is present in the structure. However, this time the hot temperature baseline (Fig. 3(d)) is analysed since the monitored case lies inside the baseline in the hot temperature baseline rather than the cold temperature baseline.

From the results obtained for damaged case 1 (Fig. 4(a) & (b)), it can be seen that the cold temperature baseline (Fig. 4(a)) should be analysed since the monitored observation lies between the extreme baselines in the first principal component for that cold baseline. Analysing the second principal component, it is concluded that damage occurs in the structure. This second component has the two extreme cases baseline observations to be separated from each other (if looking at the score values since it is difficult to see from this figure) and are at the opposite end to the monitored case. This indicates that another consequent factor(s) is affecting the data set, hence requiring a second principal component to represent the variance in the data set. This consequent factor(s) is attributed to damage. Looking at the third principal component, it can be seen that the baseline observations are mixed together, thus removing the need for further analysis.

For the second damaged case (Fig. 4(c) & (d)), it can be seen that the monitored observation lies outside the baseline in the first principal component for both cold and hot temperature baselines. This is because the natural frequencies of the structure are lower than those of the baselines. This is due to the fact that a factor(s) other than the temperature

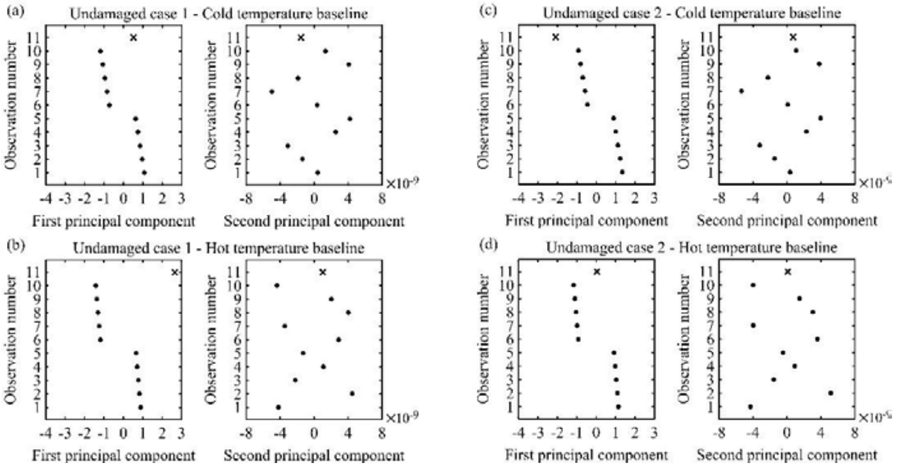

Fig. 3. First and second principal component for undamaged case 1 ((a) & (b)) and undamaged case 2 ((c) & (d)).

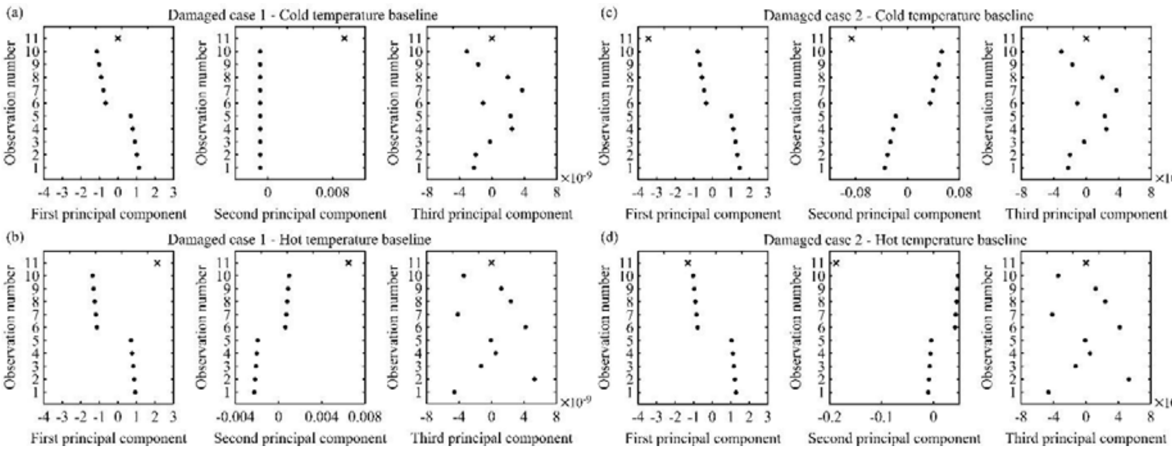

Fig. 4. First, second and third principal component for damaged case 1 ((a) & (b)) and damaged case 2 ((c) & (d)).

condition has a consequent effect on the structure to reduce the natural frequencies below the baselines' frequencies. This factor(s) is attributed to damage. Looking at the second principal component of both baselines, it can be seen that the two extreme cases baseline observations are separated from each other and are at the opposite end to the monitored observation. This indicates that a consequent factor(s) is affecting the baselines and the monitored case differently, and this is attributed to damage. This represents the different patterns of the natural frequencies from that of the baselines which occur due to damage. Generally, the first principal component will give each observation a rough representation of the values of the natural frequencies of the structure relative to each other. Hence, the first principal component of the undamaged cases will represent the temperature factor since only the temperature condition will affect the natural frequencies, while for the damaged cases, it will represent the temperature and damage factor.

3.2. Z24 Bridge

The Z24 Bridge was monitored for nearly a year with realistic damaged scenarios applied to it near the end of the monitoring period (see Kramer et al. [10] for details). The first four natural frequencies of the bridge had nonlinear relationship with the temperature condition [11], therefore presenting a good case study for the proposed method.

The first 1500 data points (first four natural frequencies used as damage sensitivity features) obtained from the undamaged structure which contain observations gathered from temperatures below and above 0 °C, along with all the damage cases data points are used to test the proposed method. PCA and GMM are first applied to the remaining undamaged cases data points to create the baseline of the undamaged structure. Two linear regions are obtained from the data set as shown in Fig. 5 which represent the behaviour of the structure at temperatures below (blue) and above (red) 0 °C. From the plot, the first 24 observations are abnormal since they have a sharp change in gradient and do not follow the trend line. Thus, these 24 points are omitted in the creation of the baseline. Two baselines using extreme and opposite cases are created to represent all possible conditions the structure may encounter

The results of a randomly chosen undamaged case and a randomly chosen damaged case are given in Fig. 6(a) & (b) and Fig. 6(c) & (d), respectively. From the first principal component of the undamaged case, it can be seen that the monitored observation lies between the baselines in the hot temperature baseline (Fig. 6(b)). Analysing the second principal component of that baseline, it can be seen that the baselines are mixed together, thus stating that the structure is undamaged. For the damaged case, the monitored observation lies outside the baseline in both cold and hot temperature baselines (Fig. 6(c) & (d)) in the first principal component. In the second principal component of both baselines, all the observations are mixed, thus removing the need for further analysis. Therefore, it is concluded that damage is present in the structure through analysis of the first principal component plots. To give a representation of the results obtained for all the cases, plots of the first principal component for all the undamaged and damaged cases are presented in Fig. 7(a) & (b) and Fig. 7(c) & (d), respectively. All the cases are adjusted to have baselines values of 0 and 1. The two horizontal dotted lines in the plots represent the baselines. From the results obtained, it is concluded that the proposed method performs well to analyse this real bridge structure, with successful rates of 87.0 % to indicate damaged for the damaged cases and 99.7 % to indicate undamaged for the undamaged cases. For the undamaged cases, either the monitored observation lies inside the cold temperature baseline Fig. 7(a) or the hot temperature baseline Fig. 7(b). Most of the damaged cases can be seen to be outside both cold and hot temperature baselines (Fig. 7(c) & (d)). From Fig. 7(c) & (d), it can be seen that damage evolution can also be obtained from the proposed method.

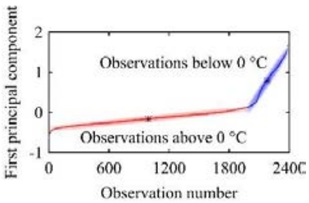

Fig. 5. Data clustering of damage sensitivity features of the Z24 Bridge.

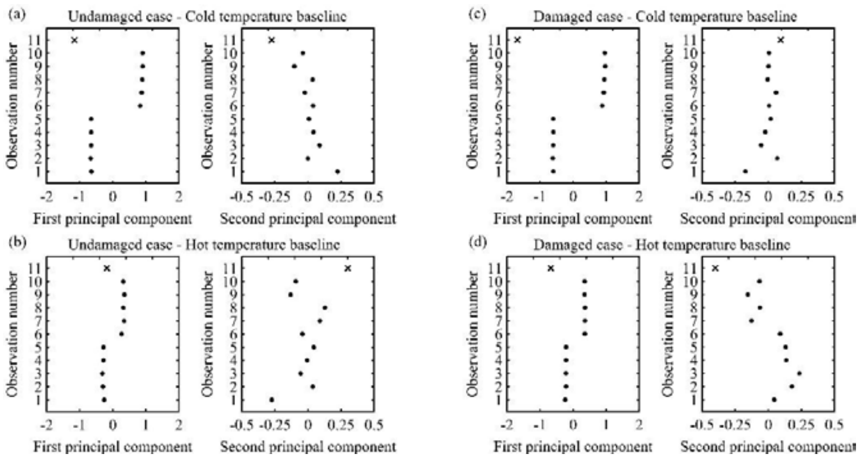

Fig. 6. First and second principal component for randomly chosen undamaged case ((a) & (b)) and damaged case ((c) & (d)).

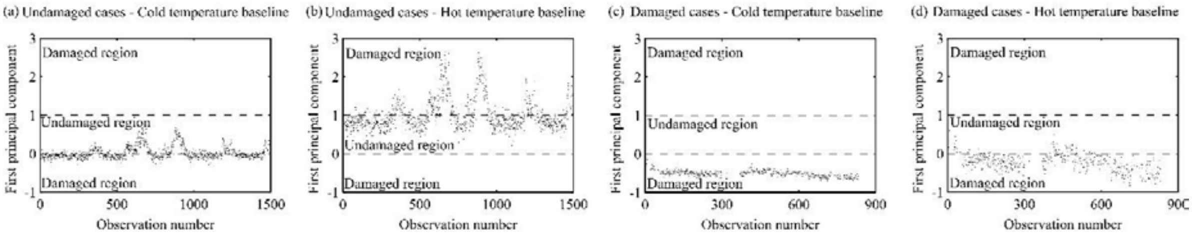

Fig. 7. Results of the first principal component for all undamaged cases ((a) & (b)) and damaged cases ((c) & (d)).

4. Conclusions

A damage detection method is presented in this paper to analyse structures that are subjected to nonlinear effects of environmental conditions. A new approach of creating the baseline of the undamaged structure which consists of damage sensitivity features obtained from two extreme and opposite environmental conditions is proposed so as to remove the need to use a large set of damage sensitivity features. The method uses Principal Component Analysis for data processing and Gaussian Mixture Model to linearize the nonlinear system. The proposed method is tested using a numerical truss structure model and the Z24 Bridge. The results obtained show the robustness of the proposed method for damage detection. The method allows near real-time monitoring since each measurement is analysed one at a time. However, it is required to form the extreme cases baseline, which can be difficult. An improvement to the proposed method is on course to allow the extreme cases to be created using data obtained from only a few months.

Acknowledgements

The authors would like to acknowledge Prof. Guido De Roeck and Prof. Edwin Reynders for providing the data of the Z24 Bridge. This work was supported by Zhejiang Province Human Resources and Social Security Department under Qiajiang Talent Scheme (Grant number QJD1402009), Ningbo Science and Engineering Bureau under Ningbo Natural Science Scheme (Grant number 2014A610025) and Ningbo Science and Technology Bureau's International Academy for the Marine Economy and Technology 'Structural Health Monitoring of Infrastructure in Logistics Cycle' (Grant number 2014A35008). The financial support is greatly appreciated.

References

[1] H. Sohn, Effects of environmental and operational variability on structural health monitoring, Philos T Roy Soc A. 365 (1851) 539-560.
[2] G. Manson, Identifying damage sensitive, environment insensitive features for damage detection, In Proceedings of the 3rd International Conference on Identification in Engineering Systems, Swansea, UK, 15-17 April 2002, 187-197.
[3] A.M. Yan, G. Kerschen, P. De Boe, J.C. Golinval, Structural damage diagnosis under varying environmental conditions part I: a linear analysis, Mech Syst Signal Pr. 19 (4) 847-864.
[4] A.M. Yan, G. Kerschen, P. De Boe, J.C. Golinval, Structural damage diagnosis under varying environmental conditions part I: local pca for non-linear cases, Mech Syst Signal Pr. 19 (4) 865-880.
[5] H. Sohn, K. Worden, C.R. Farrar, Statistical damage classification under changing environmental and operational conditions, J Intell Mater Syst Struct. 13 (9) 561-574.
[6] E. Reynders, G Wursten, G. De Roeck, Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification, Struct Health Monit. 13 (1) 82-93.
[7] S. Sharma, Applied Multivariate Techniques, John Wiley & Sons, Inc., New York, 1996.
[8] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.
[9] J. Kulla, Structural health monitoring under nonlinear environmental or operational influences, Shock Vib. 2014(2014).
[10] C. Kramer, C. De Smet, G. De Roeck, Z24 Bridge damage detection tests, In Proceedings of the 17th International Modal Analysis Conference, Kissimmee, FL, 8-11 February 1999, 1023-1329.
[11] B. Peeters, G. De Roeck, One-year monitoring of the Z24-Bridge: environmental effects versus damage events, Earthq Eng Struct Dyn. 30(2) 149-171.