D. Posenato, F. Lanata, D. Inaudi and I. F. C. Smith. Model-free Data Interpretation for Continous Monitoring of Complex Structures, in Advanced Engineering Informatics, vol. 22, num. 1, p. 135-144, 2008. doi:10.1016/j.aei.2007.02.002

Model Free Data Interpretation for Continuous Monitoring of Complex Structures

Daniele Posenato[1], Francesca Lanata[2], Daniele Inaudi[1] and Ian F.C. Smith[3]

[1] Smartec SA, via Pobiette 11
CH-6928 Manno, Switzerland
posenato@smartec.ch, inaudi@smartec.ch

[2] Department of Structural and Geotechnical Engineering
University of Genoa
Via Montallegro, 1
16145 Genoa, Italy

[3] Ecole Polytechnique Fédérale de Lausanne (EPFL)
GC G1 507, Station 18
CH-1015 Lausanne, Switzerland
Ian.Smith@epfl.ch

Abstract

Civil engineering structures are difficult to model accurately and this challenge is compounded when structures are built in uncertain environments. As consequence, their real behavior is hard to predict; such difficulties have important effects on the reliability of damage detection. Such situations encourage the enhancement of traditional approximate structural assessments through in-service measurements and interpretation of monitoring data. While some proposals have recently been made, in general, no current methodology for detection of anomalous behavior from measurement data can be reliably applied to complex structures in practical situations.

This paper presents two new methodologies for model-free data interpretation to identify and localize anomalous behavior in civil engineering structures. Two statistical methods i) moving principal component analysis and ii) moving correlation analysis have been demonstrated to be useful for damage detection during continuous static monitoring of civil structures.

The algorithms are designed to learn characteristics of time series generated by sensor data during a period called the initialization phase where the structure is assumed to behave normally. This phase subsequently helps identify those behaviors which can be classified as anomalous. In this way the new methodologies can effectively identify anomalous behaviors without explicit (and costly) knowledge of structural characteristics such as geometry and models of behavior. The methodologies have been tested on numerically simulated elements with sensors at a range of damage severities. A comparative study with wavelet and other statistical analyses demonstrates superior performance for identifying the presence of damage.

Keywords : model-free data interpretation, complex structures, moving principal components analysis, moving correlation analysis

1 Introduction

Structural health monitoring methods may employ sensors for non-destructive in-situ structural evaluations. These sensors produce data (either continuously or periodically) that are analyzed to assess the safety and performance of structures. Such data may be helpful for early identification of damage [1]. Damage usually alters i) dynamic properties (for example, modal frequencies and modal shapes) and ii) static response to loading and environmental conditions.

To assess the state of structures there are, generally, two monitoring strategies: dynamic and static monitoring. The applicability of the dynamic monitoring is limited since even significant damage may cause only small changes in natural frequencies, particularly in complex structures [1][2][3][4][5]. The success of these techniques is also affected by the presence of noise in experimental data [6-8]. For example, Gentile and Messina [8] showed that in the case of wavelet analysis, the presence of noise can mask damage.

Static monitoring can lead to damage identification by comparing static structural response (measuring displacements or strains due to environmental effects and applied loading), with predictions from behavior models [9]. This is often referred to as system identification. However, models can be expensive to create and may not accurately reflect undamaged behavior. Difficulties and uncertainties increase in presence of complex civil structures so that a well defined and unique behavior model of cannot be clearly identified [10]. Furthermore, multiple-model system identification may not succeed in identifying the right damage [11].

When no models are used, long periods are required to produce reliable information [12]. For static monitoring, the challenge of damage detection and localization has only recently been treated [13-15]. Despite continuous evolution of research, for continuous static monitoring, no reliable strategies for identifying damage have been proposed and verified for broad classes of civil structures [10][16].

Another approach is to evaluate data statistically. This approach involves examining changes in time series over time [17]. The methodology is completely data driven; the evolution of the data is estimated without information of physical processes [18]. Measurements are taken continuously during service over long periods.

The number of structures that are monitored is growing. This monitoring produces great amounts of measurement data in different formats from which it is necessary to extract knowledge. The most difficult challenge that faces the structural health monitoring community is not the lack of measurement technology but rather, finding rational methods to acquire, process, and analyze large amounts of data that are generated in order to create information on the health of structures [19-20].

The objective of this paper is to propose methodologies that discover anomalous behavior in data generated by sensors without using behavior models. We present reliable methodologies that can be applied to broad classes of civil structures with a low risk of generating false positives and false negatives. Another aspect of these methodologies is that there are applicable to entire lives of structures. Life-cycle usefulness requires adaptability to new structural states. Once the presence of an anomalous state (for example, temporary, definitive, still in progress) is detected; the methodologies adapt to new states in order to detect further anomalies.

The paper is organized as follows: Section 2 provides a description of the numerical simulation used to compare the damage detection performances of several algorithms. Section 3 introduces the methodologies, particularly how PCA and correlation analysis are modified for long measurement periods and large amounts of data. Section 4 includes a discussion of the application of these new algorithms to structural health monitoring. Section 5 presents a comparative study between these algorithms and other methods (Continuous Wavelet [21-22], Short Term Fourier Transform [22] and Instance-Based Method [23-26]) based on data derived from a numerically simulated beam in healthy and various damages states. The paper finishes with conclusions and plans for future work.

2 Model-Free Data Interpretation

2.1 Principal Components Analysis (PCA)

PCA is a data reduction tool that is capable of compressing data and reducing its dimensionality so that essential information is retained and made easier to analyse than the original data set. The main objective is to transform a number of related process variables to a smaller set of uncorrelated variables [27]. A key step is finding those principal components that contain most of the information. PCA is based on an orthogonal decomposition of the covariance matrix of the process variables along the direction that explain the maximum variation of the data.

PCA is applied to the data in this study in an effort to reduce the dimensionality of the data and to enhance the discrimination between features of undamaged and damage structures. The first step of PCA is the construction of a matrix **U** that contains the history of all the measured parameters:

$$\mathbf{U}(t) = \begin{bmatrix} u_1(t_1) & u_2(t_1) & \cdots & u_{Ns}(t_1) \\ u_1(t_2) & u_2(t_2) & \cdots & u_{Ns}(t_2) \\ \cdots & \cdots & \cdots & \cdots \\ u_1(t_N) & u_2(t_N) & & u_{Ns}(t_N) \end{bmatrix} \tag{1}$$

where N represents the total number of time observations during the monitoring period, and N_S represents the total number of sensors on the structure. Thus, each column of the matrix **U** is the time history of each sensor.

Time histories are first normalized by subtracting mean values, $\bar{u}_i$, that are defined as:

$$\bar{u}_i = \frac{1}{n}\sum_{j=1}^{n} u_i(t_j) \qquad \text{for } i = 1,.., N_s \tag{2}$$

At time t_j, the vector of the normalized measurements is:

$$\mathbf{u}(t_j) = \left[u_1(t_j) - \bar{u}_1 \quad u_2(t_j) - \bar{u}_2 \quad \cdots \quad u_{Ns}(t_j) - \bar{u}_{Ns}\right] \tag{3}$$

The N_sxN_s covariance matrix, $\mathbf{R}_{uu}$, among all measurements locations (or sensors) summed over all time samples is given by:

$$\mathbf{R}_{uu}(t_n) = \sum_{j=1}^{n} \mathbf{u}(t_j)^T \mathbf{u}(t_j) \tag{4}$$

The eigenvalues, λ_i, and eigenvectors, $\boldsymbol{\psi}_i$, of the covariance matrix satisfy:

$$\left(\mathbf{R}_{uu} - \lambda_i \mathbf{I}\right)\boldsymbol{\psi}_i = 0 \qquad \text{for } i = 1,..., N_s \tag{5}$$

where $\mathbf{I}$ denotes the $N_s \times N_s$ identity matrix. Using the fact that the covariance matrix $\mathbf{R}_{uu}$ of the data set is real, symmetric and of order N_s, there exist exactly N_s real eigenvalues with N_s real orthogonal eigenvectors $\boldsymbol{\psi}_i$ (an eigenvector is also called a *principal component*). Sorting the eigenvalues in decreasing order, $\lambda_1 > \lambda_2 > ... > \lambda_{Ns}$, the eigenvectors $\boldsymbol{\psi}_i$ represent the most persistent time function with the greatest variance. The first component corresponds to the direction in which the projected observations have the largest variance. The second component is then orthogonal to the first and again maximizes the variance of the data points projected on it. Continuing in this way, it is possible to compute all the principal components, which are the eigenvectors of the covariance matrix. To reduce the N_s-dimensional vector $\mathbf{u}$(t) into a d-dimensional vector, $\mathbf{x}_d$(t), where $d < N_s$, $\mathbf{u}$(t) is projected onto the eigenvectors that correspond to the d largest eigenvalues:

$$\mathbf{x}_v(t) = \left[\psi_1 \quad \cdots \quad \psi_d\right]^T \mathbf{u}(t) \tag{6}$$

Most of the variance is contained in the first few principal components while the remaining components are defined by measurement noise. For this reason, the analysis focuses on those components that contain most information. Generally, eigenvalues are time dependent and $\boldsymbol{\psi}_i$ are position dependent.

2.2 Correlation Analysis

This method is used to calculate the correlations on all sensor pairs for a reference period with the aim to quantify the tendency of values measured by the sensors to change in similar ways. During the reference period, variations of correlations are calculated for each sensor pair. After the initialisation phase, all correlations are calculated at each step of measurements to determine the presence of anomalies in the evolution of values. Usually anomalous behaviour is observed through correlations lying outside the thresholds defined during the initialisation phase. This parameter indicates if and by how much the behaviour of the structure changes regarding the reference period through observation those correlations outside of the valid zone. When anomalous behaviour is detected, the location of the damage is identified through identifying sensors which have the values for correlations outside of threshold values. The correlation at each step is calculated according the following formula:

$$C_{ij}(t) = \frac{\sum_{k=1}^{n}(S_i(t_k) - \overline{S}_i)(S_j(t_k) - \overline{S}_j)}{\sqrt{\sum_{k=1}^{n}(S_i(t_k) - \overline{S}_i)^2}\sqrt{\sum_{k=1}^{n}(S_j(t_k) - \overline{S}_j)^2}} \qquad \text{for } k=2,..,n \tag{7}$$

Where, n is the total number of time observation during the monitoring period, $Si(t_k)$ and $Sj(t_k)$ are the values of the sensors i and j at time t_k, $\overline{S}_i$ and $\overline{S}_j$ are the average values of the sensors i and j. In normal conditions this value should be constant or stationary. However when damage occurs, correlations between the sensors change.

2.3 Continuous Wavelet Transform (CWT)

Wavelet analysis is used to represent general functions in terms of simpler fixed building blocks at different scales and positions. A theoretical treatment of wavelets and wavelet analysis may be found in [21-22]. Using a selected analysing or mother wavelet function $\psi(t)$, the continuous wavelet transform of a function $Y(t)\in$L2(R), is defined as:

$$W(u,s) = \int_{-\infty}^{\infty} Y(t)\frac{1}{\sqrt{|s|}}\psi^*\left(\frac{t-u}{s}\right)dt \tag{8}$$

where u and s are real constants and * denotes the complex conjugate. s is a scale or dilation variable and u represents time shift. The translation parameter, u, indicates the location of the moving wavelet window in the wavelet transform. Shifting the wavelet along the axis implies examining $Y(t)$ in the neighborhood of the current window location.

Thus, information in the time domain remains, in contrast to Fourier transform where the frequency domain is used. The dilation parameter, s, indicates the width of the wavelet window. A smaller value of s implies a narrower wavelet window and a higher resolution. $W(u,s)$ is called the wavelet coefficient for the wavelet ψ and it measures the variation of the signal vicinity of the u whose size is proportional to s. This correlation between the signal and the wavelet is in the sense of frequency content. If the signal contains a spectral component corresponding to the current value of s, the products of the wavelet with the signals give relatively large values at locations where this spectral component exists.

Many types of different wavelet analyses have been applied to the data in order to select the most appropriate ones for the analysis. Gauss wavelets detect damage with the best resolution. Also, the accuracy of anomaly detection is better while increasing the scale, i.e. decreasing the frequency of the signal. Gauss wavelets detected damage with a scale of 128. For this application Gauss wavelets with a scale of 1024 were adopted in order to limit computational effort. CWT applied directly to the time series generated by the sensor produces a plot that is difficult to analyze due to periodic components of the time series, see Figure 3. For this reason CWT has been applied to the difference of the time series of the two sensors that are closest to the damage and normalized according the values of the first year.

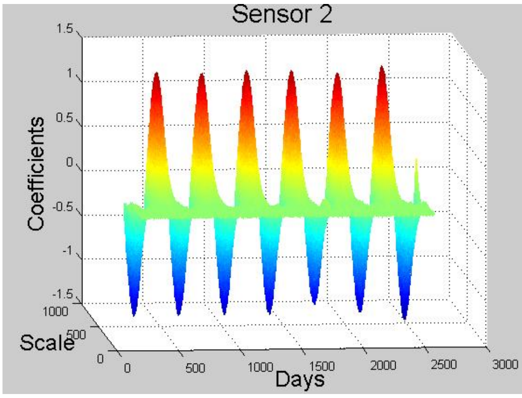


Figure 1 Plot of CWT calculated on sensor 2 in a damage scenario. The damage is simulated on 4 cells with reduction to 20% of original stiffness on sensor 2. Although something different happens the moment when damage occurred (1750 days), it is not clear whether the situation is temporary or definitive.

2.4 Short Term Fourier Transform (STFT)

Fourier Transforms (FTs) decompose a signal to complex exponential functions of different frequencies:

$$X(f) = \int_{-\infty}^{\infty} Y(t) \cdot e^{-2j\pi ft} dt \tag{9}$$

One of the problems of Fourier transforms is that while they indicate whether a certain frequency component exists or not, this information is independent of when this component appears. To use FTs, even with a non-stationary signal it is necessary to use a revised version of FT called Short Term Fourier Transform (STFT) [21]. In STFT, the signal is divided into small segments, where these segments (portions) of the signal are assumed to be stationary. A window function "w" is chosen, (a rectangle centred at t and having an amplitude of unity). The width of this window must be equal to the segment of the signal where its stationary property is valid. The formula used to compute the STFT is given by:

$$STFT_Y^{(w)}(t', f) = \int_{-\infty}^{\infty} Y(t) \cdot w^*(t - t') \cdot e^{-2j\pi ft} dt \quad (10)$$

Where $Y(t)$ is the signal, $w(t)$ is the window function, and * is the complex conjugate.

For the analysis, a window of two years has been chosen. Result of the STFT for different t's are analysed according to variations of the main components. The function, $Y(t)$ is the time series that is generated by the sensor reading.

2.5 Instance-Based Method (IBM)

The Instance-Based Method [23-26] consists of calculating, at each step, the minimum distance of a group of narrow sensors (normally 3 or 4) from the cloud of points of the training set. Damage is detected when there is a sequence of consecutive points whose values exceed their thresholds. For example, the distance calculated at the time t for the sensors: s_1, s_2, s_3 and a training set TS is given by:

$$\min_{i \in TS} \sqrt{(s_1(i) - s_1(t))^2 + (s_2(i) - s_2(t))^2 + (s_3(i) - s_3(t))^2} \quad (11)$$

3 Damage Detection Methodology

The methodology used for the damage detection includes PCA and Correlation Analysis algorithms that are adapted to follow the evolution of time series. If these algorithms are used in their original versions, a long training period is required to estimate the threshold values that are necessary to detect the presence of anomalous behaviors. To solve this problem, the two algorithms have been improved to reduce the time required for convergence. The two modified algorithms, called Moving Principal Components Analysis (Moving PCA) and Moving Correlation Analysis, are described in detail in the next sub-sections.

3.1 MOVING PCA

When equations 1-6 are used directly, there are two drawbacks:

- The time necessary to compute the covariance matrix and the principal components increases with the number of measurements
- There is a delay in detecting a new situation in the time series. This problem is due to the fact that with the increasing of the number of measurements, the effect of new points in the covariance matrix is lower and lower because they are averaged by the total number of points. Previous work has proposed solutions to this problem through following the evolution of series over time, usually through recursive strategies [28].

This paper investigates computations using PCA for fixed-sized windows that move in time. More specifically, a new method, Moving Principal Components Analysis (MPCA) that computes the covariance matrix inside a moving window of constant size is proposed. Once the dimension of the window is fixed, the number of points N_w (measurements inside the window) is constant. With this solution, the formulas used to calculate the mean and covariance matrix become:

$$\overline{u}_i = \frac{1}{N_w} \sum_{j=n-N_w}^{n} u_i(t_j) \qquad \text{for } n>N_w \tag{12}$$

$$\mathbf{R}_{uu}(t_n) = \sum_{j=n-N_w}^{n} \mathbf{u}(t_j)^T \mathbf{u}(t_j) \tag{13}$$

This means that at each step, it is necessary to calculate parameters only for the points inside the active window. The advantages of the use of a moving window are the following:

- It is possible to calculate process parameters more rapidly
- The computation time of PCAs at each step is constant being a function only of N_w
- Detection of the presence of new situations is more timely because the old measurements do not buffer results

A key issue is selecting the dimension of the window (N_w). If the process is stationary it is necessary to select a value N_w that is sufficiently large so that it is not influenced by measurement noise while providing enough rapidity for detecting new behavior. If the time series has periodic variability (for example, due to temperature cycles) the choice of the window size should be a multiple of the period. This choice ensures that mean values are stationary over time and that eigenvalues of the covariance matrix do not have periodic behavior. When damage occurs, mean values and components of the covariance matrix thus change and as consequence, eigenvalues and eigenvectors. In this study, a two-year window is chosen so that it includes the one-year environmental cycle and leaves extra time to estimate thresholds and to accommodate periods when data is missing. Good estimates of thresholds lower the chances of false alarms.

3.2 Moving Correlation Analysis

Normally the correlation is calculated for all available measurements. To follow the evolution of the time series more effectively, a moving window of fixed size is employed. Equation 7 is updated so that it can be calculated only for the last N_w points:

$$C_{ij}(t_n) = \frac{\sum_{k=n-Nw}^{n} (S_i(t_k) - \overline{S}_i)(S_j(t_k) - \overline{S}_j)}{\sqrt{\sum_{k=n-Nw}^{n} (S_i(t_k) - \overline{S}_i)^2} \sqrt{\sum_{k=n-Nw}^{n} (S_j(t_k) - \overline{S}_j)^2}} \qquad \text{for } n>N_w \tag{14}$$

$C_{ij}(t_n)$ is the correlation calculated for the couple of sensor sensors i,j at the time t_n only for the last N_w measurements. For periodic or quasi-periodic time series N_w should be a multiple of the period. For other types of time series, this value should be selected to guarantee stability of average values, to ensure rapid damage identification and to reduce the effects of noise.

3.3 Other methods

Moving windows are inherent in the CWT and STFT methods. For the IBM, no moving window was considered because initial studies indicated that sensitivity dropped significantly when a window was implemented. Therefore, all methods except IBM were compared in a moving window implementation.

4 Application to Structural Health Monitoring

Application of MPCA and Moving Correlation to structural health monitoring involves an initial phase (called initialization) where the structure is assumed to behave in an undamaged condition. The aim of this initialization period is to estimate the variability of the time series and to define thresholds for detecting anomalous behavior. This period is normally one or two years. Once thresholds have been fixed, the parameters of the process are monitored (main eigenvectors for MPCA and Correlations on all sensor pairs for the Moving Correlation) to ensure that they are inside predefined ranges. For damage localization the rule has been used that candidate damage zones are close to sensors that have measurement values exceeding a threshold.

4.1 Numerical Simulation

Due to difficulties in retrieving databases from real structures with a range of damage severities, a finite element model of a beam, studied by University of Genoa [10] has been used to evaluate the efficiency of algorithms to detect damage. The main aim of the numerical simulation task is to simulate the behavior of a bridge (two span continuous beam) in healthy and various damaged states [15]. A thermal load simulates structural behavior under varying environmental conditions.

In the FE model, it is possible to simulate both thermal and applied moving loads as well as damaged elements in one or more sections of the beam with stiffness reductions. The response is measured by means of a ‘virtual’ monitoring system installed in the structure. As shown in Figure (1), the monitoring system is composed of six pairs of elongation sensors; six sensors are located at the lower surface and six at the upper surface. Cracks were simulated in order to model local degradation in material properties. The structural behavior has been simulated both in healthy and in damaged states to test the ability of statistical algorithms to detect when and where damage occurs. Various locations and severity of damage have been simulated, see Figure (1).

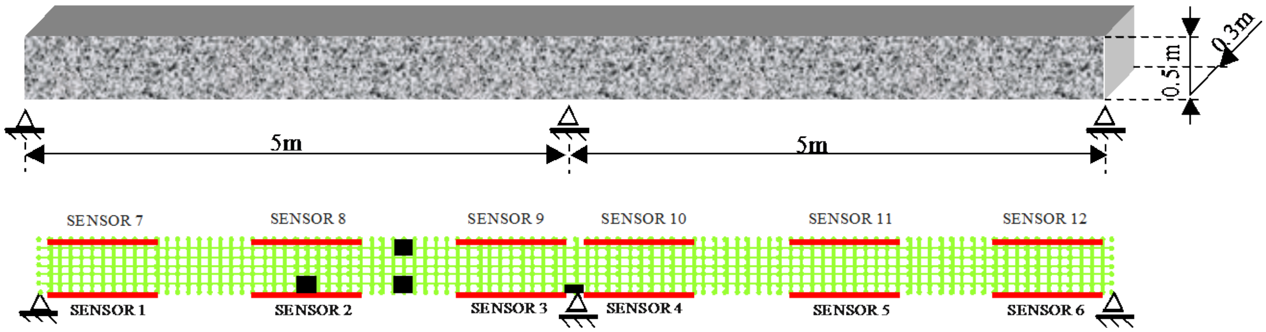


Figure 1 The FE model used to test the methodology. The lower part of the figure is a section showing where sensors are placed (red lines) and positions of simulated damage (black squares). In subsequent figures, sensors are referred to as sn1, sn2, etc.

Each simulated time series is representative of structural response measured at a given sensor location over eight years, assuming four measurements per day (11000 measurement events). The time series show harmonic variations due to seasonal variations of temperature in both healthy and damaged states, see Figure 2. However, no obvious correlation to damage formation has been observed directly from the elongation-sensor time histories. Only in cases of severe damage has a small shift of periodicity been observed. As a consequence, statistical analysis techniques that are described next have been applied to the data.

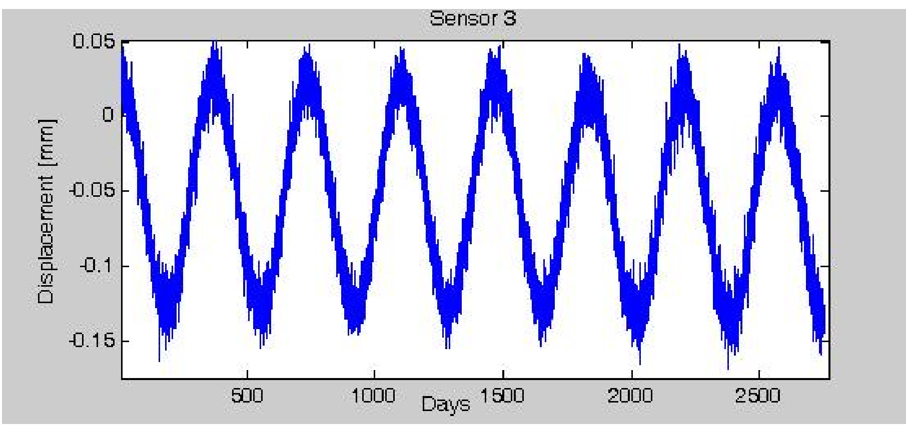


Figure 2 Typical sensor behavior. Due to seasonal effects, periodic behavior is visible. Four measurements per are taken (one every six hours). This figure presents the behavior of the sensor 2 when damage occurred after 1750 days. In the graph x-axis represents time in days and y-axis represents displacement in millimeters.

The example in this paper (Figure 1) has harmonic variability due to the seasons and therefore, the training period has been defined to be two years and the size of the moving window is one year. Thus the value of N_w is the number of measurements performed in one year. Once a sequence of values exceed a threshold, two scenarios are possible:

- The anomalous situation is temporary due, for example, to extreme conditions (hot summer, cold winter…) and after a short period measurement values come back to normal values

- This situation is definitive meaning that something is changed in a part of the structure and as consequence its daily behavior. Once this situation is detected and working conditions are stable again (for example no further damage), it is possible to redefine a training period in which thresholds for the part of structured involved in the damage are recalculated.

The MPCA algorithm has been applied to the twelve time series in the example, one for each sensor, in order to simulate measurement evolution during monitoring phases. MPCA uses all temporal values inside the active window to compute the covariance matrix and at each temporal step the window is shifted; that is, one session of measurements is added and the oldest is removed. This operation is repeated for all measurement sessions. The time series have been arranged to form the matrix **U**(t) with twelve columns and N_w rows. Each column represents a displacement history at a particular sensor location. Alternatively, each row represents the spatial distribution of the response at a given time instant.

After each measurement session, the covariance matrix and principal components have been computed. Each eigenvalue expresses the variance in time associated with the corresponding eigenvector. Orthogonal eigenvectors are time-invariant: the eigenvector associated with the maximum eigenvalue represents the spatial behavior corresponding to the time function with the maximum variance. The Moving Correlation algorithm has been applied to the time history of all sensor pairs with a moving window of one year.

Two algorithms are useful for two reasons. Firstly, using two algorithms at the same time helps reduce the likelihood of false positives and false negatives. Secondly, the use of two algorithms provides adaptability to new situations while remaining sensitive to the occurrence of damage.

The algorithms have been tested in the following damage scenarios:

a) Damage in 4 cells with reduction to 20% of original stiffness at sensor 2, Figure 4
b) Damage in 2 cells with reduction to 20% of original stiffness at sensor 2, Figure 5
c) Damage in 1 cell with reduction to 50% of original stiffness at sensor 2, Figure 6
d) Damage in 4 cells with reduction to 20% of original stiffness between sensors 2 and 3, Figure 7

Cells are the dimensions of the finite element size. In this study, each element is 10 x 8 centimeters. Damage has been introduced for all scenarios at the same time.

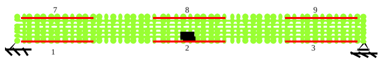


Figure 4, Scenario A : damage in 4 cells with reduction to 20% of original stiffness at sensor 2

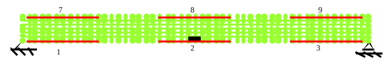


Figure 5, Scenario B : damage in 2 cells with reduction to 20% of original stiffness at sensor 2

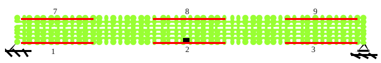


Figure 6, Scenario C : damage in 1 cell with reduction to 50% of original stiffness at sensor 2

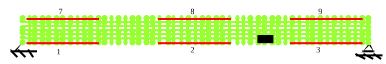


Figure 7, Scenario D : Damage in 4 cells with reduction to 20% of original stiffness between sensors 2 and 3

5 Results

Results of a comparative study between the proposed algorithms and other methods are presented in Table 1. Other methods are continuous wavelet transform (CWT), short-term Fourier transform (SFT) and Instance-Based Method (IBM). Damage is said to be detected when there is a significant statistical variation (97% confidence) in the parameters. The initial reference period is used to calculate mean and standard deviation values for parameters.

Results show that the algorithms proposed in this paper work more effectively than other methods. This does not mean that the other methods are inferior to the proposed algorithms for all applications. The results of Table 1 lead to the conclusion that MPCA and Moving Correlation are more appropriate for long-term structural health monitoring tasks such as the example simulated in this study.

Damage Scenario	**Traditional Methods**			**Innovative Methods**	
	CWT	**SFTF**	**DTS**	**MPCA**	**MOVING CORRELATION**
A	D	D	D	D	D
B	D	D	D	D	D
C	ND	ND	ND	D	ND
D	ND	ND	ND	D	D

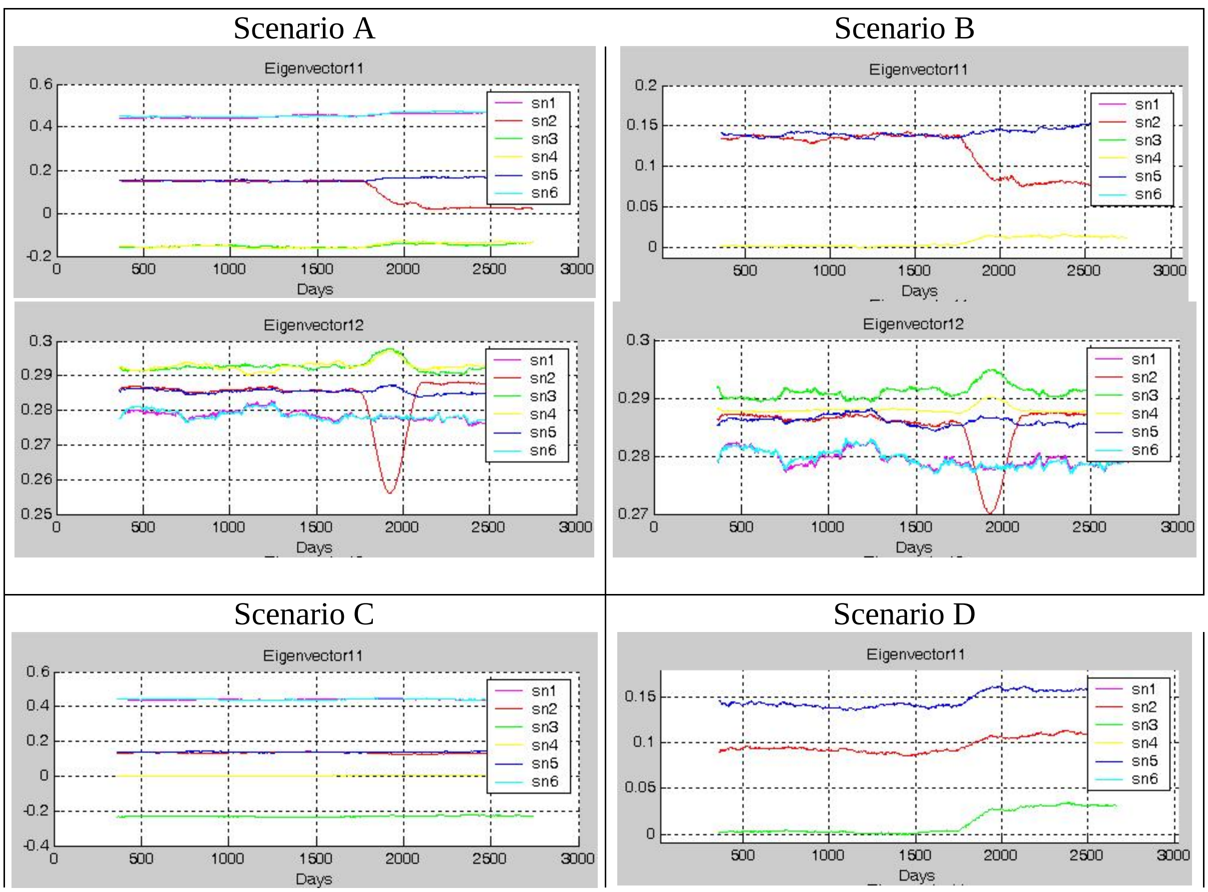
Table 1 Results of a comparative study between the algorithms proposed in this paper and other methods using data derived from a numerically simulated beam in healthy and various damages states. D = detected, ND = not detected

In Figure 8, MPCA diagnostic plots of the eigenvectors related to the main eigenvalues in all the damage scenarios are shown. The moment when damage occurs and its location are visible in all graphs through following the evolution of the two main eigenvectors. Specifically, one of the eigenvectors (eigenvector 11) indicates a new state and when it becomes stable, while the other (eigenvector 12) indicates when the damage occurred.

The location of the damage is detected by the fact that within the main eigenvectors, there are one or more rapidly changing components that are associated with sensors close to the damage.

In Figure 9, plots of the Moving Correlation related to the pair of sensors closest to the damage in all the damage scenarios are shown. In all the graphs the moments when damage occurred are visible. When the behavior of the structure can be considered to be stable, evidence is visible only for scenarios a), b) and d). This is when the sensor data is ready for new training (only for the sensor involved in the damage). In both Figures (8 and 9) the two algorithms are able to detect relatively limited damage.

In Figure 10, Distance-from-Training-Set diagnostic plots are shown. The moment when the damage occurred is clearly visible only for scenarios a) and b), while it is not visible for the other two. In Figures 11 and 12, the STFT and CWT diagnostic plots are presented. In the STFT diagnostic plots show the presence of the damage clearly for scenarios a) and b) and only slightly for scenarios c) and d). For the CWT, the moment when the damage occurred is only visible for scenarios a) and b). There is no information regarding whether the anomalies are due to a new temporary situation or to permanent damage. Although this does not mean that the CWT is not useful for the analysis of long term static monitoring time series, they cannot be applied directly to time series. Instead they could be used as support for the damage identification on the results obtained with the MPCA and Moving Correlation Analysis.



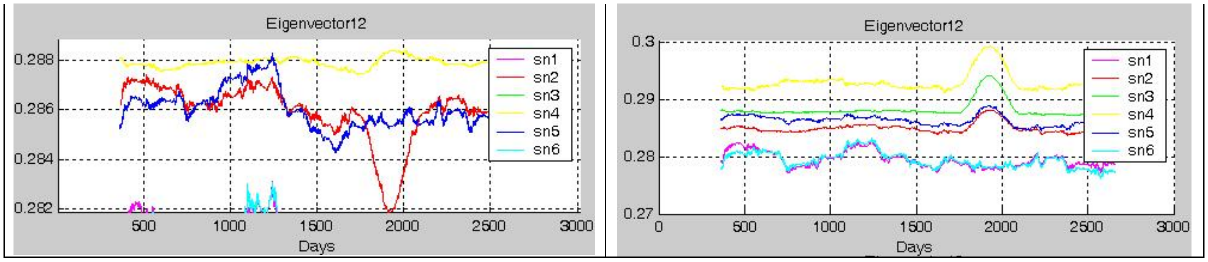


Figure 8 MPCA plots of eigenvectors related to the two main eigenvalues. They show the moment when damage occurs and its location. One eigenvector (eigenvector 11) gives gives an indication of the new state of the structure when it becomes stable while the second eigenvector (eigenvector 12) gives an indication of the damage at 1750 days. In all the graphs x-axis represents time in days and y-axis represents the eigenvector.

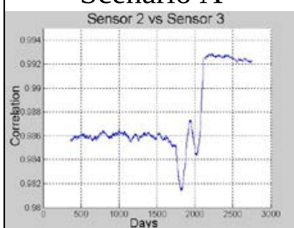
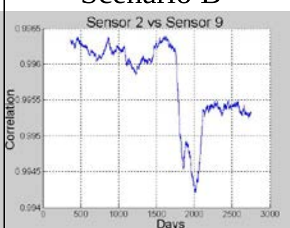
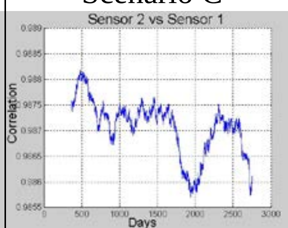
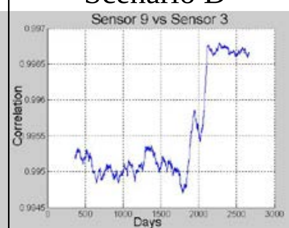
Scenario A	Scenario B	Scenario C	Scenario D
			

Figure 9 Diagnostic plots of Moving Correlation calculated from measurements of two sensors close to the damage. They show the moment when the damage occurred (1750 days) and the moment when the algorithm starts to adapt itself to the new state. Calculations were performed using a moving window of one year. In all the graphs x-axis represents time in days and y-axis represents correlation.

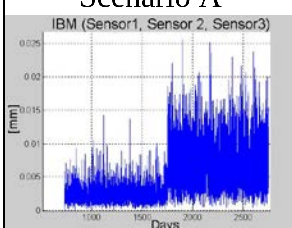
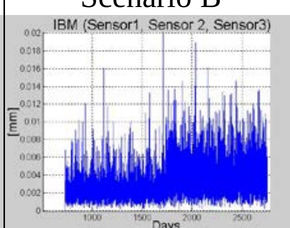
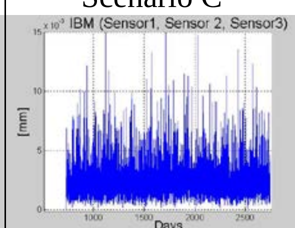
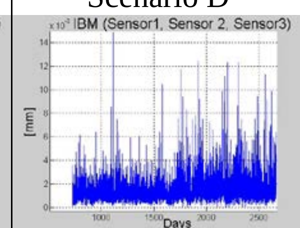
Scenario A	Scenario B	Scenario C	Scenario D
			

Figure 10 Plots of Instance-based method calculated for the sensors closest to the damage. They show the moment when damage occurred and that the situation is permanent and not temporary. The training set is composed by all the measurements done during the first two years. In all the graphs x-axis represents time in days and y-axis represents distance in millimeters.

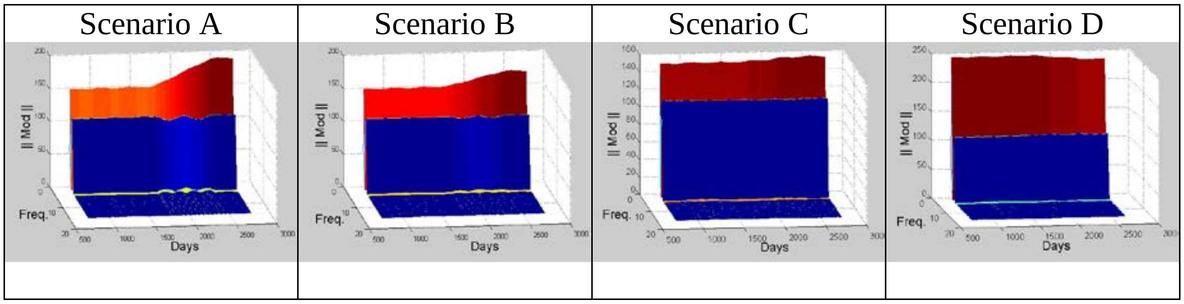


Figure 11 Plots of STFT calculated for the sensor closest to the damage. They show the moment when damage occurred and when the state of the structure is stable again. A moving window of two years was used. In all the graphs x-axis represents time in days, y-axis represents the frequency and the z-axis represents the modulus.

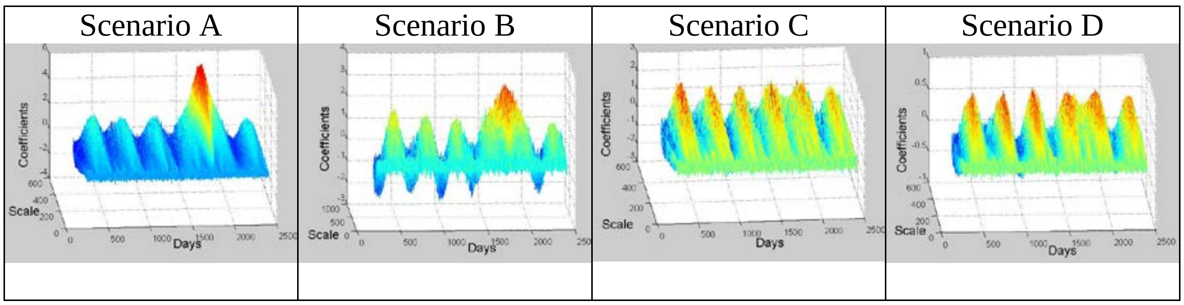


Figure 12 CWT calculated from the difference between results of the two sensors closest to the damage. Although the results show when damage occurred, there is no indication whether the situation evolves temporarily or definitively. In all the graphs x-axis represents time in days, y-axis represents the scale and the z-axis represents the coefficients of the wavelets.

5 Conclusions

Moving principal component analysis and moving correlation are useful tools for identifying and localizing anomalous behavior in civil engineering structures. These approaches can be applied over long periods to a range of structural systems in order to discover anomalous states even when there are large quantities of data. A comparative study has shown that for quasi-static monitoring of civil structures, these new methodologies perform better than wavelet methods, short term Fourier transform and the instance-based method. While these methodologies have good capacities to detect and locate damage, they also require low computational resources. Another important characteristic is adaptability. Once new behavior is identified, adaptation allows detection of further anomalies. The next step of the research is to apply the proposed methodology to a database of measurements taken from full-scale structures.

References

1. L.A. Bisby, An Introduction to Structural Health Monitoring. ISIS Educational Module 5 (2005)
2. P. Cawley and R.D. Adams, The location defects in structures from measurements natural frequencies, Journal of Strain Analysis 14 (1979) 49-57
3. E. Douka, S. Doutridis and A.Trochidis, Crack identification in plates using wavelet analysis, Journal of Sound and Vibration 270 (2004) 279-295
4. J.-T. Kim, Y.-S.Ryu, H.-M. Cho and N.Strubbs, Damage identification in beam-type structures: frequency based method vs mode-shape method, Engineering Structures 25 (2003) 57-67
5. J.-C. Hong, Y.Y. Kim, H.C. Lee and Y.W. Lee, Damage detection using Lipschitz exponent estimated by wavelet transform: applications to vibration modes of beam, International Journal of Solid and Structures 39 (2002), 1803-1846
6. A. Messina, Detecting damage in beams through digital differentiator filters and continuous wavelet transform, Journal of Sound and Vibration 272 (2004) 385-412
7. M. Rucka and K. Wilde, Crack identification using wavelets on experimental static deflection profiles, Engineering Structures, Volume 28, Issue 2, January 2006, 279-288
8. A. Gentile and A. Messina, On the continuous wavelet transforms applied to discrete vibration data for detecting open cracks in damaged beams, International Journal of Solid and Structures 40 (2003) 295-315
9. Y. Robert-Nicoud, B. Raphael, O. Burdet & I. F. C. Smith, Model Identification of Bridges Using Measurement Data, Computer-Aided Civil and Infrastructure Engineering, Volume 20 Page 118 - March 2005
10. F. Lanata Damage detection algorithms for continuous static monitoring of structures PhD Thesis Italy University of Genoa DISEG, (2005)
11. S. Saitta, B. Raphael, I.F.C. Smith, Data mining techniques for improving the reliability of system identification, Advanced Engineering Informatics 19 (2005) 289–298
12. S. F. Masri, L-H Sheng, J. P. Caffrey, R. L. Nigbor, M. Wahbeh and A. M. Abdel-Ghaffar, 2004 Application of a Web-enabled real-time structural health monitoring system for civil infrastructure systems Smart Mater. Struct. 13 1269-83
13. E. R. Biondani, C. Vardanega, G. Vaccaro, E. Mirone and E. Saveri, Observed and predicted settlements of two Italian nuclear power plants Trans. 9th Int. Conf. On Structural Mechanics in Reactor Technology (Balkema, Rotterdam), (1987) D 293-8
14. P. Omenzetter, J. M. W. Brownjohn and P.Moyo Identification of unusual events in multi channel bridge monitoring data Mech. Syst. & Sign. Proc.,(2004) 18 409-30
15. A. Del Grosso, D. Inaudi and F. Lanata Strain and displacement monitoring of a quay wall in the Port of Genoa by means of fibre optic sensors 2nd Europ. Conf. on Structural Control Paris, (2000)
16. A. Del Grosso and L. Lanata, Data analysis and interpretation for long-term monitoring of structures Int. J. for Restoration of Buildings and Monuments, (2001) 7 285-300
17. J. BROWNJOHN, S. C. TJIN, G. H.TAN, B. L. TAN, S. CHAKRABOORTY, "A Structural Health Monitoring Paradigm for Civil Infrastructure", 1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering, Nottingham, United Kingdom, 28 June – 1 July 2004
18. H. M. Jaenisch, J. W. Handley, J. C. Pooley, S. R. Murray, "DATA MODELING FOR FAULT DETECTION" , *2003 MFPT Meeting*.
19. F. Lanata and A. Del Grosso, Damage detection algorithms for continuous static monitoring: review and comparison 3rd Europ. Conf. On Structural Control (Wien, Austria), 2004
20. Sohn, H., J. A.Czarneski and C. R. Farrar. 2000. "Structural Health Monitoring Using Statistical Process Control", Journal of Structural Engineering, 126(11): 1356-1363
21. C. K. CHUI, Introduction to Wavelets, San Diego, CA: Academic Press, p.264, 1992
22. I. Daubechies, Ten Lectures on Wavelets, Philadelphia: Society for Industrial and Applied Mathematics, p. 357, 1992
23. Kaufman and Rousseeuw, 1990,L. Kaufman and P.J. Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York, 1990.
24. Mahamud and Hebert, S. Mahamud and M. Hebert. Minimum risk distance measure for object recognition. In Proceedings of the ninth IEEE International, 2003 Conference on Computer Vision (ICCV), pages 242–248, 2003.
25. Markou and Singh, M. Markou and S. Singh. Novelty detection: a review, Signal Processing, pages 2481–2521, 2003.
26. Ilmari Juutilainen and Juha Roning,Measuring Distance from a Training Data Set
27. M. Hubert and S. Verboveny, "A robust PCR method for high-dimensional regressors", Journal of Chemometrics, 17, 438-452.
29. M. Hubert , P. J. Rousseeuwy and K. V. Brandenz "ROBPCA: a New Approach to Robust Principal Component Analysis", Technometrics, Vol. 47, No. 1, February 2005 pp.64-79