Original Article

SHM Structural Health Monitoring

Structural Health Monitoring
1–16

DOI: 10.1177/1475921715624504
shm.sagepub.com
SAGE

Detection of structural changes through principal component analysis and multivariate statistical inference

Francesc Pozo[1], Ignacio Arruga[2], Luis E Mujica[1], Magda Ruiz[1] and Elena Podivilova[3]

Abstract

This article introduces a new methodology for the detection of structural changes using a statistical data-driven modeling approach by means of a distributed piezoelectric active sensor network at different actuation phases. The three main features that characterize the proposed methodology are (a) the nature of the data used in the test since vectors of principal component analysis projections are used instead of the entire measured response of the structure or the coefficients of an AutoRegressive model, (b) the number of data used since the test is based on two random samples instead of some characteristic indicators, and (c) the samples come from a multidimensional variable and therefore a test for the plausibility of a value for a normal population mean vector is performed. The framework of multivariate statistical inference is used with the objective of the classification of structures in healthy or damaged. The novel scheme for damage detection presented in this article —based on multivariate inference over the principal component analysis projections of the raw data—is applied, validated, and tested on a small aluminum plate. The results show that the presented methodology is able to accurately detect damages, that is, for each actuation phase, a unique and reliable damage detection indicator is obtained no matter the number of sensors and/or actuators. It is worth noting that a major contribution of this article is that there exists an entire range of significance levels where the multivariate statistical inference is able to offer a correct decision although all of the univariate tests make a wrong decision.

Keywords

Damage detection, principal component analysis, multivariate hypothesis testing, structural health monitoring

Introduction

Detection and classification of damages are possible by comparing the in-service dynamic time responses of a structure with respect to baseline reference responses recorded in ideal and healthy operating conditions.[1,2] These signals—even in stable environmental and operational conditions—present the main characteristic that they cannot be repeated exactly due to random measurement errors. Therefore, these signals can be considered as a *random variable*. This way, a set of dynamic responses gathered from several experiments can be defined as a *sample variable* and, all possible values of the dynamic response as the *population variable*. Therefore, the process to draw conclusions about the state of the structure from several experiments using statistical methods is usually named as statistical inference for damage diagnosis. In structural health monitoring (SHM) field, statistical inference can be considered as one of the emerging technologies that will have an impact on the damage prognosis process.[3,4]

Some works developed in the field of statistical data-driven are focused on detecting, localizing, and classifying damages in a wide collection of structures

[1]CoDAlab, Departament de Matemàtiques, Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona (EUETIB), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
[2]Departamento de Seguridad Pasiva, IDIADA Automotive Technology SA, Polígon Industrial L'Albornar, Tarragona, Spain
[3]Control Systems Department, South Ural State University (National Research University), Chelyabinsk, Russia

Corresponding author:
Francesc Pozo, CoDAlab, Departament de Matemàtiques, Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona (EUETIB), Universitat Politècnica de Catalunya (UPC), Comte d'Urgell, 187, E-08036 Barcelona, Spain
Email: francesc.pozo@upc.edu

by means of different experimental and numerical measurements. All of the approaches are based on multivariate statistical analysis.[5–13] However, to improve the accuracy of the expected results, and since these methods by themselves are unable—in some cases—to correctly classify different scenarios, the methodologies had to be enhanced using more complex and computationally expensive algorithms. Following the same goal but reducing significantly the computational cost, this article is focused on the development of a damage detection indicator—considering dynamic responses as random variables as in Mujica et al.[14]—that combines a data-driven baseline model (reference pattern obtained from the healthy structure) based on principal component analysis (PCA) and univariate hypothesis testing. As said earlier, the use of hypothesis testing is not new in this field. The contribution of the previous work[14] was based on (a) the nature of the data used in the test since we are using scores instead of the measured response of the structure or the coefficients of an AutoRegressive model,[15] (b) the number of data used since the test is based on two random samples instead of two characteristic quantities,[16] and (c) the samples come from a unidimensional variable and therefore univariate statistical hypothesis testing is performed. Although the overall accuracy of the damage detection was quite acceptable, the sensitivity of the approach is dependent on the chosen score and the transducer used as actuator. Therefore, a more complex tool is needed—but keeping a low computational cost—to combine the individual diagnoses to infer a more precise and general damage detection. With this issue in mind, the novelty of this article is the detection of structural changes through multivariate statistical inference. This way, for each actuation phase, a unique and more reliable damage detection indicator is obtained.

More precisely, the proposed development starts obtaining the baseline PCA model and the subsequent projections using the healthy structure. When the structure needs to be inspected, new experiments are carried out and they are projected onto the baseline PCA model. Each experiment is considered as a random process and the projections onto a predetermined number of principal component is a multivariate random variable. The objective is to analyze whether the distribution of the multidimensional variable associated with the current structure is related to the healthy one.

This article is organized as follows. Section "Data driven-baseline model based on PCA" describes how the baseline model is built using PCA. In section "Detection of structural changes based on multivariate statistical inference," the damage detection based on multivariate hypothesis testing is developed. Second, a multivariate hypothesis test is formulated yielding to the structural damage indicator. Section "Experimental results" presents the experimental results showing for the test the level of sensitivity, specificity, and reliability. Finally, some conclusions are drawn in section "Concluding remarks."

Data-driven baseline model based on PCA

PCA: theoretical background

Let us initiate the analysis of a physical process by measuring different variables (sensors) at a finite number of time instants. In this work, the collected data are arranged in a $n\times(N\cdot L)$ matrix as follows

$$X = \begin{pmatrix} x_{11}^1 & x_{12}^1 & \cdots & x_{1L}^1 & x_{11}^2 & \cdots & x_{1L}^2 & \cdots & x_{11}^N & \cdots & x_{1L}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1}^1 & x_{i2}^1 & \cdots & x_{iL}^1 & x_{i1}^2 & \cdots & x_{iL}^2 & \cdots & x_{i1}^N & \cdots & x_{iL}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1}^1 & x_{n2}^1 & \cdots & x_{nL}^1 & x_{n1}^2 & \cdots & x_{nL}^2 & \cdots & x_{n1}^N & \cdots & x_{nL}^N \end{pmatrix} \tag{1}$$

Matrix $\mathbf{X} \in \mathscr{M}_{n\times(N\cdot L)}(\mathbb{R})$—where $\mathscr{M}_{n\times(N\cdot L)}(\mathbb{R})$ is the vector space of $n\times(N\cdot L)$ matrices over $\mathbb{R}$—contains data from N sensors at L discretization instants with respect to n experimental trials. Consequently, each row vector $\mathbf{X}(i,:) \in \mathbb{R}^{N\cdot L}$, $i=1,\ldots,n$, represents, for a specific experimental trial, the measurements from all the sensors at every specific time instant. Equivalently, each column vector $\mathbf{X}(:,j) \in \mathbb{R}^n$, $j=1,\ldots,N\cdot L$, represents measurements from one sensor at one particular time instant in the whole set of experimental trials.

The main objective of PCA is to distinguish which dynamics are more relevant in the system, which are redundant and which can be considered as a noise.[8] This objective is essentially accomplished by defining a new coordinate space to re-express the original one. This new coordinate space is used to filter noise and redundancies according to the variance-covariance matrix of the original data. In other words, the objective is to find a linear transformation orthogonal matrix $\mathbf{P} \in \mathscr{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ that will be used to transform the original data matrix $\mathbf{X}$ according to the following matrix multiplication

$$\mathbf{T} = \mathbf{XP} \in \mathscr{M}_{n\times(N\cdot L)}(\mathbb{R}) \tag{2}$$

Matrix $\mathbf{P}$ is usually called the principal components of the data set or loading matrix and matrix $\mathbf{T}$ is the transformed or projected matrix to the principal component space, also called score matrix. Using all the $N\cdot L$ principal components, that is, in the full

dimensional case, the orthogonality of $\mathbf{P}$ implies $\mathbf{P}\mathbf{P}^T = \mathbf{I}$, where $\mathbf{I}$ is the $(N \cdot L) \times (N \cdot L)$ identity matrix. Therefore, the projection can be inverted to recover the original data as

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T$$

Matrix $\mathbf{P}$ can be computed by means of the singular value decomposition (SVD) of the covariance matrix defined in equation (3). Then, the principal components are defined by the eigenvectors and eigenvalues of the covariance matrix as follows

$$\mathbf{C}_\mathbf{X} = \frac{1}{N \cdot L - 1}\mathbf{X}^T\mathbf{X} \in \mathscr{M}_{(N \cdot L)\times(N \cdot L)}(\mathbb{R}) \quad (3)$$

$$\mathbf{C}_\mathbf{X}\mathbf{P} = \mathbf{P}\Lambda \quad (4)$$

where the columns of $\mathbf{P}$ are the eigenvectors of $\mathbf{C}_\mathbf{X}$. The diagonal terms of matrix Λ are the eigenvalues $\lambda_i, \quad i = 1, \ldots, N \cdot L$, of $C_\mathbf{X}$, whereas the off-diagonal terms are zero, that is,

$$\Lambda_{ii} = \lambda_i, \quad i = 1, \ldots, N \cdot L$$

$$\Lambda_{ij} = 0, \quad i, j = 1, \ldots, N \cdot L, \quad i \neq j$$

The eigenvectors, $p_j, \quad j = 1, \ldots, N \cdot L$, representing the columns of the transformation matrix $\mathbf{P}$ are classified according to the eigenvalues in descending order and they are called the *principal components* of the data set. The eigenvector with the highest eigenvalue, called the *first principal component*, represents the most important pattern in the data with the largest quantity of information.

However, the objective of PCA is, as said before, to reduce the dimensionality of the data set $\mathbf{X}$ by selecting only a limited number $\ell < N \cdot L$ of principal components, that is, only the eigenvectors related to the ℓ highest eigenvalues. Thus, given the reduced matrix

$$\hat{\mathbf{P}} = (p_1 | p_2 | \cdots | p_\ell) \in \mathscr{M}_{(N \cdot L)\times\ell}(\mathbb{R})$$

matrix $\hat{\mathbf{T}}$ is defined as

$$\hat{\mathbf{T}} = \mathbf{X}\hat{\mathbf{P}} \in \mathscr{M}_{n\times\ell}(\mathbb{R})$$

Note that opposite to $\mathbf{T}$, $\hat{\mathbf{T}}$ is no longer invertible. Consequently, it is not possible to fully recover $\mathbf{X}$, although $\hat{\mathbf{T}}$ can be projected back onto the original $(N \cdot L)-$dimensional space to get a data matrix X as follows

$$\mathbf{X} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T \in \mathscr{M}_{n\times(N \cdot L)}(\mathbb{R}) \quad (5)$$

The difference between the original data matrix X and X is defined as the *residual error matrix* $\mathbf{E}$ as follows

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} \quad (6)$$

or, equivalently

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T + \mathbf{E} \quad (7)$$

The residual error matrix $\mathbf{E}$ describes the variability not represented by the data matrix $\hat{\mathbf{X}}$.

Although the real measures obtained from the sensors as a function of time represent physical magnitudes, when these measures are projected and the scores are obtained, these scores no longer represent any physical magnitude.[14] The key point in this approach is that the scores from different experiments can be compared with the reference pattern to try to detect a contrasting behavior.

PCA modeling

For the PCA modeling stage, we carry out a set of experiments as stated in section "Experimental set-up." For each different phase (piezoelectric transducer 1 (PZT1) will act as an actuator in phase 1, PZT2 will act as an actuator in phase 2, and so on) and considering the signals measured by the sensors, the matrix $\mathbf{X}_\mathrm{h}$ is defined and arranged as in equation (1) and scaled as stated in Mujica et al.[8] PCA modeling basically consists of computing the projection matrix $\mathbf{P}$ for each phase as in equation (2). Matrix $\mathbf{P}$, renamed $\mathbf{P}_\mathrm{model}$, provides an improved and dimensionally limited representation of the original data $\mathbf{X}_\mathrm{h}$. $\mathbf{P}_\mathrm{model}$ is considered as the model of the healthy structure to be used to detect structural damage. The modeling stage is graphically represented in Figure 1.

Detection of structural changes based on multivariate statistical inference

A predetermined number of experiments are performed in the structure to be diagnosed and a new data matrix $\mathbf{X}_\mathrm{c}$ is constructed with the recorded data, as inequation (1). The number of experiments is not limited *a priori*. However, the number of sensors and recorded samples must correspond with the number of sensors and recorded samples in the PCA modeling stage; more precisely, the number of columns of $\mathbf{X}_\mathrm{c}$ and $\mathbf{X}_\mathrm{h}$ must agree. Matrix $\mathbf{X}_\mathrm{c}$ will be projected onto the PCA model as specified in section "Multivariate random variables and multivariate random samples." The projections onto the first components—the so-called *scores*—are used for the construction of the multivariate random samples to be compared and consequently to obtain the structural damage indicator, as illustrated in Figure 2.

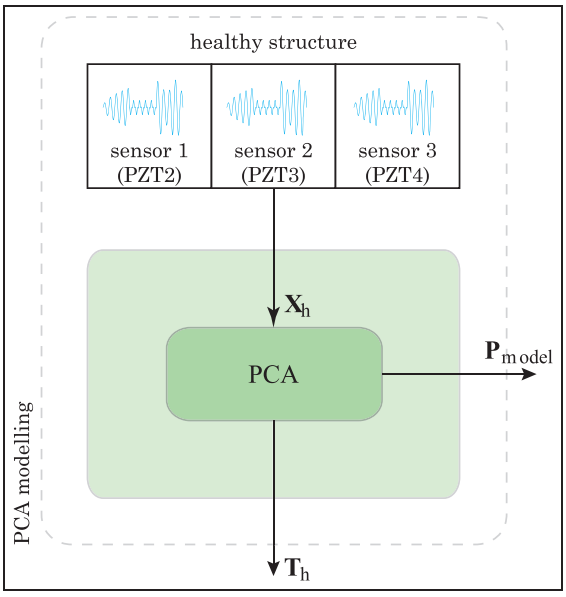

Figure 1. A principal component analysis model $\mathbf{P}_{model}$ is built for each actuator phase using the signals $\mathbf{X}_h$ recorded by sensors during the experiments with the undamaged structure.

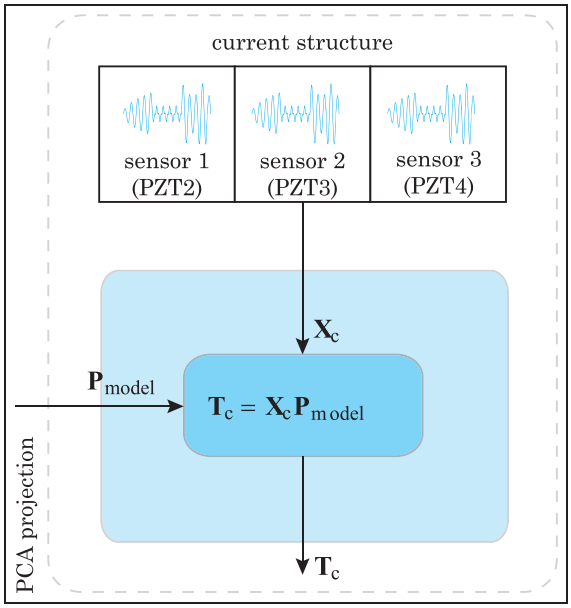

Figure 2. The structure to be diagnosed is subjected to a predefined number of experiments and a data matrix $\mathbf{X}_c$ is constructed. This matrix is projected onto the baseline principal component analysis model $\mathbf{P}_{model}$ to obtain the projections onto the first components $\mathbf{T}_c$.

Multivariate random variables and multivariate random samples

Let us start this section by specifying what we consider a random variable and how a multivariate random variable is built. Assume that for a specific actuator phase (e.g. PZT i as actuator, $i = 1,2,3,4$) and using the signals measured by the sensors in a *fully healthy state* the baseline PCA model (identified as $\mathbf{P}^i_{model}$) is built as in sections "PCA: theoretical background" and "PCA modeling." Assume also that an experiment as detailed in section "Experimental set-up" is further performed. The time responses recorded by the sensors are first discretized and then arranged in a row vector $r^i \in \mathbb{R}^{N\cdot L}$, where N is the number of sensors, L is the number of discretization instants, and i refers to the current actuator phase. The number of sensors and discretization instants must be equal to those that were used when defining $\mathbf{P}^i_{model}$. Besides, the size of each column is $N \cdot L$. Selecting the jth principal component ($j = 1, \ldots, \ell$)

$$\mathbf{P}^i_{model}(:, j) =: v^i_j \in \mathbb{R}^{N\cdot L}$$

the projection of the recorded data onto this principal component is the dot product

$$t^i_j = r^i \cdot v^i_j \in \mathbb{R} \tag{8}$$

as in equation (2).

Since the dynamic behavior of a structure depends on some indeterminacy, its dynamic response can be considered as a stochastic process and the measurements in r^i are also stochastic. On the one hand, t^i_j acquires this stochastic nature and it will be regarded as a random variable to construct the stochastic approach in this article. On the other hand, an s-dimensional random vector can be defined by considering the projections onto several principal components as follows

$$\mathbf{t}^i_{j_1,\ldots,j_s} = \begin{bmatrix} t^i_{j_1} & t^i_{j_2} & \cdots & t^i_{j_s} \end{bmatrix}^T \in \mathbb{R}^s,$$
$$s \in \mathbb{N}, \quad j_1, \ldots, j_s \in \{1, \ldots, \ell\}, \quad j_\alpha \neq j_\beta \quad \text{if } \alpha \neq \beta \tag{9}$$

By reiterating this experiment several times on the undamaged structure and using equations (8) and (9), we have a multivariate random sample of the variable $\mathbf{t}^i_{j_1,\ldots,j_s}$ that can be viewed as a baseline. When structural changes on the structure have to be detected, a new set of experiments should be performed to create the multivariate random sample that will be compared with the multivariate baseline sample. As an example, in Figure 3, two three-dimensional samples are represented; one is the three-dimensional baseline sample (left) and the other is referred to damages 1–3 (right). This illustrating example refers to actuator phase 1 and the first, second, and third principal components. More precisely, Figure 3 (right) depicts the values of the multivariate random variable $\mathbf{t}^1_{1,2,3}$. The diagnosis sample

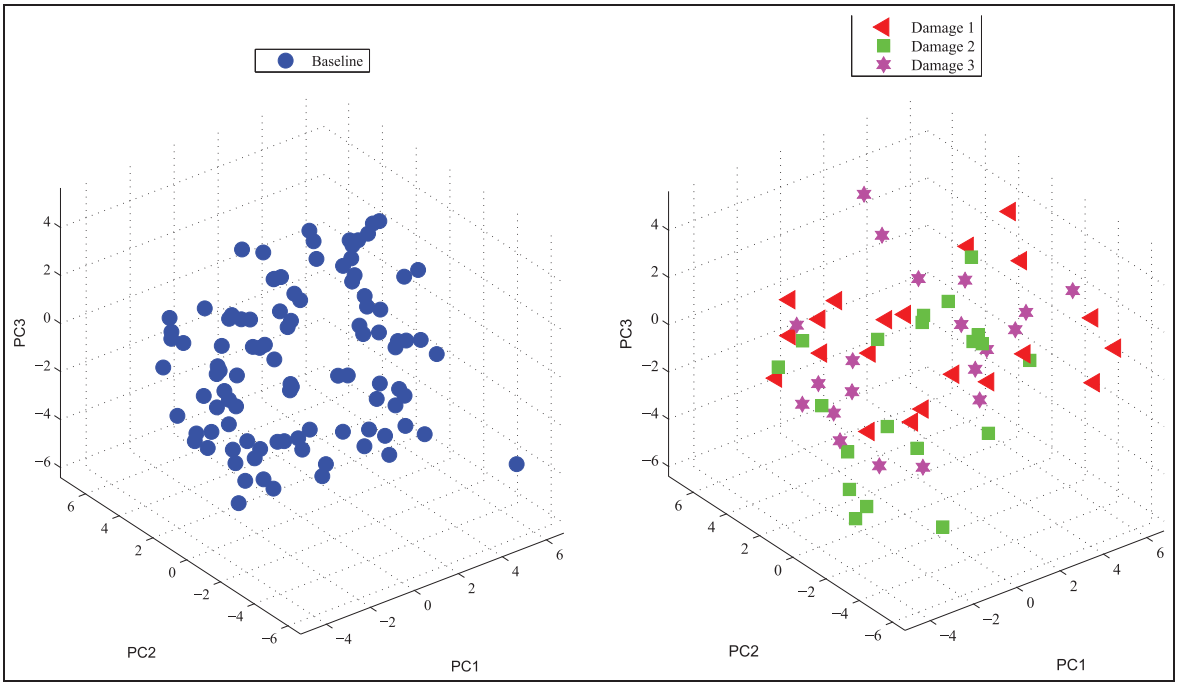

Figure 3. Baseline sample (left) and sample from the structure to be diagnosed (right).

is formed by 20 experiments and the baseline sample is made by 100 experiments.

Detection phase

In this work, the framework of multivariate statistical inference is used with the objective of the classification of structures in healthy or damaged. With this goal, a test for multivariate normality is first performed. A test for the plausibility of a value for a normal population mean vector is then performed.

Testing for multivariate normality. Many statistical tests and graphical approaches are available to check the multivariate normality assumption.[17] But there is no single most powerful test and it is recommended to perform several tests to assess the multivariate normality. Let us consider the three most widely used multivariate normality tests. That is (a) Mardia's, (b) Henze–Zirkler (HZ)'s, and (c) Royston's multivariate normality tests. We include a brief explanation of these methods for the sake of completeness.

Mardia's test

Mardia's test is based on multivariate extensions of skewness ($\hat{\gamma}_{1,s}$) and kurtosis ($\hat{\gamma}_{2,s}$) measures:[17,18]

$$\hat{\gamma}_{1,s} = \frac{1}{\nu^2}\sum_{i=1}^{\nu}\sum_{j=1}^{\nu} m_{ij}^3$$

$$\hat{\gamma}_{2,s} = \frac{1}{\nu}\sum_{i=1}^{\nu} m_{ij}^2$$

where $m_{ij} = (x_i - \bar{x})^T S^{-1}(x_j - \bar{x})$, $i,j = 1, \ldots, \nu$ is the squared Mahalanobis distance, S is the variance-covariance matrix, s is the number of variables, and v is the sample size. The test statistic for skewness, $(\nu/6)\hat{\gamma}_{1,s}$, is approximately χ^2 distributed with s $((s+1)(s+2))/6$ degrees of freedom. Similarly, the test statistic for kurtosis, $\hat{\gamma}_{2,s}$, is approximately normally distributed with mean s $(s+2)$ and variance $8s$ $(s+2)/v$. For multivariate normality, both p values of skewness and kurtosis statistics should be greater than 0.05.

For small samples, the power and the type I error could be violated. Therefore, Mardia introduced a correction term into the skewness test statistic,[19] usually when $v < 20$, in order to control type I errors. The corrected skewness statistic for small samples is $(\nu k/6)\hat{\gamma}_{1,s}$, where $k = ((s+1)(\nu+1)(\nu+3))/(\nu(\nu+1)(s+1) - 6)$. This statistic is also χ^2 distributed with s $((s+1)(s+2))/6$ degrees of freedom.

HZ's test

The HZ's test is based on a non-negative functional distance that measures the distance between two distribution functions.[18,20] If the data are multivariate normal distributed, the test statistic HZ in equation (10) is approximately lognormally distributed. It proceeds to calculate the mean, variance, and smoothness parameter. Then, mean and variance are lognormalized and

the p value is estimated. The test statistic of HZ's multivariate normality test is

$$HZ = \frac{1}{\nu}\sum_{i=1}^{\nu}\sum_{j=1}^{\nu} e^{-\frac{\beta^2}{2}D_{ij}} - 2(1+\beta^2)^{-\frac{s}{2}}\sum_{i=1}^{\nu} e^{-\frac{\beta^2}{2(1+\beta^2)}D_i} + \nu(1+\beta^2)^{-\frac{s}{2}} \quad (10)$$

where s is the number of variables, $\beta = (1/\sqrt{2})((\nu(2s+1)/4))^{1/(s+4)}$, $D_{ij} = (x_i - x_j)^T S^{-1}(x_i - x_j)$, $i,j = 1, \ldots, \nu$ and $D_i = (x_i - \bar{x})^T S^{-1}(x_i - \bar{x}) = m_{ii}$, $i = 1, \ldots, \nu$.

D_i gives the squared Mahalanobis distance of the ith observation to the centroid and D_{ij} gives the Mahalanobis distance between the ith and the jth observations. If data are multivariate normal distributed, the test statistic (HZ) is approximately lognormally distributed with mean μ and variance σ^2 as given below

$$\mu = 1 - \frac{a^{(-s/2)}\left(1 + s\beta^{(2/a)} + s(s+2)\beta^4\right)}{2a^2}$$

$$\sigma^2 = 2\left(1+4\beta^2\right)^{(-s/2)} + \frac{2a^{-s}\left(1+2s\beta^4\right)}{a^2} + \frac{3s(s+2)\beta^8}{4a^4} - 4w_\beta^{(-s/2)}\left(1 + \frac{3s\beta^4}{2w_\beta} + \frac{s(s+2)\beta^8}{2w_\beta^2}\right)$$

where $a = 1 + 2\beta^2$ and $w_\beta = (1+\beta^2)(a+3\beta^2)$. Hence, the lognormalized mean and variance of the HZ statistic can be defined as follows

$$\mu_{\log} = \ln\left(\sqrt{\frac{\mu^4}{\sigma^2+\mu^2}}\right)$$

$$\sigma^2_{\log} = \ln\left(\frac{\sigma^2+\mu^2}{\sigma^2}\right)$$

Using the lognormal distribution parameters, $\mu_{\log}$ and $\sigma^2_{\log}$, we can test the significance of multivariate normality. The Wald test statistic for multivariate normality is given in the following equation

$$z = \frac{\ln(HZ) - \mu_{\log}}{\sqrt{\sigma^2_{\log}}} \quad (11)$$

Royston's test

Royston's test uses the Shapiro–Wilk/Shapiro–Francia statistic to test multivariate normality.[18] If kurtosis of the data is greater than 3, then it uses the Shapiro–Francia test for leptokurtic distributions. Otherwise, it uses the Shapiro–Wilk test for platykurtic distributions. The Shapiro–Wilk test statistic is

$$W = \frac{\left(\sum_{i=1}^{v}(a_i \cdot x_{(i)})^2\right)^2}{\sum_{i=1}^{v}(x_i - \mu)^2}$$

where $x_{(i)}$ is the ith order statistic, that is, the ith-smallest number in the sample, μ is the mean, $a = (\mathbf{m}^T V^{-1})/(\sqrt{\mathbf{m}^T V^{-1} V^{-1}\mathbf{m}})$, and V is the covariance matrix of the order statistics of a sample of s standard normal random variables with expectation vector $\mathbf{m}$. Let W_j be the Shapiro–Wilk/Shapiro–Francia test statistic for the jth variable, $j = 1,\ldots,s$, and Z_j be the values obtained from the normality transformation proposed by[21]

$$\text{if } 4 \leqslant \nu \leqslant 11 \text{ then } x = \nu \text{ and } w_j = -\ln(\gamma - \ln(1 - W_j))$$

$$\text{if } 12 \leqslant \nu \leqslant 2000 \text{ then } x = \ln(\nu) \text{ and } w_j = \ln(1 - W_j)$$

Then transformed values of each random variable can be obtained from the following equation

$$Z_j = \frac{w_j - \mu}{\sigma} \quad (12)$$

where γ, μ, and σ are derived from the polynomial approximations given in equations[21]

$$\text{if } 4 \leqslant \nu \leqslant 11 \quad \gamma = -2.273 + 0.459x$$

$$\mu = 0.544 - 0.39978x + 0.025054x^2 - 0.0006714x^3$$

$$\ln(\sigma) = 1.3822 - 0.77857x + 0.062767x^2 - 0.0020322x^3$$

$$\text{if } 12 \leqslant \nu \leqslant 2000 \quad \mu = -1.5861 - 0.31082x - 0.083751x^2 + 0.0038915x^3$$

$$\ln(\sigma) = -0.4803 - 0.082676x + 0.0030302x^2$$

The Royston's test statistic for multivariate normality is then defined as follows

$$H = \frac{\varepsilon\sum_{j=1}^{s}\psi_j}{s} \sim \chi^2_\varepsilon$$

where ε is the equivalent degrees of freedom (edf) and $\Phi(\cdot)$ is the cumulative distribution function for standard normal distribution such that

$$\varepsilon = \frac{s}{(1 + (s-1)\bar{c})}$$

$$\psi_j = \left(\Phi^{-1}\frac{\Phi(-Z_j)}{2}\right)^2, \quad j = 1, 2, \ldots, s$$

Another extra term $\bar{c}$ has to be calculated in order to continue with the statistical significance of Royston's test statistic. Let R be the correlation matrix and r_{ij} be the correlation between ith and jth variables. Then, the extra term can be found using equation

$$\bar{c} = \sum_{i=1}^{s} \sum_{j \neq i} \frac{c_{ij}}{s(s-1)} \tag{13}$$

where

$$c_{ij} = g(r_{ij}, \nu) \tag{14}$$

with the boundaries of $g(\cdot)$ as $g(0, \nu) = 0$ and $g(1, \nu) = 1$. The function $g(\cdot)$ is defined as follows

$$g(r, \nu) = r^{\lambda}\left(1 - \frac{\mu}{\xi(\nu)}(1-r)^{\mu}\right) \tag{15}$$

The unknown parameters μ, λ, and ξ were estimated from a simulation study conducted by Royston.[21] He found $\mu = 0.715$ and $\lambda = 5$ for sample size $10 \leqslant \nu \leqslant 2000$ and ξ is a cubic function which can be obtained as follows

$$\xi(\nu) = 0.21364 + 0.015124 \ln^2(\nu) - 0.0018034 \ln^3(\nu) \tag{16}$$

Quantile–quantile plot

Apart from the multivariate normality tests, some *visual* representations can also be used to test for multivariate normality. The quantile–quantile (Q-Q) plot is a widely used graphical approach to evaluate the agreement between two probability distributions.[17,18] Each axis refers to the quantiles of probability distributions to be compared, where one of the axes indicates theoretical quantiles (hypothesized quantiles) and the other indicates the observed quantiles. If the observed data fit hypothesized distribution, the points in the Q-Q plot will approximately lie on the bisectrix $y = x$. The sample quantiles for the Q-Q plot are obtained as follows. First, we rank the observations $y_1, y_2, \ldots, y_\nu$ and denote the ordered values by $y_{(1)}, y_{(2)}, \ldots, y_{(\nu)}$; thus $y_{(1)} \leqslant y_{(2)} \leqslant \cdots \leqslant y_{(\nu)}$. Then the point $y_{(i)}$ is the i/ν sample quantile. The fraction i/ν is often changed to $(i-0.5)/\nu$ as a continuity correction. With this convention, $y_{(i)}$ is designated as the $(i-0.5)/\nu$ sample quantile. The population quantiles for the Q-Q plot are similarly defined corresponding to $(i-0.5)/\nu$. If we denote these by $q_1, q_2, \ldots, q_\nu$, then q_i is the value below which a proportion $(i-0.5)/\nu$ of the observations in the population lie; that is, $(i-0.5)/\nu$ is the probability of getting an observation less than or equal to q_i. Formally, q_i can be found for the standard normal random variable Y with distribution $N(0,1)$ by solving

$$\Phi(q_i) = P(Y < q_i) = \frac{i-0.5}{\nu} \tag{17}$$

which would require numerical integration or tables of the cumulative standard normal distribution, $\Phi(x)$. Another benefit of using $(i-0.5)/\nu$ instead of i/ν is that $\nu/\nu = 1$ would make $q_\nu = +\infty$. The population need not have the same mean and variance as the sample since changes in mean and variance merely change the slope and intercept of the plotted lie in the Q-Q plot. Therefore, we use the standard normal distribution, and the q_i values can easily be found from a table of cumulative standard normal probabilities. We then plot the pairs $(q_i, y_{(i)})$ and examine the resulting Q-Q plot for linearity.

Contour plot

In addition to Q-Q plots, creating perspective and contour plots can be also useful.[17,18] The perspective plot is an extension of the univariate probability distribution curve into a three-dimensional probability distribution surface related with bivariate distributions. It also gives information about where data are gathered and how two variables are correlated with each other. It consists of three dimensions where two dimensions refer to the values of the two variables and the third dimension, which is likely in univariate cases, is the value of the multivariate normal probability density function. Another alternative graph, which is called the *contour plot*, involves the projection of the perspective plot into a two-dimensional space and this can be used for checking multivariate normality assumption. Figure 4 shows the contour plot for bivariate normal distribution with mean $(0 \quad 0)^T \in \mathbb{R}^2$ and covariance

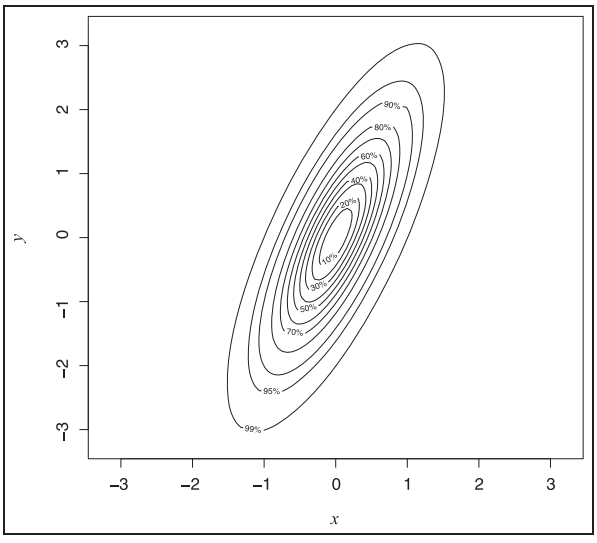

Figure 4. Contour plot for a bivariate normal distribution. The ellipses denote places of equal probability for the distribution and provide confidence regions with different probabilities.

matrix $\begin{pmatrix} 0.250 & 0.375 \\ 0.375 & 1.000 \end{pmatrix} \in \mathscr{M}_{2\times 2}(\mathbb{R})$. For bivariate normally distributed data, we expect to obtain a three-dimensional bell-shaped graph from the perspective plot. Similarly, in the contour plot, we can observe a similar pattern.

Testing a multivariate mean vector. The objective of this article is to determine whether the distribution of the multivariate random samples that are obtained from the structure to be diagnosed (undamaged or not) is connected to the distribution of the baseline. To this end, a test for the plausibility of a value for a normal population mean vector will be performed. Let $s \in \mathbb{N}$ be the number of principal components that will be considered jointly. We will also consider that (a) the baseline projection is a multivariate random sample of a multivariate random variable following a multivariate normal distribution with known population mean vector $\mu_{\mathrm{h}} \in \mathbb{R}^s$ and known variance-covariance matrix $\Sigma \in \mathscr{M}_{s\times s}(\mathbb{R})$ and (b) the multivariate random sample of the structure to be diagnosed also follows a multivariate normal distribution with unknown multivariate mean vector $\mu_{\mathrm{c}} \in \mathbb{R}^s$ and known variance-covariance matrix $\Sigma \in \mathscr{M}_{s\times s}(\mathbb{R})$.

As said previously, the problem that we will consider is to determine whether a given s-dimensional vector μ_{c} is a plausible value for the mean of a multivariate normal distribution $N_s(\mu_{\mathrm{h}}, \Sigma)$. This statement leads immediately to a test of the hypothesis

$$H_0 : \mu_{\mathrm{c}} = \mu_{\mathrm{h}} \textit{versus}$$
$$H_1 : \mu_{\mathrm{c}} \neq \mu_{\mathrm{h}}$$

that is, the null hypothesis is "the multivariate random sample of the structure to be diagnosed is distributed as the baseline projection" and the alternative hypothesis is "the multivariate random sample of the structure to be diagnosed is not distributed as the baseline projection." In other words, if the result of the test is that the null hypothesis is not rejected, the current structure is categorized as healthy. Otherwise, if the null hypothesis is rejected in favor of the alternative, this would indicate the presence of some structural changes in the structure.

The test is based on the statistic T^2—also called Hotelling's T^2—and it is summarized below. When a multivariate random sample of size $\nu \in \mathbb{N}$ is taken from a multivariate normal distribution $N_s(\mu_{\mathrm{h}}, \sum)$, the random variable

$$T^2 = \nu(\bar{\mathbf{X}} - \mu_{\mathrm{h}})^T \mathbf{S}^{-1}(\bar{\mathbf{X}} - \mu_{\mathrm{h}})$$

is distributed as

$$T^2 \rightarrowtail \frac{(\nu-1)s}{\nu - s} F_{s,\,\nu-s}$$

where $F_{s,\,\nu-s}$ denotes a random variable with an F-distribution with s and $\nu - s$ degrees of freedom, $\bar{\mathbf{X}}$ is the sample vector mean as a multivariate random variable; and $(1/n)(\mathbf{S} \in \mathscr{M}_{s\times s}(\mathbb{R}))$ is the estimated covariance matrix of $\bar{\mathbf{X}}$.

At the α level of significance, we reject H_0 in favor of H_1 if the observed

$$t^2_{obs} = \nu(\bar{\mathbf{X}} - \mu_{\mathrm{h}})^T \mathbf{S}^{-1}(\bar{\mathbf{X}} - \mu_{\mathrm{h}})$$

is greater than $(((\nu-1)s)/(\nu - s))F_{s,\,\nu-s}(\alpha)$, where $F_{s,\,\nu-s}(\alpha)$ is the upper (100α)th percentile of the $F_{s,\,\nu-s}$ distribution. In other words, the quantity t^2_{obs} is the damage indicator and the test is summarized as follows

$$t^2_{obs} \leqslant \frac{(\nu-1)s}{\nu - s} F_{s,\,\nu-s}(\alpha) \;\Rightarrow\; \text{Fail to reject } H_0 \qquad (18)$$

$$t^2_{obs} > \frac{(\nu-1)s}{\nu - s} F_{s,\,\nu-s}(\alpha) \;\Rightarrow\; \text{Reject } H_0 \qquad (19)$$

where $F_{s,\,\nu-s}(\alpha)$ is such that

$$\mathbb{P}(F_{s,\,\nu-s} > F_{s,\,\nu-s}(\alpha)) = \alpha$$

where $\mathbb{P}$ is a probability measure. More precisely, we fail to reject the null hypothesis if $t^2_{obs} \leqslant (((\nu-1)s)/(\nu - s))F_{s,\,\nu-s}(\alpha)$, thus indicating that no structural changes in the structure have been found. Otherwise, the null hypothesis is rejected in favor of the alternative hypothesis if $t^2_{obs} > (((\nu-1)s)/(\nu - s))$ $F_{s,\,\nu-s}(\alpha)$, thus indicating the existence of some structural changes in the structure.

Experimental results

In this work, a particular experimental set-up based on the analysis of vibrational changes is used as an exemplifying configuration in order to justify, validate, and test the methodology. The proposed methodology can also be applied to a more general structure.

Experimental set-up

Some experiments were performed in order to test the methods presented on this article. In these experiments, four PZT discs were attached to the surface of a thin aluminum plate, with dimensions $250 \times 250 \times 1$ mm^3. Those PZTs formed a square with 144 mm per side. The plate was suspended by two elastic ropes, being isolated from environmental influences. Figures 5 (left) and 6 show the plate hanging on the elastic ropes.

As a response to an electrical excitation, a PZT produces a mechanical vibration, propagating, in this case,

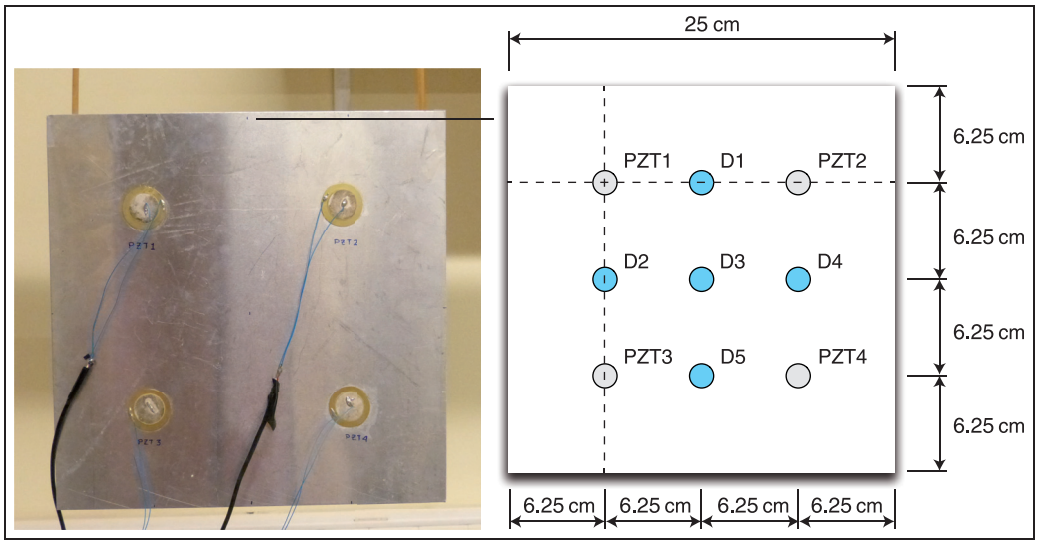

Figure 5. Aluminum plate (left). Dimensions and piezoelectric transducers location (right).

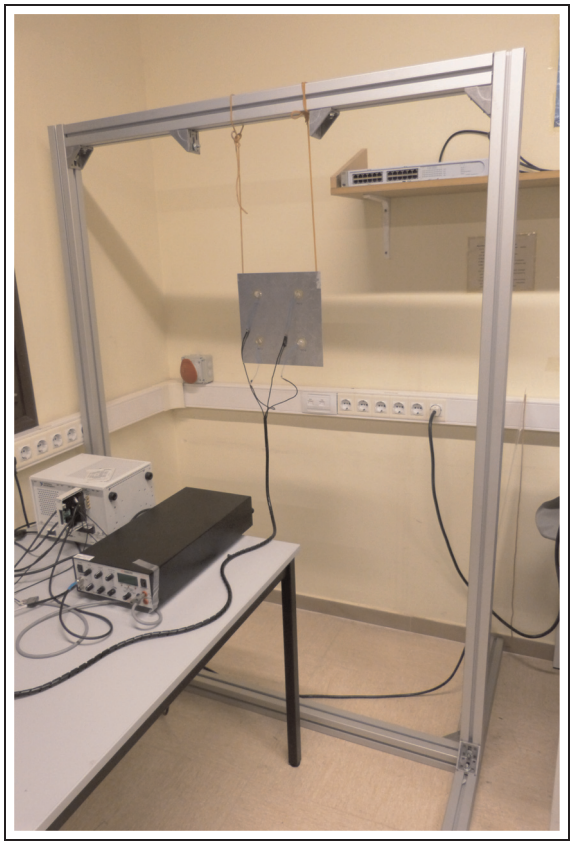

Figure 6. The plate is suspended by two elastic ropes in a metallic frame.

across the plate (forming Lamb waves since a thin plate has been used). PZTs are also able to generate an electrical signal as a response to a mechanical vibration. In every excitation phase of an experiment, one PZT was used as actuator and the other three PZTs were used as sensors, recording the dynamical response of the plate.

In all, 500 experiments were performed over the healthy structure, and another 500 experiments were performed over the damaged structure with five damage types (100 experiments per damage type). Figure 5 (right) shows the position of damages 1–5 (D1–D5). As excitation, a 50 kHz sinusoidal signal modulated by a hamming window was used. Figure 7 shows the excitation signal and an example of the signal collected by PZT1.

Multivariate normality

As said in section "Experimental set-up," the experiments are performed in four independent phases: (a) PZT1 is configured as actuator and the rest of PZTs as sensors, (b) PZT2 as actuator, (c) PZT3 as actuator, and (d) PZT4 as actuator. In order to analyze the influence of each set of projections to the PCA model (score), the results of scores 1–3 (jointly), scores 1–5 (jointly), and scores 1–10 (jointly) have been considered. In this way, a total of 12 scenarios were examined.

The multivariate normality tests described in section "Testing for multivariate normality" were performed for all the data. We summarize in Table 1 the results of the multivariate normality test when considering the first three principal components (PC1–PC3) for all the actuator phases.

Some examples of Q-Q plots for the data we consider in this article are shown in Figure 8. It can be observed that the points are distributed closely following the bisectrix, thus indicating the multivariate normality of the data as stated in Table 1.

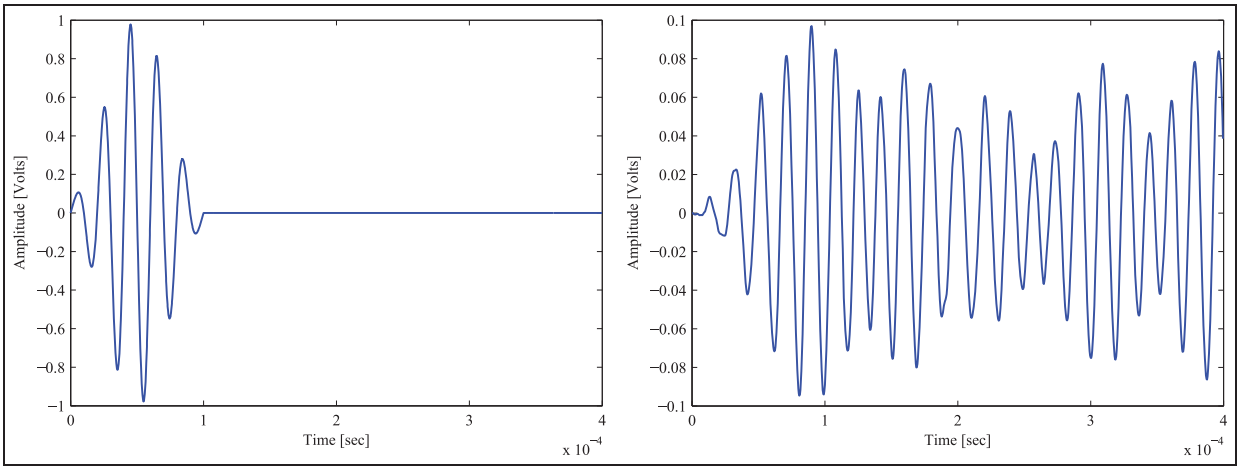

Figure 7. Excitation signal (left), and dynamic response recorded by piezoelectric transducer 1 (right).

Table 1. Results of the multivariate normality tests when considering the first three principal components (PC1–PC3) in the four actuator phases. "−" means that all the tests rejected multivariate normality, "+" means that at least one test indicated multivariate normality while the subindex shows the tests that indicated normality: 1 (Mardia's test), 2 (Henze–Zirkler's test), or 3 (Royston's test).

	PZT1 act.	PZT2 act.	PZT3 act.	PZT4 act.
Undamaged (baseline)	−	$+_{2}$	$+_{2}$	−
Undamaged (first set to test)	−	$+_{1,2,3}$	$+_{2}$	−
Undamaged (second set to test)	$+_{1}$	$+_{1,2}$	−	−
Undamaged (third set to test)	−	$+_{2}$	−	$+_{2,3}$
Undamaged (fourth set to test)	−	−	−	$+_{1,2,3}$
Undamaged (fifth set to test)	−	−	$+_{1}$	$+_{1,3}$
D1	$+_{1,2,3}$	$+_{2}$	$+_{1,2}$	$+_{3}$
D2	$+_{1,2,3}$	$+_{1,2,3}$	$+_{1}$	$+_{1,3}$
D3	$+_{1,2,3}$	$+_{2}$	$+_{1,2}$	−
D4	$+_{2}$	$+_{2,3}$	−	$+_{3}$
D5	$+_{1,2,3}$	−	$+_{1}$	−

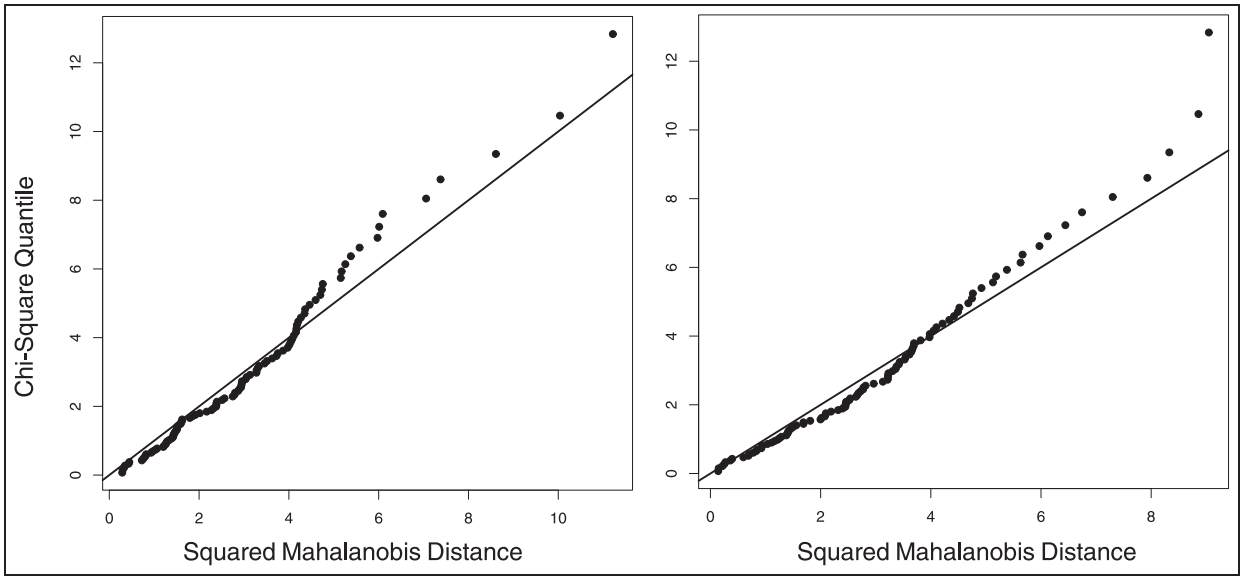

Figure 8. Q-Q plots corresponding to: (i) fourth set of undamaged data to test, using the first three principal components (PC1–PC3) in the actuator phase 4 (left) and (ii) damage 2 data, using the first three principal components (PC1–PC3) in the actuator phase 1 (right). The points of these Q-Q plots are close to the line $y = x$ thus indicating the multivariate normality of the data.

Table 2. Results of univariate normality tests when considering the first five principal components separately in the four actuator phases. "−" means lack of normality while "+" means normality.

	PZT1 act.	PZT2 act.	PZT3 act.	PZT4 act.
Undamaged (baseline)	− + − + +	− + + + +	− + + + +	− + + − +
Undamaged (first set to test)	− − − + −	− + + + −	− + + − +	+ − + + +
Undamaged (second set to test)	− + + + +	− + + + +	− + + + +	− + + − +
Undamaged (third set to test)	− − + + +	− + + + +	− + + + +	− + + + +
Undamaged (fourth set to test)	− + − + +	− + + + +	− + + + −	− + + − +
Undamaged (fifth set to test)	− + − + +	− + + + +	− + + + +	+ + + + +
D1	− + + + +	− + + − +	− + + − −	+ + + + +
D2	− + + + +	− + + + +	− + + + −	+ + + + +
D3	+ + + + +	− + + + +	− + + + +	+ + + + +
D4	− + + + +	+ + + − +	− + + + +	− + + + +
D5	+ + + + −	− + + + +	− + − + −	− + + + +

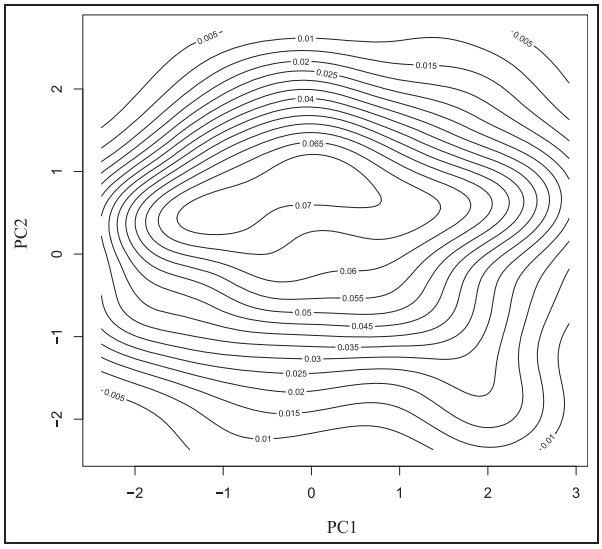

Figure 9. Contour plot for undamaged case (fourth set to test), piezoelectric transducer 4 act., principal components (PC1 and PC2). The contour lines are similar to ellipses of normal bivariate distribution from Figure 4 that means that the distribution in this case is normal.

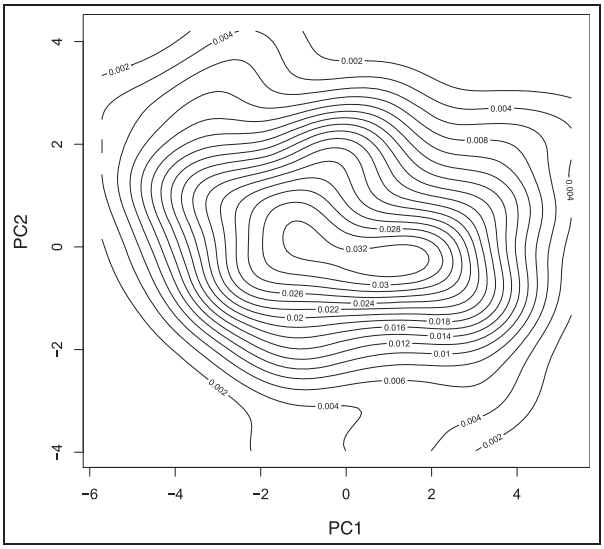

Figure 10. Contour plot for case D3, piezoelectric transducer 1 act., and principal components (PC1 and PC2). The contour lines are similar to ellipses of normal bivariate distribution from Figure 4 that means that the distribution in this case is normal.

Moreover, some other examples of contour plots for the data we consider in this article are given in Figures 9 and 10. These plots are similar to the contour plot of the bivariate normal distribution in Figure 4.

Finally, the univariate normality for each principal component and for each actuator phase is also tested. The results are presented in Table 2. As it can be observed, the univariate data are normally distributed in most of the cases. However, this does not imply multivariate normality.

Type I and Type II errors

For each scenario, a total of 50 samples of 20 experiments each one (25 for the undamaged structure and 5 for the damaged structure with respect to each of the 5 different types of damages) plus the baseline are used to test for the plausibility of a value for a normal population mean vector, with a level of significance $\alpha = 0.60$. Each set of 50 testing samples are categorized as follows: (i) number of samples from the healthy structure (undamaged sample) which were classified by the hypothesis test as "healthy" (fail to reject H_0); (ii) undamaged sample classified by the test as "damaged" (reject H_0); (iii) samples from the damaged structure (damaged sample) classified as "healthy"; and (iv) damaged sample classified as "damaged." The results for the 12 different scenarios presented in Table 4 are organized according to the scheme in Table 3. It can be stressed from each scenario in Table 4 that the sum of

Table 3. Scheme for the presentation of the results in Table 4.

	Undamaged sample (H_0)	Damaged sample (H_1)
Fail to reject H_0	Correct decision	Type II error (missing fault)
Reject H_0	Type I error (false alarm)	Correct decision

Table 4. Categorization of the samples with respect to presence or absence of damage and the result of the test, for each of the four phases and considering the first score, the second score, scores 1–3 (jointly), scores 1–5 (jointly), and scores 1–10 (jointly).

	PZT1 act.		PZT2 act.		PZT3 act.		PZT4 act.	
	H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1
Score 1								
Fail to reject H_0	22	13	21	7	18	13	22	12
Reject H_0	3	12	4	18	7	12	3	13
Score 2								
Fail to reject H_0	21	2	24	18	18	5	22	14
Reject H_0	4	23	1	7	7	20	3	11
Scores 1–3								
Fail to reject H_0	24	0	24	13	25	9	24	4
Reject H_0	1	25	1	12	0	16	1	21
Scores 1–5								
Fail to reject H_0	21	0	23	0	21	0	20	0
Reject H_0	4	25	2	25	4	25	5	25
Scores 1–10								
Fail to reject H_0	25	0	25	0	25	0	25	0
Reject H_0	0	25	0	25	0	25	0	25

the columns is constant: 25 samples in the first column (undamaged structure) and 25 more samples in the second column (damaged structure).

In this table, it is worth noting that two kinds of misclassification are presented which are denoted as follows:

1. Type I error (or *false positive*), when the structure is healthy but the null hypothesis is rejected and therefore classified as damaged.
2. Type II error (or *false negative*), when the structure is damaged but the null hypothesis is not rejected and therefore classified as healthy. The probability of committing a type II error is called β.

It can be observed from Table 4 that Type I errors (false alarms) appear only when we consider scores 1–3 (jointly) and scores 1–5 (jointly), while in the last case (scores 1–10), all the decisions are correct.

Sensitivity and specificity

Two more statistical measures can be selected here to study the performance of the test: *the sensitivity* and *the specificity*. The sensitivity, also called as the power of the test, is defined, in the context of this work, as the proportion of samples from the damaged structure which are correctly identified as such. Thus, the sensitivity can be computed as $1 - \beta$. The specificity of the test is defined, also in this context, as the proportion of samples from the undamaged structure that are correctly identified and can be expressed as $1 - \alpha$.

The sensitivity and the specificity of the test with respect to the 50 samples in each scenario have been included in Table 5. For each scenario in this table, the results are organized as shown in Table 6.

It is worth noting that type I errors are frequently considered to be more serious than type II errors. However, in this application, a type II error is related to a *missing fault*, whereas a type I error is related to a *false alarm*. In consequence, type II errors should be minimized. Therefore, a small level of significance of 1%, 5%, or even 10% would lead to a reduced number of *false alarms* but to a higher rate of *missing faults*. That is the reason of the choice of a level of significance of 60% in the hypothesis test.

The results show that the sensitivity of the test $1 - \beta$ is close to 100%, as desired, with an average value of 78%. The sensitivity with respect to scores 1–5 and scores 1–10 is increased, in mean, to a 100%. The average value of the specificity is 90%.

Table 5. Sensitivity and specificity of the test for each scenario.

	PZT1 act.		PZT2 act.		PZT3 act.		PZT4 act.	
	H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1
Score 1								
Fail to reject H_0	0.88	0.52	0.84	0.28	0.72	0.52	0.88	0.48
Reject H_0	0.12	0.48	0.16	0.72	0.28	0.48	0.12	0.52
Score 2								
Fail to reject H_0	0.84	0.08	0.96	0.72	0.72	0.20	0.88	0.56
Reject H_0	0.16	0.92	0.04	0.28	0.28	0.80	0.12	0.44
Scores 1–3								
Fail to reject H_0	0.96	0.00	0.96	0.52	1.00	0.36	0.96	0.16
Reject H_0	0.04	1.00	0.04	0.48	0.00	0.64	0.04	0.84
Scores 1–5								
Fail to reject H_0	0.84	0.00	0.92	0.00	0.84	0.00	0.80	0.00
Reject H_0	0.16	1.00	0.08	1.00	0.16	1.00	0.20	1.00
Scores 1–10								
Fail to reject H_0	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
Reject H_0	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00

Table 6. Relationship between type I and type II errors.

	Undamaged sample (H_0)	Damaged sample (H_1)
Fail to reject H_0	Specificity ($1 - \alpha$)	False-negative rate (β)
Reject H_0	False-positive rate (α)	Sensitivity ($1 - \beta$)

Table 7. True rate of false positives and false negatives.

	PZT1 act.		PZT2 act.		PZT3 act.		PZT4 act.	
	H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1
Score 1								
Fail to reject H_0	0.63	0.37	0.75	0.25	0.58	0.42	0.65	0.35
Reject H_0	0.20	0.80	0.18	0.82	0.37	0.63	0.19	0.81
Score 2								
Fail to reject H_0	0.91	0.09	0.57	0.43	0.78	0.22	0.61	0.39
Reject H_0	0.15	0.85	0.13	0.88	0.26	0.74	0.21	0.79
Scores 1–3								
Fail to reject H_0	1.00	0.00	0.65	0.35	0.74	0.26	0.86	0.14
Reject H_0	0.04	0.96	0.08	0.92	0.00	1.00	0.05	0.95
Scores 1–5								
Fail to reject H_0	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
Reject H_0	0.14	0.86	0.07	0.93	0.14	0.86	0.17	0.83
Scores 1–10								
Fail to reject H_0	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
Reject H_0	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00

Reliability of the results

The results in Table 7 are computed using the scheme in Table 8. This table is based on the Bayes' theorem,[22] where $P(H_1|accept\ H_0)$ is the proportion of samples from the damaged structure that has been incorrectly classified as healthy (*true rate of false negatives*) and $P(H_0|accept\ H_1)$ is the proportion of samples from the undamaged structure that has been incorrectly classified as damaged (*true rate of false positives*).

The receiver operating characteristic (ROC) curves

An additional study has been developed based on the ROC curves to determine the overall accuracy of the proposed method. These curves represent the trade-off

Table 8. Relationship between proportion of false negative and false positives.

	Undamaged sample (H_0)	Damaged sample (H_1)
Fail to reject H_0	$P(H_0 \mid accept\ H_0)$	True rate of false negatives $P(H_1 \mid accept\ H_0)$
Reject H_0	True rate of false positives $P(H_0 \mid accept\ H_1)$	$P(H_1 \mid accept\ H_1)$

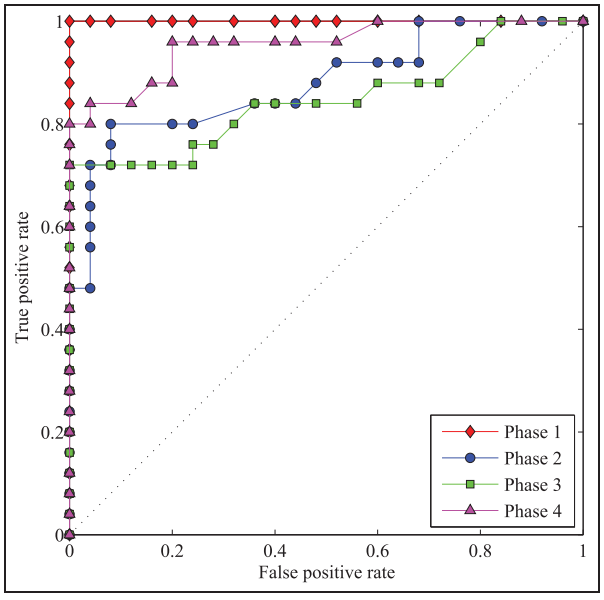

Figure 11. The receiver operating characteristic curves for the scores 1–3 (jointly) in the four actuator phases.

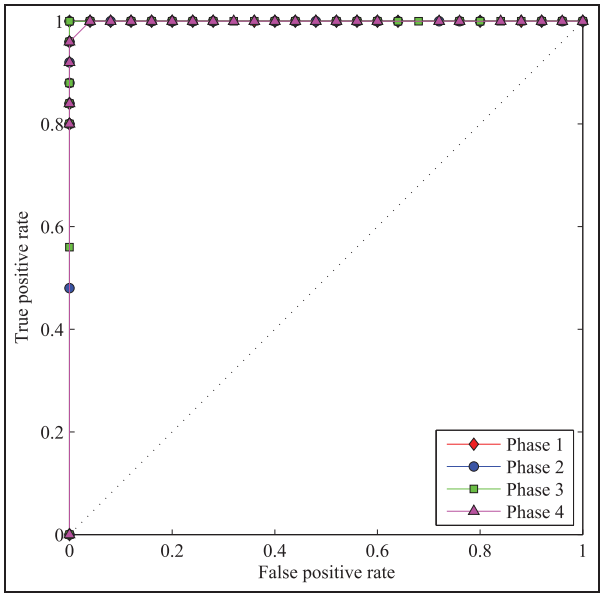

Figure 12. The receiver operating characteristic curves for the scores 1–5 (jointly) in the four actuator phases.

between the *false-positive rate* and the *sensitivity* in Table 6 for different values of the level of significance that is used in the statistical hypothesis testing. Note that the false-positive rate is defined as the complementary of the specificity, and therefore these curves can also be used to visualize the close relationship between specificity and sensitivity. It can also be remarked that the sensitivity is also called true positive rate or probability of detection.[23] More precisely, for each scenario and for a given level of significance, the pair of numbers

$$(\text{false} - \text{positive rate},\ \text{sensitivity}) \in [0,1]\times[0,1] \subset \mathbb{R}^2 \tag{20}$$

is plotted. We have considered 99 levels of significance within the range [001, 0.99] and with a difference of 0.01. Therefore, for each scenario 99 connected points are depicted, as can be seen in Figures 11–13 when we consider scores 1–3 (jointly), scores 1–5 (jointly), and scores 1–10 (respectively).

The placement of these points can be interpreted as follows. Since we are interested in minimizing the number of false positives while we maximize the number of true positives, these points must be placed in the upper-left corner as much as possible. However, this is not always possible because there is also a relationship between the level of significance and the false-positive rate. Therefore, a method can be considered acceptable if those points lie within the upper-left half-plane.

As said earlier, the ROC curves for the 12 possible scenarios are depicted in Figures 11–13. The best performance is achieved for the case of scores 1–3 in phase 1 (Figure 11) because all of the points are placed in the upper-left corner. In phases 2–4, the points lie in the upper-left half-plain but not in the corner, which represents a very good behavior of the proposed method. When we consider the case of scores 1–5 (jointly) in Figure 12 and the case of scores 1–10 (jointly) in Figure 13, it can be observed that the area under the ROC curves is close to 1 in all of the actuator phases thus representing an excellent test.

Finally, we can say that the ROC curves provide a statistical assessment of the efficacy of a method and can be used to visualize and compare the performance of multiple scenarios.

Analysis and discussion

Multivariate tests allow to get better results in damage detection than univariate tests. This is perfectly illustrated in Figure 14 where correct or wrong detections are represented as a function of the level of significance α used in the test. We can clearly characterize four different regions:

- $0 < \alpha \leqslant 0.13$. In this region, *all* the five univariate tests and the multivariate statistical inference *pass* (correct decision).
- $0.13 < \alpha \leqslant 0.62$. In this region, *some* of the five univariate tests *fail* (wrong decision) while the multivariate statistical inference *pass* (correct decision).
- $0.62 < \alpha \leqslant 0.71$. In this region, *all* the five univariate tests fail (wrong decision) while the multivariate statistical inference pass (correct decision).
- $0.71 < \alpha < 1$. In this region, *all* the five univariate tests and the multivariate statistical inference *fail* (wrong decision).

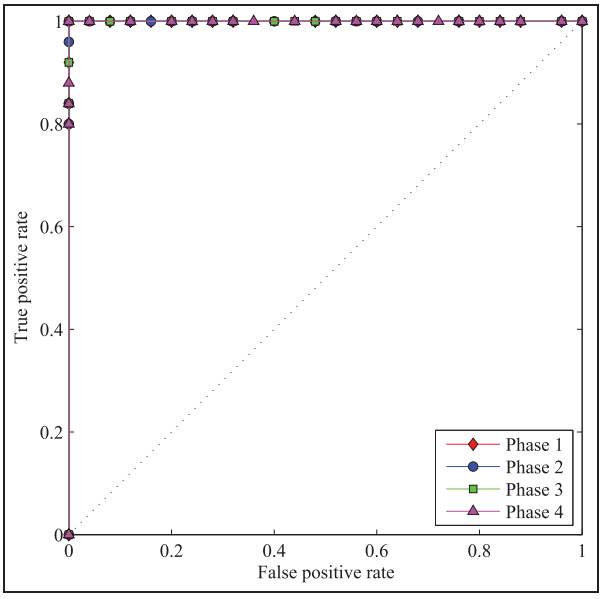

Figure 13. The receiver operating characteristic curves for the scores 1–10 (jointly) in the four actuator phases.

It is worth noting that in the region $0.62 < \alpha \leqslant 0.71$, that is, when the level of significance lies within the range (0.62, 0.71], the multivariate statistical inference using scores 1–5 (jointly) is able to offer a correct decision although all of the univariate tests make a wrong decision.

The scenarios with the best results are those that considers scores 1–10, because the false-negative rate is 0% and the false-positive rate is 0% for all the actuator phases. The results for scores 1–5 (jointly) are quite good, because the false-negative rate is 0% for all actuators and the false-positive rate is 7%–17%.

Concluding remarks

In this article, a new methodology to detect structural changes has been introduced. The three main features that characterize the proposed methodology are (a) the nature of the data used in the test since we are using vectors of scores instead of the measured response of the structure or the coefficients of an AutoRegressive model, (b) the number of data used since the test is based on two random samples instead of two

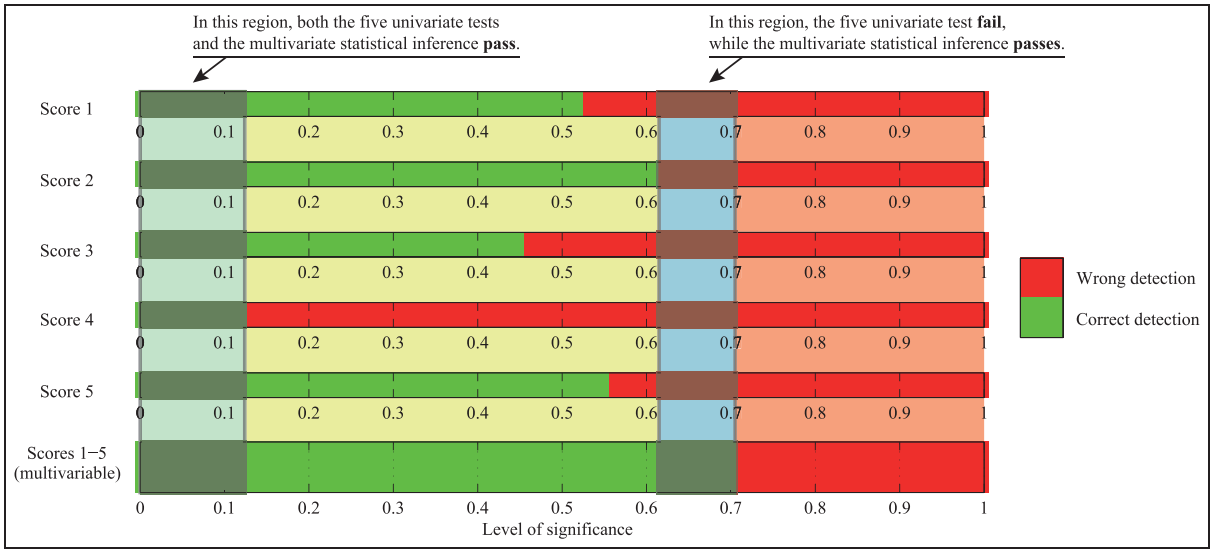

Figure 14. Multivariate tests allow to get better results in damage detection that univariate tests. A correct or wrong detection is represented as a function of the level of significance where four regions can be identified.

characteristic quantities, and (c) the samples come from a multidimensional variable and therefore a test for the plausibility of a value for a normal population mean vector is performed. In this work, the framework of multivariate statistical inference is used with the objective of the classification of structures in healthy or damaged. With this goal, a test for multivariate normality was first performed.

The proposed methodology has been applied and validated on a small aluminum plate. The results indicate that the presented methodology is able to accurately detect damages, that is, for each actuation phase, a unique and reliable damage detection indicator is obtained.

It is worth noting that there is an entire range of significance levels where the multivariate statistical inference using, for instance, scores 1–5 (jointly) is able to offer a correct decision although all of the univariate tests make a wrong decision.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by CICYT (Spanish Ministry of Economy and Competitiveness) through grants DPI2011-28033-C03-01 and DPI2014-58427-C2-1-R.

References

1. Farrar CR, Doebling SW and Nix DA. Vibration-based structural damage identification. *Philos T R Soc A* 2001; 359(1778): 131–149.
2. Farrar CR and Worden K. An introduction to structural health monitoring. *Philos T R Soc A* 2007; 365(1851): 303–315.
3. Farrar CR and Lieven NAJ. Damage prognosis: the future of structural health monitoring. *Philos T R Soc A* 2007; 365(1851): 623–632.
4. Farrar CR and Worden K. *Structural health monitoring: a machine learning perspective*. Hoboken, NJ: Wiley, 2012.
5. Anaya M, Tibaduiza DA, Torres M, et al. Data-driven methodology to detect and classify structural changes under temperature variations. *Smart Mater Struct* 2014; 23(4): 045006.
6. Gharibnezhad F, Mujica LE and Rodellar J. Applying robust variant of principal component analysis as a damage detector in the presence of outliers. *Mech Syst Signal Pr* 2015; 50–51: 467–479.
7. Mujica LE, Ruiz M, Güemes A, et al. Contribution plots on PCA based indices for damage identification on structures. In: *4th ECCOMAS thematic conference on smart structures and materials*, Porto, Portugal, 13–15 July 2009.
8. Mujica LE, Rodellar J, Fernández A, et al. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. *Struct Health Monit* 2011; 10(5): 539–553.
9. Sierra-Pérez J, Güemes A, Mujica LE, et al. Damage detection in composite materials structures under variable loads conditions by using fiber Bragg gratings and principal component analysis, involving new unfolding and scaling methods. *J Intel Mat Syst Str*, http://jim.sagepub.com/content/early/2014/07/09/1045389X14541493
10. Tibaduiza DA, Mujica LE and Rodellar J. Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Struct Control Hlth* 2013; 20: 1303–1316.
11. Tibaduiza DA, Mujica LE, Rodellar J, et al. Structural damage detection using principal component analysis and damage indices. *J Intel Mat Syst Str*. Epub ahead of print 8 January 2015. DOI: 10.1177/1045389X14566520.
12. Torres-Arredondo MA, Tibaduiza DA, McGugan M, et al. Multivariate data-driven modelling and pattern recognition for damage detection and identification for acoustic emission and acousto-ultrasonics. *Smart Mater Struct* 2013; 22: 105023.
13. Torres-Arredondo MA, Tibaduiza DA, Mujica LE, et al. Data-driven multivariate algorithms for damage detection and identification: evaluation and comparison. *Struct Health Monit* 2014; 13(1): 19–32.
14. Mujica LE, Ruiz M, Pozo F, et al. A structural damage detection indicator based on principal component analysis and statistical hypothesis testing. *Smart Mater Struct* 2014; 23(2): 025014.
15. Zugasti E, Gómez González A, Anduaga J, et al. Null-Space and AutoRegressive damage detection: a comparative study. *Smart Mater Struct* 2012; 21: 1–9.
16. Fassois SP and Sakellariou JS. Time-series methods for fault detection and identification in vibrating structures. *Philos T R Soc A* 2007; 365: 411–448.
17. Rencher AC and Christensen WF. *Methods of multivariate analysis*. Hoboken, NJ: Wiley, 2012.
18. Korkmaz S, Goksuluk D and Zararsiz G. *MVN: an R package for assessing multivariate normality* (MVN version 3.8), 2015, https://journal.r-project.org/archive/2014-2/korkmaz-goksuluk-zararsiz.pdf
19. Mardia KV. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya Ser B* 1974; 36(2): 115–128.
20. Henze N and Zirkler B. A class of invariant consistent tests for multivariate normality. *Commun Stat Theory* 1990; 19(10): 3595–3617.
21. Royston P. Approximating the Shapiro-Wilk W test for non-normality. *Stat Comput* 1992; 2(3): 117–119.
22. DeGroot MH and Schervish MJ. *Probability and statistics*. London: Pearson, 2012.
23. Yinghui L and Michaels JE. Feature extraction and sensor fusion for ultrasonic structural health monitoring under changing environmental conditions. *IEEE Sens J* 2009; 9(11): 1462–1471.