Comparison of several methods for damage localization using indices and contributions based on PCA

D.A. Tibaduiza, L.E. Mujica, J. Rodellar
Control, Dynamics and Applications Group (CoDAlab).Department of Applied Mathematics III. Escola Universitària d’Enginyeria Tècnica Industrial de Barcelona (EUETIB). Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
E-mail: {diego.tibaduiza, luis.eduardo.mujica, jose.rodellar}@upc.edu

Abstract. In previous works by the authors, it was shown the advantages of applying an active piezoelectric system in combination with Principal Component Analysis and Neural Networks as methodology for damage detection in structures.

An active piezoelectric system considers the advantage of using piezoelectric transducers (PZT’s) as actuator as well as sensor. In each phase of the diagnosis procedure, one PZT is used as actuator (a known electrical signal is applied) and the others are used as sensors (collecting the wave propagated through the structure at different points). An initial baseline model for undamaged structure is built applying Principal Component Analysis (PCA) to the data collected by several experiments. Current structure (damaged or not) is subjected to the same experiments, and the collected data are projected in the PCA model. In this paper, two indices are used to detect damages; these indices are calculated from the information obtained from the projection of the experiments in the PCA model (baseline). Besides the localization is performed using five different methods. These methods are based on the contribution of each sensor to each index, in this way, according to these contributions, the damage can be localized. The combination of all indices and all contributions (a total of 2 x 5) are analyzed and compared. To validate the approach, the methods are applied to an aluminum plate which is instrumented with several PZT’s.

1. Introduction
Structural Health Monitoring (SHM) is the integration of elements of actuation and sensing with different mathematics and computational techniques in order to know the health of a structure using non-destructive techniques. All the data obtained from the structure are analyzed in order to detect abnormal characteristics to define the health of the structure. The obtained information can be used to define whether the structure can operate and, in which conditions. SHM has been applied in different areas that include civil engineering, many works has been reported for more than three decades [1] with excellent results, for instance, tests is bridges[2] [3], buildings [4] and other structures [5][6],[7].

In other areas like aeronautical and aerospace engineering [8], SHM is very important to evaluate constantly the health of the structures to guarantee appropriate operation. Probably, the use of SHM in these two areas is most important than other, for instance some damages in civil
structures are really dangerous when they start to have a considerable size. In aeronautical and aerospace structures, damages that are imperceptible when subjected to extreme changes in their working conditions can cause catastrophes.

In SHM exist various techniques for monitoring of structures, some of them include the use of statistics techniques such as Principal Component Analysis (PCA) in time domain or frequency domain, which usage is depend on for instance of the experimental configuration. The methodology reported in this paper uses vibrational responses collected from PZT’s using an active piezoelectric system [9] and PCA as pattern recognition technique. Initially, a base-line model of the structure without damage is built, when the system is ready (operating mode), the vibrational response of the current structure under test is compared with the model and using two indices PCA based ($T^2$ and $Q$) which give us information about damages. The methodology include the use of five methods for damage localization, these methods define the contributions of each damage to the different indices.

2. Principal Component Analysis (PCA)
Principal Component Analysis is a technique of multivariable and megavariable analysis [10] which may provide arguments for how to reduce a complex data set to a lower dimension and reveal some hidden and simplified structure/patterns that often underlie it. The goal of PCA is to discern which dynamics are more important in the system, which are redundant and which are just noise [11]. This goal is essentially achieved by determining a new space (coordinates) to re-express the original data filtering that noise and redundancies based on the variance-covariance structure of the original data. PCA can be also considered as a simple, non-parametric method for data compression and information extraction, which finds combinations of variables or factors that describe major trends in a confusing data set [12]. Among their objectives it can be mentioned: to generate new variables that could express the information contained in the original set of data, to reduce the dimensionality of the problem that is studied, to eliminate some original variables if its information is not relevant. In order to develop a PCA model it is necessary to arrange the collected data in a matrix $X$, this $m \times n$ matrix contains information from $n$ sensors and $m$ experimental trials [13]. Since physical variables and sensors have different magnitudes and scales, each data-point is scaled using the mean of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor. Once the variables are normalized the covariance matrix $C_x$ is calculated as follows:

$$C_x = \frac{1}{m-1}X^TX$$

where $C_x$ is a square symmetric $m \times m$ matrix that measures the degree of linear relationship within the data set between all possible pairs of variables (sensors). The subspaces in PCA are defined by the eigenvectors and eigenvalues of the covariance matrix as follow:

$$C_xP = \Lambda P,$$

where the eigenvectors of $C_x$ are the columns of $P$, and the eigenvalues are the diagonal terms of $\Lambda$ (the off-diagonal terms are zero). Columns of matrix $P$ are sorted according to the eigenvalues by descending order and they are called the Principal Components of the data set. The eigenvectors with highest eigenvalue represents the most important pattern in the data with the largest quantity of information. Choosing only a reduced number $r$ of principal components, those corresponding to the first eigenvalues, the reduced transformation matrix could be imagined as a model for the structure. Geometrically, the transformed data matrix $T$
(score matrix) is the projection of the original data over the direction of the principal components \( P \).

\[
T = XP
\]  

(3)

In the full dimension case, this projection is invertible (since \( PP^T = I \)) and the original data can be recovered as \( X = TP^T \). Now, with the given \( T \), it is not possible to fully recover \( X \), but \( T \) can be projected back onto the original \( m \)-dimensional space and obtain another data matrix as follow:

\[
\hat{X} = TP^T = X(PP^T)
\]  

(4)

Considering \( \hat{X} \) as the projection of the data matrix \( X \) onto the selected \( r \) principal components and \( \tilde{X} \) as the projection onto the residual left components, the following decomposition can be performed:

\[
X = \hat{X} + \tilde{X}
\]  

(5)

\[
\hat{X} = X(PP^T)
\]  

(6)

\[
\tilde{X} = X(I - PP^T)
\]  

(7)

2.1. Damage Detection and Localization

There are several kind of fault detection indices [14]. Two well-known indices are commonly used to this aim: the \( Q - index \) (or SPE - index), the Hotelling’s \( T^2 - statistic \) (\( D - statistic \)). The first is based on analyzing the residual data matrix \( \tilde{X} \) to represent the variability of the data projection in the residual subspace[15]. The second method is based on the analysis of the score matrix \( T \) to check the variability of the projected data in the new space of the principal components. In general terms, every model is a representation of the original data and according to the accuracy of model, it can be a good representation with minimal errors. In any case, errors can be large or small and it is possible to quantify. Based on the previous statement, we can define the equation

\[
X_{original} = X_{model} + E
\]  

(8)

where, \( X_{original} \) are the original data, \( X_{model} \) are the projected data on the PCA model and \( E \) is the error in the model.

Due to the projected data can be represented as equation 9:

\[
X_{model} = TP^T = X(PP^T)
\]  

(9)

Replacing equation 9 in 8 and rearranging we can obtain the error (equation 10). This error is known as \( Q - statistic \).

\[
E = X(I - PP^T)
\]  

(10)

\( T^2 - statistic \) can be obtained from the concept of Euclidean distance normalized using as normalization factor the covariance matrix \( C \). The normalized Euclidean distance between two points is defined as:

\[
d_{ij} = [(x_i - x_j)^T M^{-1} (x_i - x_j)]
\]  

(11)

where, \( M \) is a diagonal matrix to normalize the variables. Using the equation 2 and replacing in 11, we obtain the equation

\[
T^2 = [(x_i - x_i)^T (P \Lambda^{-1} P^T)^{-1} (x_i - x_i)]
\]  

(12)

In a general way it is possible to define any index as it appears in the equation

\[
\text{Index} = x^T M x
\]  

(13)

Where the vector \( x \) represents measurements from all the sensors at a specific experiment trial, besides the matrix \( M \) depends of the type of index.
2.2. Contribution Methods for Localization

According to [14] five methods can be used for fault detection in process monitoring. Authors of this work adapted these methods for use in damage detection and localization in structures. These methodologies are used to calculate the contribution of each sensor to each index in each experiment trial. In this way, the damage will be located between actuator and sensor with largest contribution.

All the indices can determine if there are damages and distinguish between them, however they do not provide reasons for it. The main idea is to determine which variable or variables are responsible. The variables with the largest contribution are considered major contributors to the damage.

(i) Complete Decomposition Contributions (CDC)

Complete decomposition Contributions also called contribution plots are well known diagnostic tools for fault identification [16]. In each index is indicated the significance of the effect of each variable on the index. The contribution of the variable (or sensor) $j$ to the index is defined as:

$$\text{Index} = x^T M x = \| M^{\frac{1}{2}} x \|^2 $$  \hspace{1cm} (14)

$$\text{Index} = \sum_{j=1}^{n} (\xi_j^T M^{\frac{1}{2}} x)^2 = \sum_{j=1}^{n} CDC_{j}^{Index} $$  \hspace{1cm} (15)

$$CDC_{j}^{Index} = x^T M^{\frac{1}{2}} \xi_j^T M^{\frac{1}{2}} x $$  \hspace{1cm} (16)

where $\xi_j$ is the $j^{th}$ column of the identity matrix, and represents the direction of $x_i$.

(ii) Partial Decomposition Contributions (PDC)

This method decomposes a damage detection index as the summation of variable contributions.

$$PDC_{j}^{Index} = x^T M \xi_j^T x $$  \hspace{1cm} (17)

(iii) Diagonal Contributions (DC)

The diagonal contribution remove the cross-talk among variables. The DC is defined as:

$$DC_{j}^{Index} = x^T \xi_j^T M \xi_j \xi_j^T x $$  \hspace{1cm} (18)

(iv) Reconstruction Based Contributions (RBC)

The Reconstruction-Based Contribution [17] is an approach that uses the amount of reconstruction of a damage detection index along a variable direction as the contribution of that variable to the index. The RBC is defined as:

$$RBC_{j}^{Index} = x^T M \xi_j (\xi_j^T M \xi_j)^{-1} \xi_j^T M x $$  \hspace{1cm} (19)

$$RBC_{j}^{Index} = (\xi_j^T M x)^2 \frac{1}{(\xi_j^T M \xi_j)} $$  \hspace{1cm} (20)

(v) Angle-Based Contributions (ABC)

$$\xi_j = M^{\frac{1}{2}} \xi_j $$  \hspace{1cm} (21)

$$\bar{x} = M^{\frac{1}{2}} x $$  \hspace{1cm} (22)

The ABC of Variable $j$ is the squared cosine of the angle between

$$ABC_{j}^{Index} = (\frac{\xi_j^T \bar{x}}{\|\xi_j\| \|\bar{x}\|})^2 = \frac{(\xi_j^T M x)^2}{\xi_j^T M \xi_j x^T M x} $$  \hspace{1cm} (23)
Applying the measures explained in section 2.1 the methods defined in this section we obtain the results of the table 1.

| Table 1. Damage diagnosis methods. |
|-------------------------------------|
| Q | T² |
|----|---|
| CDC | \( x^TQ\bar{T}_j\xi_jQ\bar{T}_j^T\bar{x} \) | \( x^T\bar{T}_j\xi_jQ\bar{T}_j^T\bar{x} \) |
| PDC | \( x^TQ\xi_jQ\bar{T}_j^T\bar{x} \) | \( x^T\xi_jQ\bar{T}_j^T\bar{x} \) |
| DC | \( \xi_jQ\bar{T}_j^T\bar{x} \) | \( \xi_jQ\bar{T}_j^T\bar{x} \) |
| RBC | \( (\xi_jQ\bar{T}_j^T\bar{x})^2 \) | \( (\xi_jQ\bar{T}_j^T\bar{x})^2 \) |
| ABC | \( \frac{RBC_j^Q}{x^TQ\bar{x}} \) | \( \frac{RBC_j^Q}{x^T\bar{T}_j^T\bar{x}} \) |

According to [14] is possible to group these five methodologies in three general diagnosis methods. These are: General Decompositive Contributions, Reconstruction Based Contributions, Diagonal Contributions.

The complete and partial decomposition can be defined as special cases of General Decompositive Contributions. This can be defined as:

\[ \text{GDC}^\text{Index}_j = x^TM^{1-\beta}\xi_jT\bar{T}_j^T\bar{M}^\beta\bar{x}, \quad 0 \leq \beta \leq 1 \] (25)

When \( \beta = 0 \) or \( \beta = 1 \) \( \text{PDC} = \text{GDC} \) as is shown in the equations 26,29.

\[ \text{GDC}^\text{Index}_j = x^TM^1\xi_jT\bar{T}_j^T\bar{x}, \quad \text{if} \quad \beta = 0 \] (26)

\[ \text{GDC}^\text{Index}_j = x^TM^1\xi_jT(\bar{I})\bar{x}, \quad \text{if} \quad \beta = 0 \] (27)

\[ \text{GDC}^\text{Index}_j = x^TM^1\xi_j\bar{T}_j^T\bar{x} = \text{PDC}^\text{Index}_j, \] (28)

For \( \beta = 1 \) is possible to obtain:

\[ \text{GDC}^\text{Index}_j = x^TM^1\xi_j\bar{T}_j^T\bar{M}^1\bar{x}, \quad \text{if} \quad \beta = 1 \] (29)

Here, \( I = \Sigma_j(1)^{n}\xi_j\bar{T}_j^T \), then is possible to reorganize the equation to obtain:

\[ \text{GDC}^\text{Index}_j = x^TM\xi_j\bar{T}_j^T\bar{x} = \text{PDC}^\text{Index}_j \] (30)

In the same way, when \( \beta = \frac{1}{2} \) is possible to obtain CDC.

\[ \text{GDC}^\text{Index}_j = x^TM^{1-\frac{1}{2}}\xi_j\bar{T}_j^T\bar{M}^{\frac{1}{2}}\bar{x}, \quad \text{if} \quad \beta = \frac{1}{2} \] (31)

\[ \text{GDC}^\text{Index}_j = \text{CDC}^\text{Index}_j \] (32)

Since ABC is a scaled version of RBC, it is possible to use RBC as a general case for both diagnosis methods.
3. Experimental Setup

The application of the methodologies for damage localization was tested in an aluminium plate (figure 1). The dimensions of this plate is $25cm \times 25cm \times 0.2cm$. This plate was instrumented with four Piezoelectric transducer discs (PZT’s) attached on the surface. In previous work [18], it was conclude that the farther is the damage from the actuator, the more difficult is its detection. Therefore, to detect damage on a larger area and being useful the fact that PZT’s can be used as much as actuators as sensors, the experiment to assess the structure is performed in several phases. In every phase, just one PZT is used as actuator (an electrical excitation signal is applied) and the others are used as sensors. A real damage was made between PZT2 and PZT4 as is shown in the figure 2, 300 experiments were performed and recorded: 100 using the undamaged structure, and 200 using the structure with different size of the damage (increasing the depth). The PCA model was created using 100% of the whole dataset collected using the undamaged structure, 50% of the dataset of undamaged structure and the whole dataset of the damaged structure were used for testing the approach.

![Figure 1. Aluminium plate used](image1)

4. Damage Localization

To detect and distinguish damages a PCA model is built using the data collected in the different phases, each phase (PZT1 as actuator, PZT2 as actuator, and so on) use the signals recorded by sensors during the experiments with the undamaged structure. Data from experiments using the current structure (damaged or not) is projected on the model (figure 3). Projections onto the first principal components (scores), $T^2$-index and Q-index are calculated by each PCA model.
To localize the damage, the contribution of each sensor in the different phases to each index in each phase is calculated and finally are added in order to obtain one measure that shown the region with more abnormalities, this contribution is a measure of the level of influence of the damage to the sensor. Therefore, it is expected that the damage is located between the actuator and the most influenced sensor. In each phase, a region of the structure is selected as the region where the damage is located. Considering all phases, a general diagnosis could be performed (intersection of all the regions).

5. Experimental Results
The five methods explained in section 2.2 were applied to each index and compared using the damage located between PZT 2 and PZT 4 with different depth, figures 4-8 shows the results of using the DC, CDC, PDC, ABC and RBC for the $Q$ index and damage 1, the higher the value of the color, the more probability of the localization of the damage. As it is expected, the damage is located between PZT2 and PZT4.
Figure 6. damage 1 using PDC

Figure 7. damage 1 using ABC

Figure 8. damage 1 using RBC

Figure 9. damage 1: comparison of methods

Figure 10. damage 2 using DC

Figure 11. damage 2 using CDC

Figure 9 shows the contributions obtained for each path between the different PZT’s for the five methods. As shown, in each method, the path between PZT 2- PZT 4 contains the highest values of contribution, this is because the damage is located in this path. With all the methods
is possible to locate the damage with different values, the lowest contribution is found with CDC method, additionally, it is possible to see that in all the methods the difference between the values of each path for each method are significantly greater for the path between PTZ 2 and PZT 4 except for CDC method where the values are similar.

**Figure 12.** damage 2 using PDC  
**Figure 13.** damage 2 using ABC

The same five methodologies are applied for the $T^2 - index$ in the same damage with different depth, results are shown in figures 10-14, the higher the value of the color, the more probability of the localization of the damage. As it is expected, the damage is located between PZT2 and PZT4.

Figure 15 shows the contributions obtained for each path between the different PZT’s for the five methods. As for the $Q - index$, in each method, the path between PZT 2- PZT 4 contains the highest values of contribution, it happens because the damage is in this path. In this case, comparing with the results for $Q - index$, exist more differences between the different paths using CDC method.

**Figure 14.** damage 2 using RBC  
**Figure 15.** damage 2: comparison of methods
6. Conclusions
Five different methods (DC, RBC, ABC, PDC, CDC) for damage localization were presented using $T^2$ and $Q$-index. The methodology developed to localize damages is based in PCA models built from vibrational responses of the structure using a active system [9].

The region which contain the damage is obtained by finding the highest value area, for this, the sum of the contributions obtained for each sensor to each index is calculated.

For the $T$–index and $Q$–index the lowest contributions were obtained with the CDC method, in the $Q$–index case, the results obtained with the other methods present similar results, but in the $T$–index case was obtained best result using ABC method. Also was found that in the CDC method for $Q$–index the values for each path between the sensors present similar results.

Acknowledgments
This work has been supported by the ”Ministerio de Ciencia e Innovación” in Spain through the coordinated research project DPI2008-06564-C02-01/02. The authors would like to thank the support from the ”Agència de Gestió d’Ajuts Universitaris i de Recerca” of the ”Generalitat de Catalunya”, ”Escola Universitaria d’Enginyeria Tècnica Industrial de Barcelona (EUETIB)”, and ”Universitat Politècnica de Catalunya”.

References
[1] Chang P, Flatau A and Liu S 2003 Structural Health Monitoring 2 257–267
[2] Azevedo S, Mast J, Nelson S, Rosenbury E, Jones H, McEwan T, Mullenhoff D, Hugenerber R E, Stever R D, Warhus J P and Wieting M G 1996 SPIE 294(6), 195304.
[3] Lee J W, Kim J D, Yun C B, Yi J H and Shim J M 2002 Journal of Sound and Vibration 257 247–264
[4] Garziara R, Amabili M and Collini L 2007 IV Conferencia Panamericana de END Buenos Aires
[5] H Sohn K L 2000 Probabilistic Engineering Mechanics 15 139–153
[6] Mujica L, Vehi J, Rodellar J and Kolakowski P 2005 Smart Materials and Structures 14 1554–1562
[7] Sohn H, Farrar C R, Hunter N and Worden K 2001 Applying the lanl statistical pattern recognition paradigm for structural health monitoring to data from a surface-effect fast patrol boat Tech. Rep. LA-13761-MS Los Alamos National Laboratory
[8] Staszewski W, Boller C and Tomlinson G 2004 Health Monitoring of Aerospace Structures: Smart Sensor Technologies and Signal Processing (Chichester: John Wiley & Sons, Ltd) URL http://www3.interscience.wiley.com/cgi-bin/bookhome/107064020/
[9] Burgos D, Mujica L, Guemes A and Rodellar J 2010 European Workshop on Structural Health Monitoring, 5 th edition. Sorrento-Italy.
[10] Jolliffe I 2002 Principal Component Analysis (Springer)
[11] Worden K and Farrar C 2007 Philosophical Transactions of the Royal Society, Series A 365(1861) 299–632
[12] Mujica L, Ruiz M, Guemes A and Rodellar J 2009 Proc of the 4th. ECCOMAS thematic conference on smart structures and materials. Porto, Portugal
[13] Li G, Qin S, Ji Y and Zhou D 2009 7 th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Process. Barcelona-Spain.
[14] Alcala C and Qin S 2009 7 th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes. Barcelona-Spain.
[15] Mujica L, Rodellar J, Fernandez A and Guemes A 2010 Structural Health Monitoring online version 1–15
[16] Mnasri B, Adel E E and Ouladsine M 2008 16th Mediterranean Conference on Control and Automation Congress Centre, ajaccio, France. June 25-27
[17] Alcala C and Qin J 2009 Automatica 45 1593–1600
[18] Mujica L, Tibaduiza D and Rodellar J 2010 5 th World Conference on Structural Control and Monitoring. Tokio-Japan.