A robust procedure for Damage detection from strain measurements based on Principal Component Analysis

A. Guemes[1, a], J. Sierra-Pérez[2], J. Rodellar[3] and L. Mujica[3]

[1]Department of Aerospace Materials and Processes, Universidad Politécnica de Madrid, Madrid, Spain.

[2]Aerospace Engineering Research Group, Universidad Pontificia Bolivariana, Medellin, Colombia.

[3]Department of Applied Mathematics III, Univ. Politécnica de Catalunya, Barcelona, Spain

[a]alfredo.guemes@upm.es (corresponding author)

Keywords: Principal component Analysis, fiber optic sensors.

Abstract. FBGs are excellent strain sensors, because of its low size and multiplexing capability. Tens to hundred of sensors may be embedded into a structure, as it has already been demonstrated. Nevertheless, they only afford strain measurements at local points, so unless the damage affects the strain readings in a distinguishable manner, damage will go undetected. This paper show the experimental results obtained on the wing of a UAV, instrumented with 32 FBGs, before and after small damages were introduced. The PCA algorithm was able to distinguish the damage cases, even for small cracks. Principal Component Analysis (PCA) is a technique of multivariable analysis to reduce a complex data set to a lower dimension and reveal some hidden patterns that underlie.

Introduction

Structural Health Monitoring (SHM) aims for automatic procedures for the assessment of the load conditions and damage occurrence in structures, to ensure proper performances during service life. When fully implemented, it will allow a change of paradigm, from scheduled maintenance to on-condition maintenance, with significant savings. The basic characteristic of SHM that distinguish it from conventional Non Destructive Evaluation (NDE) is that sensors are permanently attached to the structure. Therefore, sensors must be working for the whole operational life of the structure without requiring operator intervention. A useful system usually requires a large number of sensors distributed throughout the structure. The large amount of data produced must be automatically processed. Warning signals need to be filtered to the user when there is a structural overload, initiation of damage, or to produce reports on the fatigue accumulated by the structure.

For thin shells and similar aircraft structures, Lamb waves produced and detected by a PZTs network is the commonest approach. Comparatively few articles are using FBGs, due to the local nature of the sensor, because it only detects strains at its position. Unless there is a coincidence of the sensor location and damage initiation point, the local crack would produce a quite small change into the general strain field, and may go undetected. Only by comparing the strains readings at several positions (technique called differential strains) some information may be obtained, but the procedure is very manual. PCA offers a more robust approach to handle the strain signals.

Fiber optic sensors

Concerning fiber optic sensors, the Bragg gratings are the mostly used today because of its high performances and comparative advantages compared to other strain measurement techniques. Some of this advantages are: high sensitivity, small size that allow to embed it into composite materials, low weight, less signal degradation (immunity to electromagnetic interference and radio frequency), low power consumption, non-flammable, user friendly, moderate cost, high operating temperatures, high fatigue resistance.

The Bragg gratings are regions in the optical fiber where the core has a periodic variation of its refractive index with a period Λ. Such modulation is induced in a special type of optical fiber (photosensitive) by exposing the core to light from an ultraviolet laser. The laser interacts with a diffraction grating (phase mask) of determined wavelength (period), placed between the laser and the optical fiber. In this way, a Bragg grating with a given period and a length ranging between 1 and 20 mm is 'written' on the fiber optics.

When light goes through a FBG, either from a white light source or from a tunable laser, the FBG behaves as a narrow filter, reflecting back only a specific wavelength, 2 n Λ, where n is the average refractive index of the O.F., and Λ the formerly said spacing period. By varying Λ, several FBGs can easily been engraved and interrogated at the same optical fiber.

The Bragg gratings can be used as strain and temperature sensors. To be employed as strain sensors, the FBGs must be bonded or embedded into the material where the strain measurement is sought. Thus, the FBG will deform with the substrate to which it has been bonded. As a result of the strain, the modulation period Λ will change, as well as the refractive index, and consequently, the reflected wavelength. Similar effects happen with temperature, the strain readings needs to be compensated from thermal drifting, as with electrical strain gages.

Principal component analysis

The measurements performed in SHM techniques normally use different sensors, measuring continuous dynamic signals as a function of time. Therefore, it is necessary to perform a treatment of experimental data prior to application of the PCA technique. In its first instance it is necessary to discretize the signals in order to obtain a manageable data set. A X matrix with all data information from measuring several variables at a number of time instants (one experiment) can be arranged as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{im} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{nm} \end{bmatrix} = \begin{pmatrix} v_1 & v_2 & \cdots & v_j & \cdots & v_m \end{pmatrix} \quad (1)$$

This matrix contains information from n experimental trials × m variables or sensors.

If more than one experiment is done, at the end, a tridimensional matrix (X_{3D}) with all the obtained information can be arranged as follows: I experiments × K samples per experiment × J sensors (see Figure 1). Each frontal slice represents all measurements of one sensor for the whole experiment. This matrix must be unfolded for perform a PCA study. [1]

There are several ways to unfold 3D data arrays in the literature; each one allows studying a different kind of variability by means of the principal component analysis. According to Nomikos and MacGregor [2], the most used way to unfolding 3D data arrays is the called 'D type unfolding'.

The main reason is unfolding in this way the whole batch is consider as an object and it is possible to compare batches between them. i.e. baseline for healthy structure with subsequent states of the structure during operation [2], [3]. The methodology for 'D type unfolding' is illustrated in Figure 1.

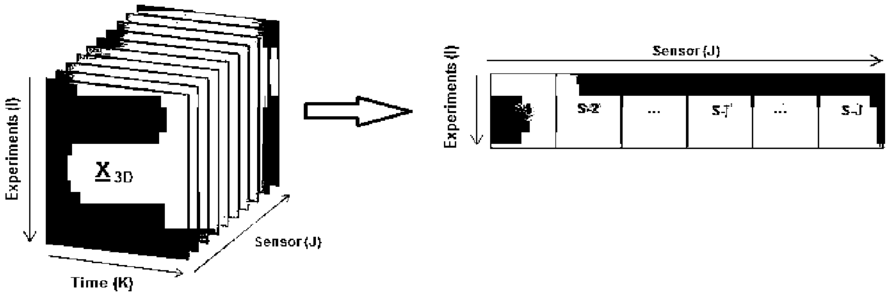

Figure 1. `Type D' unfolding model.

Physical variables present in most experiments do not have the same magnitudes and scales. It is necessary to treat the initial experimental data before any statistical analysis takes part. In the literature there are a variety of techniques to rescale experimental data. For unfolded matrices several scaling techniques has been studied. Among the most used techniques are the Continuous Scaling (CS), Variable Scaling (VS), Group Scaling (GS), Auto Scaling (AS), etc. However, the most common method used for unfolded matrices is the Group Scaling. The main reason is Group Scaling considers the interaction between different sensors and process them alltogether. All the methods are rescaling each of the variables so that they have a mean of zero magnitude and a same variance or variance equal to one (depends on technique). [1]-[7].

Once the matrix X_{3D} is unfolded into a new matrix X, centered and scaled, a PCA study is performed based on the covariance matrix which quantifies the degree of linearity between all possible pairs of variables. Then, is possible to order the eigenvectors associated to the covariance matrix in descending order (according to their associated eigenvalues), this way, only a few principal eigenvectors can be selected; these represent the more important system dynamics. As result, an important dimension reduction can be obtained.

The covariance matrix is given by:

$$C_{\bar{X}} = \frac{1}{n-1} X^T X \tag{2}$$

The main diagonal terms of the covariance matrix are the variance and the off diagonal terms represent the covariance between pairs of variables:

$$\sigma_{\upsilon_j}{}^2 = \frac{1}{n-1} \upsilon_j^T \upsilon_j = \frac{1}{n-1} \sum_{i=1}^{n} x_{ij}{}^2 \tag{3}$$

$$\sigma_{\upsilon_j,\upsilon_k}{}^2 = \frac{1}{n-1} \upsilon_j^T \upsilon_k = \frac{1}{n-1} \sum_{i=1}^{n} x_{ij} x_{ik} \tag{4}$$

Once the covariance matrix is obtained, the data matrix X can be transformed using a linear transformation in order to achieve the minimal redundancy.

$$T = XP \quad (5)$$

This linear transformation must be such that the new data matrix T is diagonal, i.e.:

$$C_T = \frac{1}{n-1} P^T X^T XP = P^T C_{\overline{X}} P \quad (6)$$

The transformation matrix is selected for having their eigenvectors by columns, i.e.:

$$P = (p_1 \; p_2 \cdots p_j \cdots p_m) \quad (7)$$

Since eigenvectors are ordered according to the amount of information, the dimensionality of the data matrix X can be reduced if only a certain number of r principal components are chosen.

$$P_r = (p_1 \; p_2 \cdots p_r) \quad (8)$$

The T matrix (called 'score matrix') has uncorrelated row vectors and its column vectors are the projection of the original data over the direction of the jth principal component. This column vectors are called 'scores'.

A baseline must be constructed using data for the healthy structure. This means to calculate matrix P_r for the healthy structure. Later, results for an unknown structure condition (X) should be projected into the baseline model. The methodology is outlined in the Figure 2.

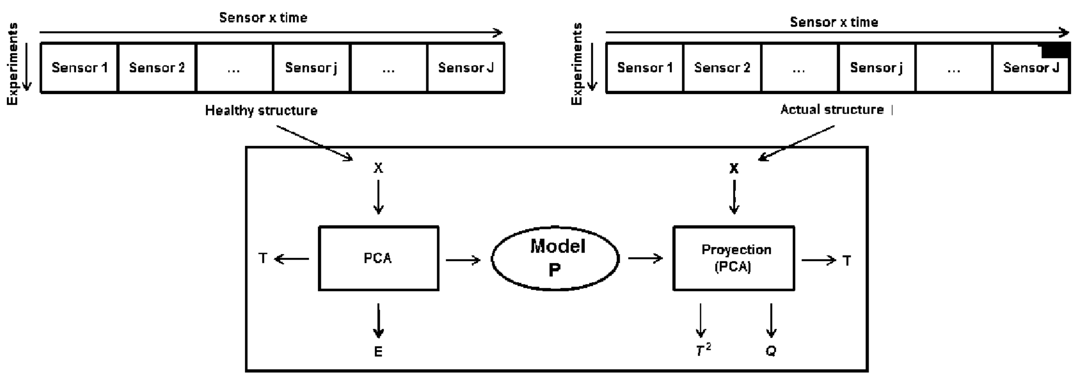

Figure 2: Scheme of PCA model for structural damage identification.

$$T = XP_r \quad (9)$$

From this projection, is possible to calculate different damage indices and detection thresholds [6].

There are statistical tools that used along with PCA, allow detection of anomalous behavior in systems. The two most common tools are the Q index (or Squared Prediction Error index) and the T^2 index (or D index). The index Q indicates how well each sample fits the PCA model. It is a measure of the difference between a sample and its projection in the main components retained by the PCA model. [8], [5], [9].

The T^2 index is a measurement of the variation of each sample in the PCA model. It is based on analysis of the score matrix (T) which allows to study the variability of the projected data in the new principal components space. [10].

Q index is given by:

$$Q_i = \tilde{x}_i \tilde{x}_i^T = x_i (I - PP^T) x_i^T \tag{10}$$

Where, $\tilde{x}_i$ is the projection into the residual subspace.

And the T^2 index is given by:

$$T_i^2 = \sum_{j=1}^{r} \frac{t_{sij}^2}{\lambda_j} = \frac{t_{si} \, t_{si}^T}{\Lambda} = \frac{x_i PP^T x_i^T}{\Lambda} \tag{11}$$

Experimental set-up

To validate the methodology, a 1.5 meters wing section fully made in composite materials, belonging to an unmanned air vehicle, was used. Two fiber optics were bonded at the intrados and two at the extrados, each one having 8 FBGs. In total, 32 FBGs were used. The wing was fixed to a testing bench by mean of screws in the same way is fixed to the fuselage on the aircraft. Once the wing was fixed, the testing phase begun. The first step consisted in the model building (baseline) for the healthy structure. Each experiment consisted in loading the structure in bending mode, progressive loading from zero load to a specific load. After waiting 10 seconds for the load stabilization, the load was removed progressive. The sampling rate was 10 Hz. 4 different loads were used in bending mode. Each load case and the zero load case were repeated 10 times for the baseline model building and 10 times more for model validation. That is, additional data for the healthy structure were taken in order to validate the baseline model. Figure 3 illustrate the used wing section. The zero load cases were useful in order to verify that no residual stresses appeared after damages promoting.

A Micron Optics Si 425 interrogator was used for data acquisition. This equipment has a tunable laser and can interrogate up to 512 optic sensors in four channels at same time at maximum sampling rate of 250 Hz with a resolution less than 0.2 pm.

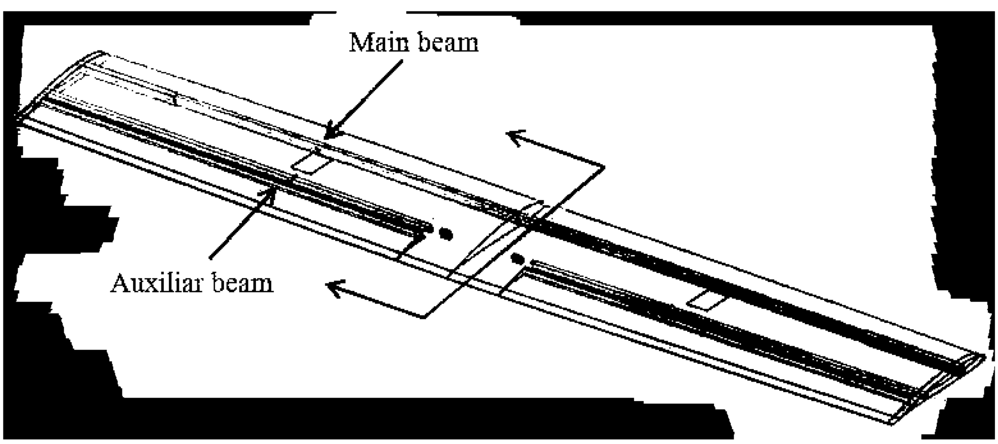

Figure 3. Scheme of wing section used.

After building the model for the healthy structure, two kinds of accumulative artificial damages were induced in the locations schemed in Figure 4. The first damage case consisted in a longitudinal skin cutting of 1 cm. The second damage case consisted in increase the size of the first crack to 3 cm. The third damage case consisted in a transversal skin cutting of 1 cm without cutting the spar cap. From the fourth to the seventh cases, the transversal crack was increased 1 cm each time and the spar cap was also cut superficially. Again, for each damage case, the zero load and the four different load cases were used and each experiment was performed 10 times. In total, 400 experiments were performed, each one consisting in the signal of 32 sensors in more than 400 instants of time.

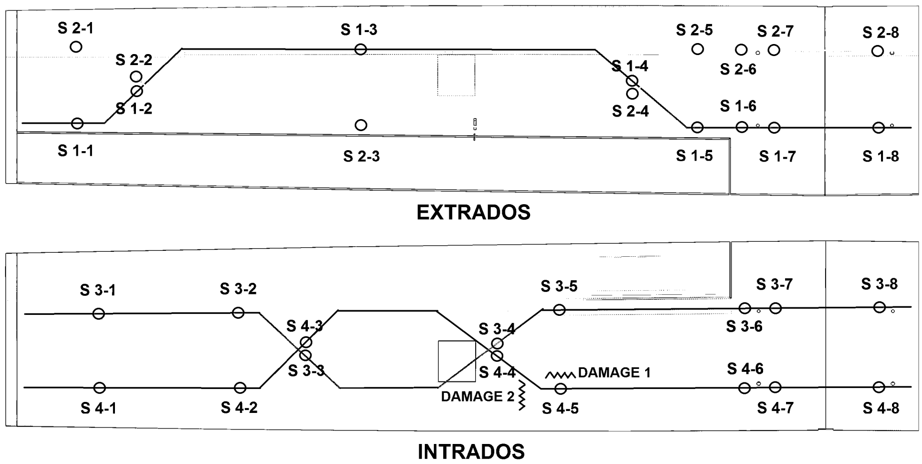

Figure 4. Sensors and induced damages locations.

From the signal of each sensor, the initial zero load region and the stable load region (zone 1 and 2 respectively in Figure 5) were isolated and preprocessed in order to remove outliers. The Ferguson test was used (kurtosis coefficient given by equation (12)) (ASTM E178 1972).

$$b_2 = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n-1)^2 s^4} = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2} \quad (12)$$

Where s represents the standard deviation, $\bar{x}$ represents the arithmetic average and n represents the number of samples.

To apply the test proposed by Ferguson, the b_2 value must be computed and if the value exceed the desired significance level, the observation farthest from the mean is rejected and same procedure is repeated until no more values are judged to be outliers.

After removing the outliers, the average of zone 1 in Figure 5 was taken as reference (initial wavelength) and the average of zone 2 in same figure, was taken as final wavelength for strain calculation. In this way, since the experiments were performed in room with controlled temperature and each one takes no more than 45 seconds, thermal effects can be neglected. No wavelength shifting between zone 1 and zone 3 (Figure 5) means no thermal effects and no residual stresses appears during the experiments.

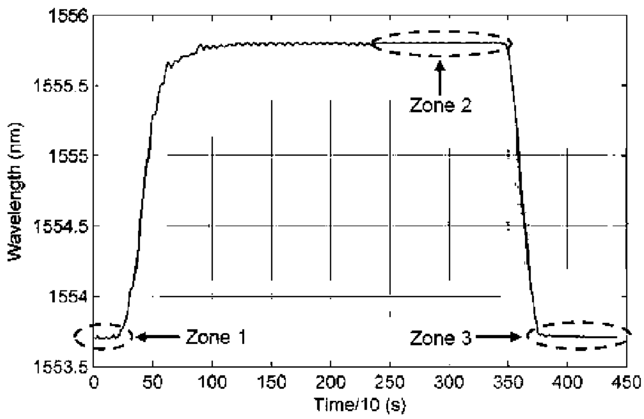

Figure 5. Example of load spectrum and interest zones.

The strains for each experiment were calculated using the equation (13).

$$\frac{\Delta\lambda_B}{\lambda_B} = (1-\rho_\alpha)\Delta\varepsilon + (1+\xi)\Delta T \tag{13}$$

Where λ_B is the Bragg wavelength, ρ_α is the photo elastic coefficient of the fiber optics and ε is the thermo optic coefficient for the fiber optics.

For the same kind of fiber optics used in this work, García obtained the experimental constants given in equations (14) and (15) [11].

$$\Delta\varepsilon = (803.9 \pm 5.6)\frac{\mu\varepsilon}{nm}(\Delta\lambda) \rightarrow k_\varepsilon = (0.7991 \pm 0.0055)\mu\varepsilon^{-1} \tag{14}$$

$$\Delta T = (101.9 \pm 1.2)\frac{K}{nm}(\Delta\lambda) \rightarrow k_T = (6.334 \pm 0.074)\times 10^{-6}\, K^{-1} \tag{15}$$

Where $k_\varepsilon = 1-\rho_\alpha$ and $k_T = 1+\xi$

Before proceeding to PCA, the data was unfolded using the methodology schematized in Figure 1, centered and scaled using Group Scaling (GS). By mean of Group Scaling each data is scaled using the mean of all measurements of a sensor at the same instant of time and the standard deviation of all measurements of a sensor as follows: [1], [12].

$$\bar{x}_{ijk} = \frac{x_{ijk} - \mu_{jk}}{\sigma_j} \tag{16}$$

Where x_{ijk} is the kth sample of the jth sensor in the ith experiment, μ_{jk} is the mean of all kth samples of the jth sensor, μ_j is the mean of all measurements of the jth sensor, σ_j is the standard deviation of all measurements of the jth sensor and $\bar{x}_{ijk}$ is the scaled sample. The coeficients μ_{jk}, μ_j and σ_j are given by:

$$\mu_{jk} = \frac{1}{I}\sum_{i=1}^{I} x_{ijk} \qquad (17)$$

$$\sigma_j = \sqrt{\frac{1}{IK}\sum_{i=1}^{I}\sum_{k=1}^{K}(x_{ijk} - \mu_j)^2} \qquad (18)$$

$$\mu_j = \frac{1}{IK}\sum_{i=1}^{I}\sum_{k=1}^{K} x_{ijk} \qquad (19)$$

Finally the PCA was performed for the healthy structure and the different damage cases were projected onto the model. In this study was selected a number of 20 principal components to build the model.

Analysis of the results

At figure 6, a few of the raw strain data are plotted, for the undamaged and a damage case, and for a load condition. Differences in strain values are very small, and both cases can hardly been distinguished.

For all the damage cases the scores 1 against scores 2 were plotted. The damage indices Q and T^2 were calculated and plotted for each case and one versus the other were also plotted.

In order to verify the baseline model, data for healthy case was projected into the PCA model (for all the load cases). These additional data were no taken into account to build the PCA model. In this way it was possible to compare between real damage cases and healthy case, when the data was projected into the model.

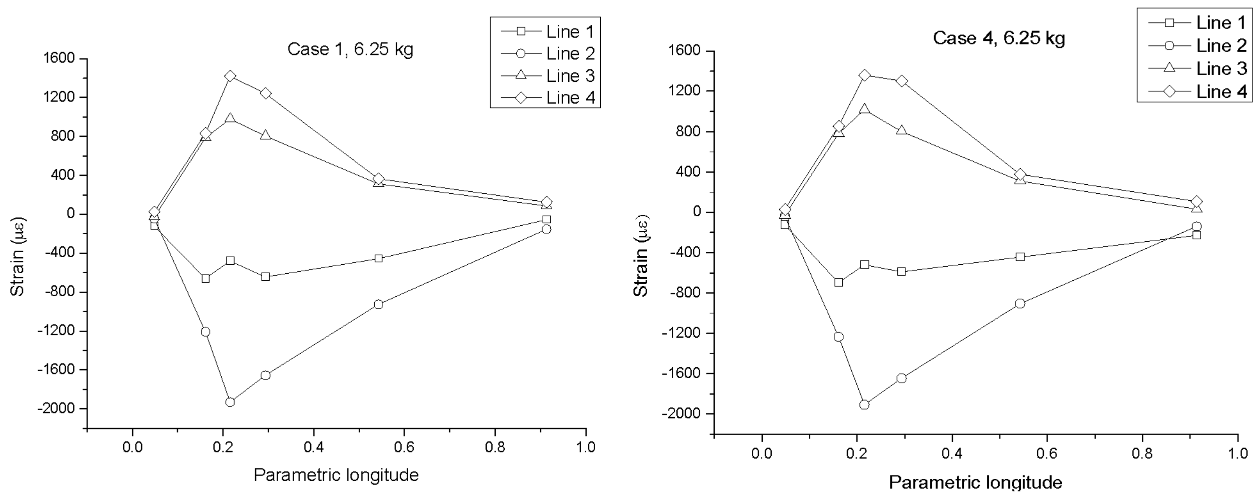

Figure 6. Raw strain data for the undamaged and a damaged case (2 cm length)

Figure 7 shows the projection into the first principal component against the projection into the second principal component. Trying to represent more than two principal components in single plot is more complex and getting valuable information from this can be difficult.

Looking at Figure 7, it is possible to distinguish between four different segregated groups of data. These data groups corresponding to each of the various loading conditions studied. Inside each segregated group is possible to distinguish tendencies for different damage cases. If it is desired to get more precise information is necessary to study each of these planes containing the different principal component projections. As Westerhuis shown, the sum of the variances of the two first principal components, exceed 80% of the original data variance [3].

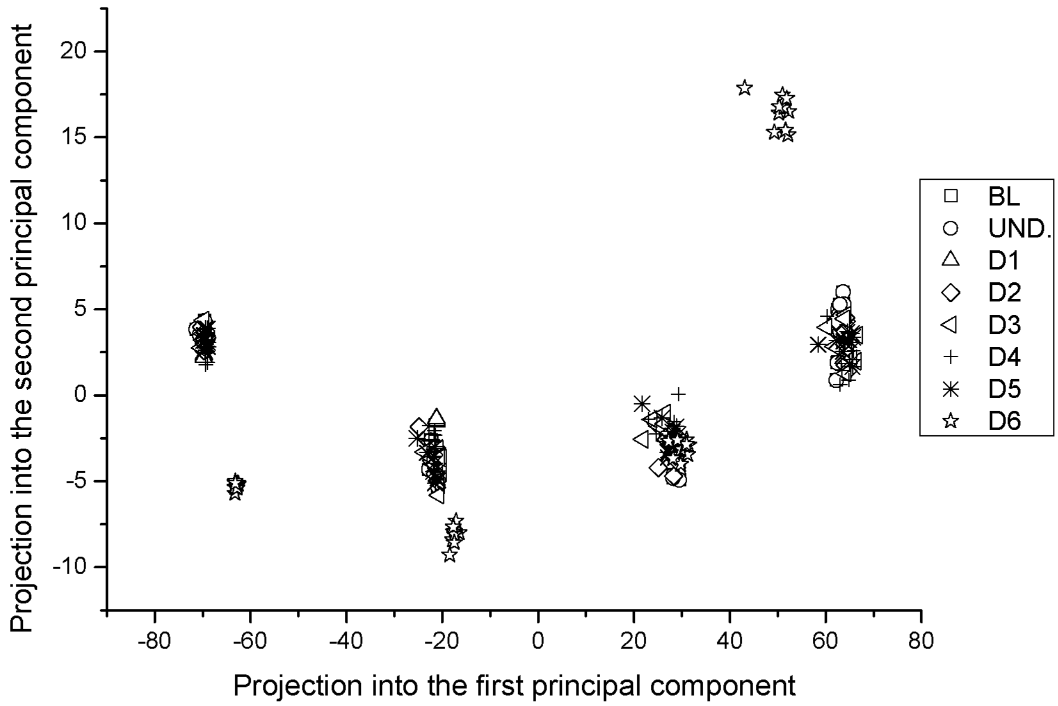

Figure 7. Projections into the first two principal components for all cases.

In Figure 7 is evident a separation between the baseline and different damage cases and a very good fit between the baseline and the undamaged case.

Due to the apparition of non linear effects during some load cases, the differences between baseline and damage cases are not the same for the four load cases studied. In the third data group (from the left to the right in Figure 7) for example, the data corresponding to the most severe damage case is closer to other data than in the other three load cases.

At figure 8 both damage index are represented, for a load condition. The Q index distinguish the undamaged case from the others, and for case 6, when damage affects not only the skin but also the main spar, the damage identification is neat,

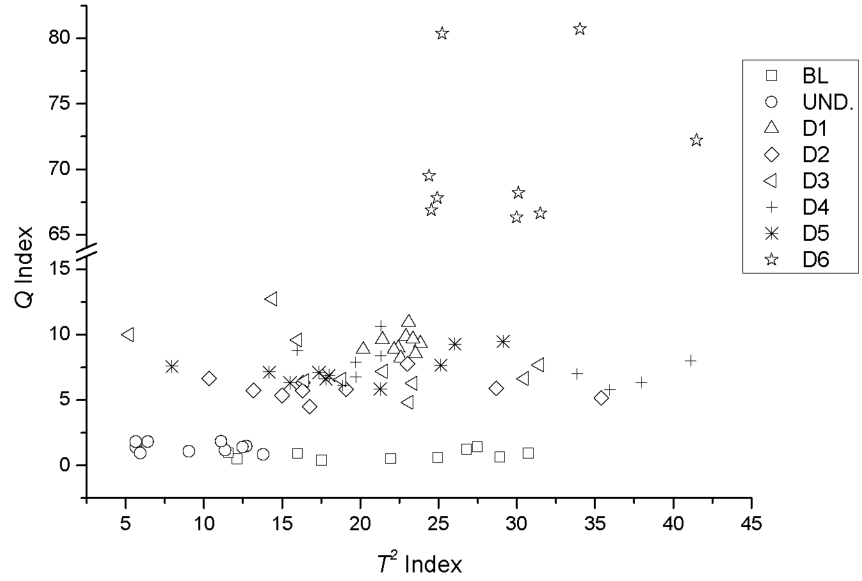

Figure 8. Damage indexes for all damage cases.

Conclusions

A PCA model is built using the signals recorded by strain FBGs sensors during the experiments with the undamaged structure. PCA modelling essentially consists of calculating the matrix **P**. During the subsequent step, the experiments are performed using the structure in the different possible states (undamaged and 6 damages). These signals are projected on the PCA model, thus obtaining a selected number of the first principal components (scores-**T**). In addition, the Q-statistic and T^2-statistic are calculated. The approach has been experimentally analyzed showing good results in classifying different states of the structure: healthy structure and six different damages.

Acknowledgement

The authors would like to thank the support from the "Ministerio de Ciencia e Innovación" in Spain through the coordinated project DPI2011-28033-C03-02

References

[1] Kourti, T, and John MacGregor. Process analysis, monitoring and diagnosis, using multivariate projection methods. Chemometrics and intelligent laboratory systems, (1995) 3-21.

[2] Nomikos, Paul, and Jhon F. MacGregor. Monitoring batch Processes Using Multiway Principal Component Analysis. AIChE Journal, (1994) 1361-1375.

[3] Westerhuis, Johan A, Theodora Kourti, and Jhon F MacGregor. Comparing alternative approaches for multivariate statistical analysis of batch process data. Journal of Chemometrics, (1999) 397-413.

[4] Gurden, S, J Westerhuis, R Bro, and A Smilde. A comparision of multiway regression and scaling methods. Chemometrics and Intellingent Laboratory Systems. 59 (2001) 121-136.

[5] Villez, K, K Steppe, and D De Pauw. Use of unfold PCA for on-line plant stress monitoring and sensor failure detection. Biosystems Engineering, (2009) 23-34.

[6] Mujica, L, D Tibaduiza, and J Rodellar. Data driven multiactuator piezoelectric system for structural damage localization. Fifth world conference on structural control and monitoring (2010).

[7] Wold, S, N Kettaneh, H Friden, and A Holmberg. Modelling and diagnostics of batch processes and analogous kinetic experiments. Chemometrics and intelligent laboratory systems. (1998) 331-340.

[8] Jackson, E, and G Mudholkar. Control procedures for residual associates with PCA. Technometrics, (1979) 341-349.

[9] Burgos, D.A, L Mujica, A Güemes, and J Rodellar. Active piezoelectric system using PCA. Fifth European Workshop on Structural Health Monitoring. (2010) 164-169.

[10] Mujica, L.E., J. Rodellar, A. Fernandez, and A. Guemes. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. Structural Health Monitoring.(2010) 1-15.

[11] García, Carlos E. Caracterización de coeficientes de Expansion termica. Informe Técnico, (2010).

[12] Sierra, J, and A Güemes. Detección de daño en materiales compuestos mediante fibra óptica. Actas del IX congreso nacional de materiales compuestos. Girona: AEMAC, 2011.