Principal Component Analysis and Self-organizing Maps for Damage Detection and Classification under Temperature Variations

MARIBEL ANAYA VEJAR, DIEGO A. TIBADUIZA-BURGOS, MIGUEL A. TORRES-ARREDONDO and FRANCESC POZO

ABSTRACT

The use of statistical techniques for data driven has proven very useful in multivariable analysis as a pattern recognition approach. Among their multiple advantages such as data reduction, multivariable analysis and the definition of statistical models built with data from experimental trials, they provide robustness and allow avoiding the need of the development of physical models which sometimes are difficult for modelling especially when the system is complex. In this paper, a methodology based on Principal Component Analysis (PCA) is developed and used for building statistical baseline models comprising the dynamics from the monitored healthy structure under different temperature conditions. In a second step, for testing the proposed methodology, data from the structure at different structural states and under different temperature conditions are projected into the baseline models in order to obtain statistical measures (Scores and Q-index) which are included as feature vectors in a Self-Organizing Map for the damage detection and classification tasks.

The methodology is evaluated using ultrasonic signals collected from an aluminium plate and a stiffened composite panel. Results show that all the simulated states are successfully classified no matter what the kind of damage or the temperature is present in both structures.

INTRODUCTION

Structural Health Monitoring (SHM) is an important area which seeks to assess and monitor structures in order to ensure the reliability, and security of service in its continuous operation. This is done by the use of sensors, which are permanently installed in the structure and computational algorithms to evaluate the information from the structure. In SHM there are four levels of damage diagnosis [1]. The first

Maribel Anaya and Diego A. Tibaduiza, Faculty of Electronics Engineering, Universidad Santo Tomás, Bogotá, Colombia.

M.A. Torres-Arredondo, MAN Diesel & Turbo SE, Engineering 4 Stroke-Test and Validation-Test Bed, Management-Measurement (EETTM), Augsburg, Germany.

Francesc Pozo, CoDAlab, Departament de Matemàtica Aplicada III, Escola Universitària d'Enginyeria Tècnica Industrial de Barcelona (EUETIB), Universitat Politècnica de Catalunya (UPC), Comte d'Urgell, 187, 08036 Barcelona, Spain.

level includes the damage detection and its main objective is to know whether there is any abnormality in the structure and if this abnormality corresponds to damage. The second level includes the damage localization and allows determining the position of the damage in the structure. Third step includes classification tasks to define the type of damage and its size. Finally, the level 4 is focused on knowing the structure remaining lifetime. In general, the damage identification can be performed following two main approaches. The first approach consists in obtaining a reliable physics-based model of the structure, while the second is based on data-driven approaches that normally tackle the problem as one of pattern recognition. One advantage of the use of data-driven approaches is the reliability in the analysis since the indication of damage could be directly determined with the comparison between a baseline and the collected data. However, to ensure the reliability of the analysis performed to the collected signals in several experiments, it is necessary to consider variations in the environmental and operational conditions, and the proper functioning of the sensors, actuators and hardware that are used to inspect the structure. This paper is addressed to the levels one and three of the damage identification task by means of a methodology which uses a baseline with several data from different temperatures to classify structural states and fault sensor damages. As several works have shown, in metallic and composite materials the temperature influence is presented in the wave propagation. In fact, wave propagation is affected when the elastic properties and density of the propagation medium are shifted. In ultrasonic signals, changes in the temperature stretch or compress the signal and distort the shape [2],[3],[4]. These changes directly affect the performance of the methodologies and need to be considered. The methodology presented in this work considers the use of Discrete Wavelet Transform (DWT), Multiway Principal Component Analysis (MPCA) and Self-Organizing Maps (SOM) to solve the temperature changes problem and it is tested using an aluminium plate instrumented with five piezoelectric transducers and a simplified aircraft composite skin panel which is instrumented with four piezoelectric transducers. On the one hand, seven different states are studied in the classification in the first structure: the undamaged state, 4 simulated damages (structural modifications by adding a mass at different positions of the structure) and two fault sensor (sensor breakage at 25% and 50%). On the other hand, six structural states are studied in the second structure: these states include the undamaged structure and 5 structural modifications that are simulated by means of adding a mass in different locations. The experimental results show that all these states are successfully detected and classified no matter the kind of damage or the temperature in both structures.

DAMAGE DETECTION METHODOLOGY

The proposed methodology considers the use of a multi-actuator piezoelectric system (distributed piezoelectric active network) working in several actuation phases, Discrete Wavelet Transform, Multiway Principal Component Analysis (PCA), SPE-index and Self-Organizing Maps for the classification of different structural states by considering temperature changes using robust baselines.

The signals propagated through the structure are collected in different points using the rest of the sensors and pre-processed by using the DWT. In this work the family of Daubechies wavelet basis function 'db8' was chosen for the methodology presented here since it proved to be adequate to encode and approximate the ultrasonic waveforms. Figure 1 shows the scheme of the methodology. In a first step, when working with one temperature, the dynamic responses collected from each actuator phase are stored by the data acquisition system into a matrix with dimensions (I×K), where I represents the number of experiments and K the number of sampling times. Denoting J as the number of PZT transducers that are receiving the dynamical responses for each experiment, J matrices with the information from each sensor by each actuator phase are finally stored. Therefore, the whole set of the data collected in each actuator phase and with a specific temperature can be organized in a three-dimensional matrix (I×K×J) or in a two-dimensional unfolded matrix (I×JK), where data from each sensor are located besides the other sensors. To manage the use of different temperatures, in this paper the data (approximation coefficient) from each temperature are unfolded and organized in a matrix in order to obtain a matrix by each actuator phase with the information from all the sensors at different temperatures.

The collected data in each actuation phase must be pre-processed before the computation of the PCA model. For this kind of data sets (unfolded matrices), several ways of scaling have been presented in the literature: continuous scaling (CS), group scaling (GS) and auto-scaling (AS) [6]. According to these references, GS is selected for this work because this method considers changes between sensors and these sensors are not processed independently. Using this normalized data, a PCA model is built by each actuator phase.

In a second step, the experiments are performed by evaluating the structure in the different possible states or scenarios (undamaged and with different kind of damages) under different temperatures [7]. The collected signals are pre-processed and organized in the same manner as in the first step. Afterwards, these signals are projected onto the corresponding PCA model, and the scores and the Q-index are obtained. With a predetermined number of scores and with the SPE-index, combining the results of all the actuator phases, a feature vector is built. This vector is used to apply data fusion and obtain a pattern with the information for the classification using all the structural states as shown in Figure 1. This feature vector is then introduced as the input to a classifier. A SOM has been chosen as classifier because its characteristics can provide a good support for the classification and graphical representation and grouping input data with similar features in clusters [5]. One important characteristic of this kind of ANN is that it does not need previous knowledge about the state of the structure (healthy or with some damage) to obtain the final clustering.

To visualize the results of the classification the U-matrix surface and the cluster map are used. The U-matrix surface allows the visualization of the distances between neurons by means of colours between adjacent neurons and the cluster map corresponds to another representation that can be used as a tool to show the different data set grouped with similar characteristics showing the clustering tendency. However, it is not possible to provide a multi-damage detection which is able to identify several occurring damages independently. Multi-damage scenarios will just be detected as an additional damage and generate an additional cluster into the SOM.

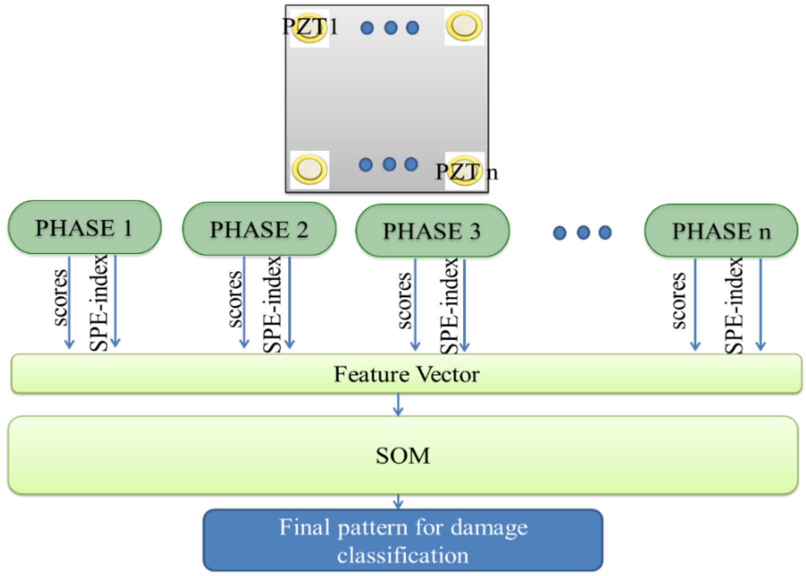
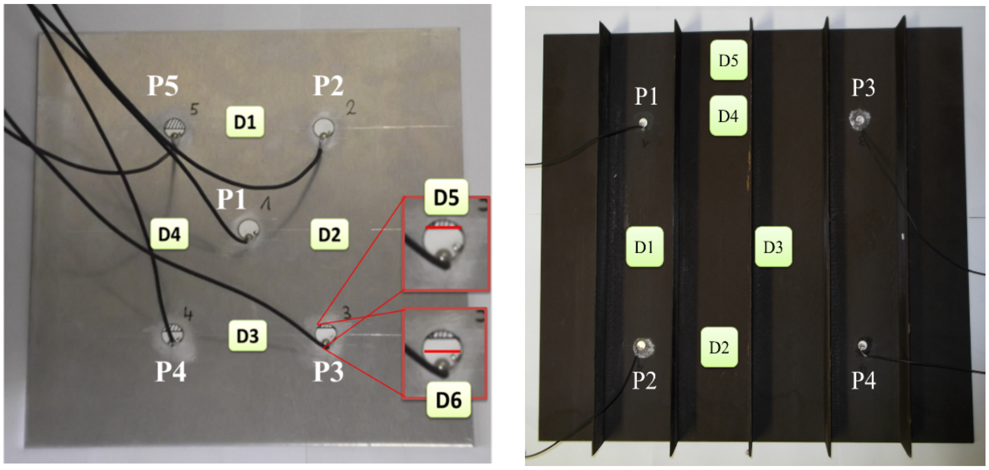


Figure 1. Final pattern for damage classification.

EXPERIMENT SETUP AND RESULTS

The validation of the methodology is carried out by using data collected from experiments performed on two different specimens. The first specimen corresponds to an aluminium plate with dimensions 200 mm × 200 mm which is instrumented with 5 PZT transducers (PIC-151) bonded on the surface as shown in Figure 2a. Six damages have been simulated on the structure. The first four damages are simulated by placing magnets on both sides of the structure at different positions. The other two damages correspond to a break or cut of the sensor patch to reduce the total area; 25% for damage 5 and, 50% for damage 6. The location of the damages is shown in Figure 2a. The second structure is a simplified aircraft composite skin panel made of carbon fibre reinforced plastic (CFRP).The structure is depicted in Figure 2b. The overall size of the plate is approximately 500×500×1.9mm and its weight is about 1.125 kg. The stringers are 36 mm high and 2.5 mm thick. The properties of the unidirectional (UD) material are VFibre = 60%, E1=142.6GPa, E2=9.65GPa, ν12 = 0.334, ν13 = 0.328, ν23 = 0.54 and G12 = G13 = 6.0GPa. The plate and the stringers consist of 9 plies. All plies are aligned in the same direction. Damage on the tested composite plate was simulated by localized masses at different positions as in the previous case. Figure 2b outlines the coordinates for the simulated damage on the composite skin panel.

Both structures were inspected with a 12V Hamming windowed cosine train signal with 5 cycles and 50 KHz as central frequency. The first structure was subjected to temperature changes. To perform these experiments, the structure was placed in an oven with controlled temperature. Data from the structure under six different temperatures (24ºC, 30ºC, 35ºC, 40ºC, 45ºC and 50ºC) for each structural state were collected and 100 experiments were saved for each state and for each temperature. In the second structure, 120 experiments were recorded per sensor-actuator configuration at five different temperatures (35ºC, 45ºC, 55ºC, 65ºC and 75ºC).



a. ***b.***

Figure 2: Aluminium plate and damage description.

The wavelet approximation coefficients are calculated by using the data of the healthy structure in different temperatures. With this approximation coefficients the baselines (PCA models) are built. To determine the number of principal components, an analysis of the retained variance is performed. From this analysis, the first ten principal components are selected and used to define the PCA model by each actuator phase. After that, new data from the structure to be diagnosed, in different states (healthy and damaged) are then collected and projected into the corresponding PCA model. These projections (scores) and the SPE-index are calculated by each actuation phase and used to define the feature vector and perform the data fusion.

In order to define the optimal set of parameters to configure the map such as the normalization, the shape of the cluster map and its size, several SOMs are trained and validated. As a result, normalization type histD is selected on this work to normalize the feature vector. From this analysis an hexagonal lattice with a flat sheet shape is also defined. Different shapes such as sheet, cylinder or toroid can be chosen. For ease, a flat sheet shape is considered here. Finally, a cluster size of 30 x 10 is obtained to train the SOM.

To test the methodology, two kind of test were performed. First, the classification of the different states using all temperatures in the aluminium plate and second the Classification using the Baseline with all temperatures and Damages at different temperatures and Positions. The results obtained are shown in the following subsections.

Damage detection and classification using all temperatures

This test is performed to the aluminium plate and in this case, the optimal set of parameters to configure the map was found to be as follows: 150 scores, normalization HistD, a hexagonal lattice with a flat sheet shape and a cluster size of 30 x 30. In spite of the number of scores is high, the feature vector built with this number of scores and the SPE-index corresponds to a reduced version compared with the use of the raw signals or all the approximation coefficients. The results are presented in the Figures 3 and 4 by means of the cluster map and the U-matrix surface. Results show a clear distinction between the different structural states in both figures. This distinction can be observed by the different data sets with different colours in the cluster map. The main difference with the previous results is that the damage 3 is now separated in two groups. This result can also be confirmed by evaluating the U-matrix surface. It is also worth remarking the presence of small zones inside of each damage case in the U-matrix surface and the cluster map. In the U-matrix surface, the damage cases are separated by the highest boundaries, the sub-groups are represented by the dark blue colour and the separation between these sub-groups is represented by the light blue colour. More precisely, six sub-cases can be identified for each damage case. These six cases correspond to the data at the six temperatures. These results confirm that the methodology allows a proper identification of the different type of damage despite the changes of the temperature.

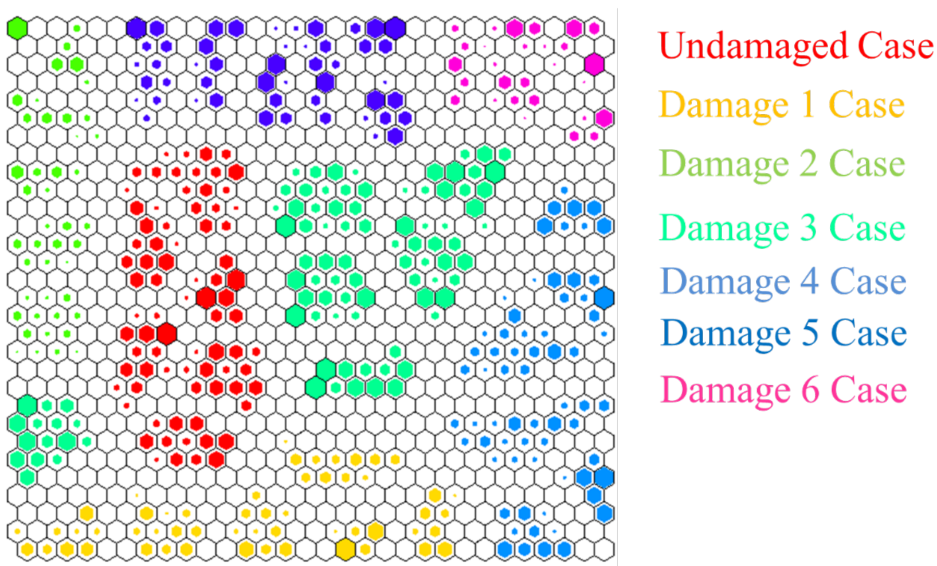


Figure 3: Classification of the different states using all temperatures, 150 scores, SPE-index, normalization type histD and map size 30 x 30 in the cluster map.

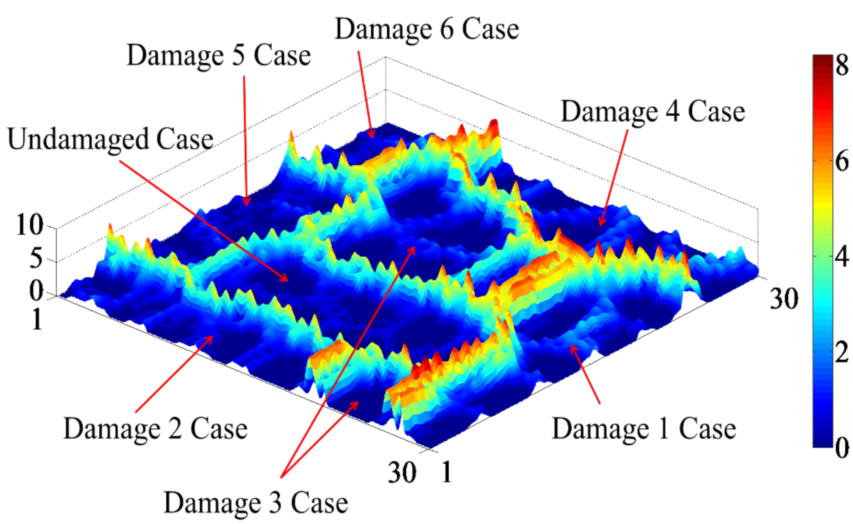


Figure 4. Classification of the different states using all temperatures, 150 scores, SPE-index, normalization type histD and map size 30 x 30 in the U-matrix.

Classification using the baseline with all temperatures and damages at different temperatures and positions

As second study, the stiffened composite panel is used, in this case, the baseline is built using data from all temperatures. Moreover, 5 different damages from different temperatures are used. In this respect, damage 1 corresponds to damage 1 when the plate is exposed to a temperature of 35º C; damage 2 is the damage 2 at 45º; damage 3 is the damage 3 at 55º; damage 4 is the damage 4 at 65º; and damage 5 is the damage 5 at 75º. The feature vector is formed by 30 scores and the SPE-index of each actuation phase. Results from the second study are depicted in Figures 5 and 6. In this case. The observation of the cluster map and the U-matrix surface allows us to identify the different states despite the presence of two outliers (in the undamaged case and in the damage 1). In contrast to the previous results, the boundary in the undamaged state is less clear with respect to the rest of boundaries in the U-matrix surface.

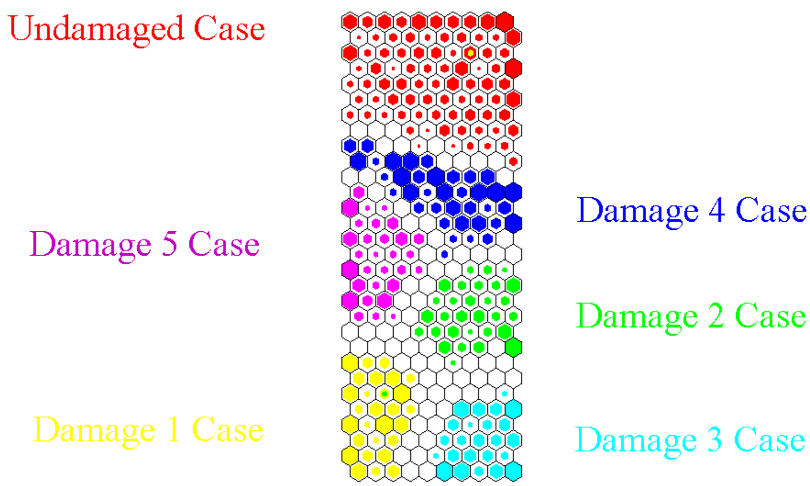


Figure 5. Classification using the Baseline with all temperatures and Damages at different temperatures and Positions.

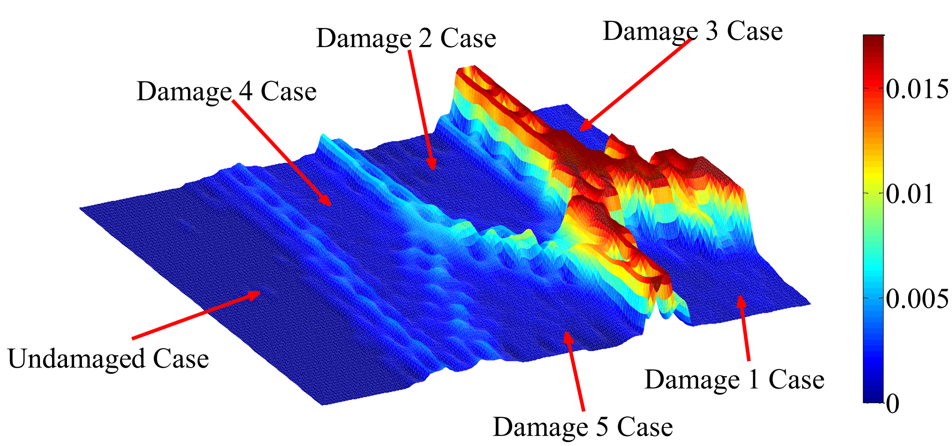


Figure 6. Classification using the baseline with all temperatures and damages at different temperatures and positions.

CONCLUSIONS

A new methodology for the analysis of ultrasonic signals in detection and classification of damages under temperature variations it was proposed. The results obtained showed that there is an influence of the temperature in the variability of the dynamics in the data gathered from the structure when it is subjected to environmental changes in spite of the structural state studied. This result demonstrates that the temperature is an important environmental effect to bear in mind in the design of a SHM system, however the methodology allowed in all the cases the damage detection in spite of temperature changes, the complexity and type of material of the evaluated structures. According to the results, in all the evaluated cases there was a clear separation between the healthy state and the damage states in the cluster map and the U-matrix surface. Finally, it is necessary to highlight that with the application of this approach, the problem of evaluate all the phases to define the existence of damage, especially, in large structures instrumented with several PZT transducers is solved. Now, the solution implies only the evaluation of the cluster map or the U-matrix surface obtained by data fusion.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Professor C.-P. Fritzen and the Group of Technical Mechanics, Institute of Mechanics for their support during the experimental trials. Also to the “Unidad de Investigaciones” from the “Universidad Santo Tomás-Colombia” for sponsoring the research presented herein through the projects FODEIN 047 and FODEIN 8410083003-036. This work is also supported by CICYT (Spanish Ministry of Economy and Competitiveness) through grants DPI2011- 28033-C03-01 and DPI2014-58427-C2-1-R.

REFERENCES

[1] Rytter, A. 1993. “Vibration based inspection of Civil Engineering Structures”. PhD Thesis. Department of Building Technology and Structural Engineering. Aalborg University, Denmark.

[2] Clarke, T., F. Simonetti and P. Cawley. 2010. “Guided wave health monitoring of complex structures by sparse array systems: Influence of temperature changes on performance,” *Journal of Sound and Vibration,* 329(12):2306-2322.

[3] Dodson, J.C. and D.J. Inman. 2013. “Thermal sensitivity of Lamb waves for structural health monitoring applications,” *Ultrasonics,* 53(3): 677-685.

[4] Lu, Y. and J.E. Michaels. 2005. “A methodology for structural health monitoring with diffuse ultrasonic waves in the presence of temperature variations,” *Ultrasonic*, 43(9):717-731.

[5] Tibaduiza, D.A., L.E. Mujica, and J. Rodellar. 2012. “Damage classification in structural health monitoring using principal component analysis and self-organizing maps,” *Structural Control and Health Monitoring*, 20(10):1303-1316.

[6] Westerhuis, J., T. Kourti, and J. MacGregor. 1999. “Comparing alternative approaches for multivariate statistical analysis of batch process data,” *Journal of Chemometrics*, 13(3-4):397–413.

[7] Anaya, M., D.A. Tibaduiza, M.A. Torres, F. Pozo, M. Ruiz, L.E.Mujica, J. Rodellar, C.P. Fritzen. 2014. “Data-driven methodology to detect and classify structural changes under temperature variations,” *Smart Materials and Structures*, 23(4), paper 045006.