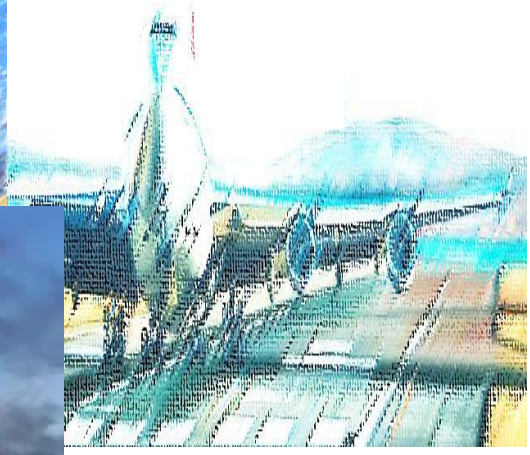
DESIGN AND VALIDATION OF A STRUCTURAL HEALTH MONITORING SYSTEM BASED ON BIO-INSPIRED ALGORITHMS



Maribel Anaya Vejar

Barcelona-Spain, June 2016

DOCTORAL THESIS

DESIGN AND VALIDATION OF A STRUCTURAL HEALTH MONITORING SYSTEM BASED ON BIO-INSPIRED ALGORITHMS

Maribel Anaya Vejar

CoDAlab
Control, Dynamics and Applications

UNIVERSITAT POLITÈCNICA DE CATALUNYA
DEPARTAMENT DE MATEMÀTIQUES

ABSTRACT

DESIGN AND VALIDATION OF A STRUCTURAL HEALTH MONITORING SYSTEM BASED ON BIO-INSPIRED ALGORITHMS

by Maribel Anaya Vejar

ADVISORS:
Dr. Francesc Pozo Montero
Dr. Diego Alexander Tibaduiza Burgos

May, 2016
Barcelona, Spain

The need of ensure the proper performance of the structures in service has made of structural health monitoring (SHM) a priority research area. Researchers all around the world have focused efforts on the development of new ways to continuous monitoring the structures and analyze the data collected from the inspection process in order to provide information about the current state and avoid possible catastrophes. To perform an effective analysis of the data, the development of methodologies is crucial in order to assess the structures with a low computational cost and with a high reliability. These desirable features can be found in biological systems, and these can be emulated by means of computational systems. The use of bio-inspired algorithms is a recent approach that has demonstrated its effectiveness in data analysis in different areas. Since these algorithms are based in the emulation of biological systems that have demonstrated its effectiveness for several generations, it is possible to mimic the evolution process and its adaptability characteristics by using computational algorithms. Specially in pattern recognition, several algorithms have shown good performance. Some widely used examples are the neural networks, the fuzzy systems and the genetic algorithms. This thesis is concerned about the development of bio-inspired methodologies for structural damage detection and classification. This document is organized in five chapters. First, an overview of the problem statement, the objectives, general results, a brief theoretical background and the description of the different experimental setups are included in Chapter 1 (Introduction). Chapters 2 to 4 include the journal papers published by the author of this thesis. The discussion of the results, some conclusions and the future work can be found on Chapter 5. Finally, Appendix A includes

other contributions such as a book chapter and some conference papers.

UNIVERSITAT POLITÈCNICA DE CATALUNYA
DEPARTAMENT DE MATEMÀTIQUES

RESUMEN

DISEÑO Y VALIDACIÓN DE UN SISTEMA DE MONITORIZACIÓN DE DAÑOS EN ESTRUCTURAS BASADO EN SISTEMAS BIO-INSPIRADOS

por Maribel Anaya Vejar

DIRECTORES:
Dr. Francesc Pozo Montero
Dr. Diego Alexander Tibaduiza Burgos

Mayo, 2016
Barcelona, España

La necesidad de asegurar el correcto funcionamiento de las estructuras en servicio ha hecho de la monitorización de la integridad estructural un área de gran interés. Investigadores en todas las partes del mundo centran sus esfuerzos en el desarrollo de nuevas formas de monitorización continua de estructuras que permitan analizar e interpretar los datos recogidos durante el proceso de inspección con el objetivo de proveer información sobre el estado actual de la estructura y evitar posibles catástrofes. Para desarrollar un análisis efectivo de los datos, es necesario el desarrollo de metodologías para inspeccionar la estructura con un bajo coste computacional y alta fiabilidad. Estas características deseadas pueden ser encontradas en los sistemas biológicos y pueden ser emuladas mediante herramientas computacionales. El uso de algoritmos bio-inspirados es una reciente técnica que ha demostrado su efectividad en el análisis de datos en diferentes áreas. Dado que estos algoritmos se basan en la emulación de sistemas biológicos que han demostrado su efectividad a lo largo de muchas generaciones, es posible imitar el proceso de evolución y sus características de adaptabilidad al medio usando algoritmos computacionales. Esto es así, especialmente, en reconocimiento de patrones, donde muchos de estos algoritmos brindan excelentes resultados. Algunos ejemplos ampliamente usados son las redes neuronales, los sistemas *fuzzy* y los algoritmos genéticos. Esta tesis involucra el desarrollo de unas metodologías bio-inspiradas para la detección y clasificación de daños estructurales. El documento está organizado en cinco capítulos. En primer lugar, se incluye una descripción general del problema, los objetivos del trabajo, los resultados obtenidos, un breve marco conceptual y la descripción de los diferentes escenarios experimentales en el Capítulo

1 (Introducción). Los Capítulos 2 a 4 incluyen los artículos publicados en diferentes revistas indexadas. La revisión de los resultados, conclusiones y el trabajo futuro se encuentra en el Capítulo 5. Finalmente, el Anexo A incluye otras contribuciones tales como un capítulo de libro y algunos trabajos publicados en conferencias.

ACKNOWLEDGMENTS

The present thesis is an effort of many people who has supported my work over these years who have led me to further progress in my professional life, by this reason, I would hereby like to express my profound sense of gratitude to my supervisors: Francesc Pozo Montero and Diego Alexander Tibaduiza Burgos, for their many suggestions and their constant support during this research, without their valuable advice and guidance this thesis would not have been possible. Thank you so much for always being there when I needed you.
I'd also like to extend my gratitude to the Professor C-P. Fritzen and to my colleague and friend Miguel Angel Torres for their collaboration, suggestions and hospitality during my research visit in the Siegen University. In the same way to the professor Alfredo Güemes for his kind help and co-operation from "Universidad Politécnica de Madrid"

To all my friends, family and people who have believed in me and supported me all these years, among them, to all the members of CoDAlab research group for their friendship, specially to professor José Rodellar, Luis Eduardo Mújica and Magda Ruiz who gave me valuable guidance and important suggestions in the first stage of my work.

In addition, I would like to thank the support from the " Ministerio de Educación, Cultura y Deporte " in Spain for the support in the research visit through the scholarship: "Movilidad de estudiantes en programas de doctorado con Mención hacia la Excelencia" and to the "Universidad Santo Tomás" for the support during the last year of my doctoral thesis.

To God, the greatest scientific who gave us endless things to discover, learn and investigate. Finally, I'd like to give my special thank to my parents, my sister and my brother who encouraged me to finish this thesis. Last, but not least, Id like to express my deep gratitude to my husband and best friend Diego, to my son Juan Diego for their constantly encouragement when I felt frustrated with problems in this thesis, they wouldn't let me quit.

Maribel Anaya Vejar

Contents

Contents **vii**

1 Introduction **1**
1.1 Main contribution . . . 3
1.2 Objectives . . . 3
1.2.1 Specific Objectives . . . 3
1.3 General results . . . 3
1.4 Theoretical background . . . 5
1.4.1 Bioinspired systems . . . 6
1.4.2 Natural immune systems . . . 7
1.4.3 Artificial Immune System . . . 9
1.4.4 Discrete Wavelet Transform . . . 10
1.4.5 Multiway Principal Component Analysis (MPCA) . . . 11
1.4.6 Principal Component Analysis (PCA) . . . 11
1.4.7 Damage detection indices based on PCA . . . 15
1.4.8 Self-organizing maps (SOM) . . . 15
1.5 Cases studies . . . 16
1.5.1 Aircraft skin panel . . . 16
1.5.2 Simplified aircraft skin panel . . . 18
1.5.3 Aluminium plate with reversible damages and temperature variations . 18
1.5.4 Vertical axis wind generator ALEKO WGV75W . . . 19
1.6 Research framework . . . 20
1.7 Organization . . . 21

Bibliography **23**

2 Data-driven methodology for damage detection and classification under temperature changes **27**
2.1 Introduction . . . 29
2.2 Experimental setup . . . 32
2.3 Theoretical background . . . 33
2.3.1 Discrete Wavelet Transform . . . 34
2.3.2 Multiway Principal Component Analysis (MPCA) . . . 35
2.3.3 Principal Component Analysis . . . 36
2.3.4 Self-Organizing Map (SOM) . . . 37
2.4 Damage detection and classification . . . 38

2.5 Analysis and discussion of the results . . . 41
2.5.1 Aluminium plate . . . 41
2.5.2 Multi-layered composite plate . . . 50
2.6 Conclusions . . . 57
2.7 Acknowledgment . . . 57

Bibliography **59**

3 A bio-inspired methodology based on an artificial immune system for damage detection in structural health monitoring **63**
3.1 Introduction . . . 65
3.2 General framework . . . 68
3.2.1 Bio-inspired systems . . . 68
3.2.2 Natural immune systems . . . 68
3.2.3 Artificial immune systems . . . 71
3.2.4 Principal Component Analysis (PCA) . . . 72
3.2.5 Damage detection indices based on PCA . . . 75
3.3 Damage detection methodology . . . 75
3.3.1 Data pre-processing, PCA modeling and feature extraction . . . 76
3.3.2 Training step . . . 77
3.3.3 Testing step . . . 79
3.4 Experimental setup and experimental results . . . 80
3.4.1 Experimental setup . . . 80
3.4.2 Experimental results . . . 81
3.5 Concluding remarks . . . 87

Bibliography **89**

4 Detection and Classification of Structural Changes using Artificial Immune Systems and Fuzzy Clustering **93**
4.1 Introduction . . . 95
4.2 General framework . . . 98
4.2.1 Bio-inspired systems . . . 98
4.2.2 Natural immune systems . . . 99
4.2.3 Artificial immune systems . . . 101
4.2.4 Principal Component Analysis (PCA) . . . 102
4.2.5 Damage detection indices based on PCA . . . 105
4.3 Damage detection methodology . . . 106
4.3.1 Data pre-processing, PCA modeling and feature extraction . . . 106
4.3.2 Training step . . . 107
4.3.3 Testing step . . . 110
4.3.4 Damage classification . . . 110
4.4 Experimental setup and experimental results . . . 112
4.4.1 Experimental setup . . . 112
4.4.2 Experimental results . . . 113
4.5 Concluding remarks . . . 119

Bibliography **121**

5 Conclusions and future research **125**
5.1 Comment and concluding remarks 125
5.1.1 Instrumentation and data acquisition 125
5.1.2 Data pre-processing 126
5.1.3 Structural damage detection and classification 126
5.1.4 Sensor fault detection and classification 127
5.2 Future research 128

A Publications **129**
A.1 Book chapters 129
A.2 Journals 129
A.3 Conferences 130

Chapter 1

Introduction

Structural health monitoring (SHM) is an important research area which seeks to assess the proper performance of a structure. To achieve this objective, SHM makes use of sensors permanently installed in the structure for inspecting and defining its current state based on the analysis of structural responses. As a result, the collected structural responses are analysed and compared with baseline patterns in order to detect abnormal characteristics and define the structural integrity. The obtained information can be used to define whether the structure can operate and under which conditions.

In general, the damage identification can be performed by two main approaches: the first consists in obtaining a reliable physics-based model of the structure, while the second is based on data-driven approaches which 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 signals collected in several experiments, it is necessary to ensure the proper functioning of the sensors, actuators and hardware used to inspect the structure. Among the big quantity of damages that can be presented in the normal service of a structure, the following categories can be distinguished [1, 2]:

- Gradual damage such as fatigue, corrosion and aging.
- Sudden and predictable damage like aircraft landing and planned explosions in confinement vessels.
- Sudden and unpredictable damage originating from foreign-object-impact, earthquakes and wind loads.

At the same time, these different kinds of damage can also be classified depending on its severity in three big groups [3]:

- *Light damage*. This corresponds to the initial stage of a damage, which can be relatively easily-repairable and is not dangerous for the normal operation of the structure.
- *Moderate damage*. With respect to the previous one, this damage requires major repairs and needs to be evaluated more carefully in order to define if the structure can operate in normal conditions.

- *Severe damage*. This type of damage, unlike previous damages, requires big reparations or the replacement of the structure.

The variability of the dynamic properties of an in-service structure, can result in changing the environmental and operational conditions. [4, 5]. Some of the environmental conditions to consider are humidity, wind loads, temperature, pressure, among others. Operational conditions include loading conditions, operational speed and mass loading [4]. The damage identification techniques need, in this way, to consider several variables and in most of the cases the damage identification procedure depends on the critical damage admissible in the structure.

The damage diagnosis is often grouped in four levels [6], starting by the damage detection. In this level the objective is to know whether there are some changes in the structure and if these changes are due to a damage. The second level considers the damage localization. The third level is used to define the type of damage and its size. The last level is defined for calculating the remaining lifetime of the structure. Recently, an extra level is considered which includes intelligent structures with self-healing [7] (Figure 1.1).

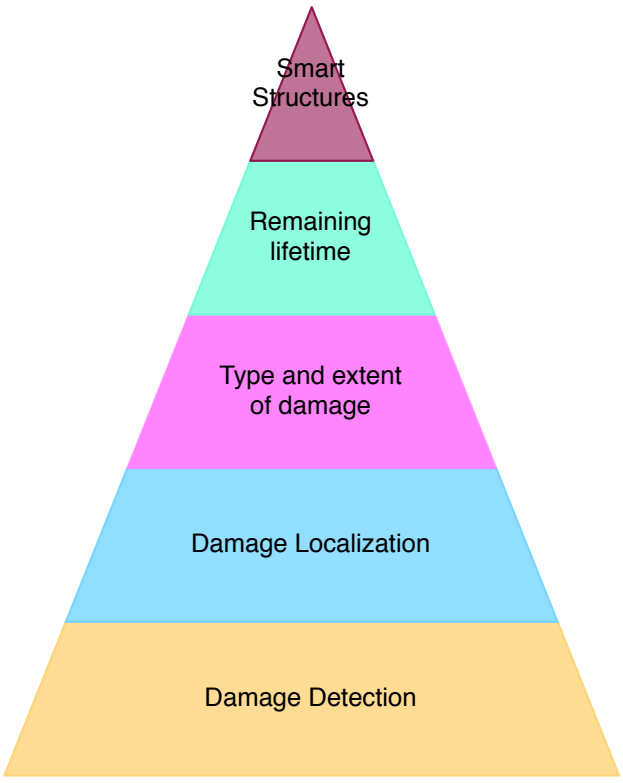


Figure 1.1: Levels of damage diagnosis in structural health monitoring [6]

In this thesis, the damage detection and classification are tackled by using bio-inspired data-driven approaches. These bio-inspired approaches are applied to analyze the signals collected by several experiments from a piezoelectric system which is permanently attached to the structures.

1.1 Main contribution

As a contribution to solve the problem of damage detection and classification in structural health monitoring, this thesis presents the development of two bio-inspired methodologies for damage detection and classification by using signals from a piezoelectric system which is permanently attached to the structure under test. These methodologies consider these tasks as a pattern recognition problem by comparing the data collected from the structures without damages and the structure under different structural states (with and without damages) using bio-inspired data-driven algorithms.

The developed methodologies have been tested with different experimental setups at real and small-scale structures and the results show promising developments for implementing in continuous monitoring tasks. All the results are included and discussed in each section and in the conclusions.

1.2 Objectives

The main objective of this thesis is to develop a methodology for the detection and classification of damages in mechanical structures using bio-inspired systems. This damage identification methodology will work under the paradigm that any damage in the structure will produce changes in the vibrational responses and these damages can be quantified by pattern recognition techniques.

1.2.1 Specific Objectives

1. To study the problem of monitoring and damage detection in structures.
2. To study the application of bio-inspired algorithms for the damage identification as a tool for pattern recognition.
3. Design and validate a damage detection and classification methodology based on bio-inspired systems.
4. Design and validate a methodology for damage detection and classification under temperature changes.
5. To validate and adapt the proposed methodologies using experiments with small-scale structures or components.
6. To validate the methodologies using aircraft and aerospace structures in real-scale.

1.3 General results

All the proposed objectives were achieved and presented within the framework of this thesis. However some brief comments about each result associated to each objective are summarized below.

To study the problem of monitoring and damage detection in structures.

The damage detection problem and the need for continuous monitoring have been studied along this thesis. In particular, the research stage and the international conferences where all these works have been presented served as scenario to understand the need for the application and development of structural health monitoring systems. As a result of this study, a state-of-the-art with the more relevant works on bio-inspired systems is presented in each chapter of this thesis.

To study the application of bio-inspired algorithms for the damage identification as a tool for pattern recognition.

The bio-inspired algorithms are built mimicking some biological systems functions such as processing of information, decision-making and others. In this thesis the artificial immune system (AIS), fuzzy clustering and self organized maps (SOM) are studied in order to know and learn their principles and performance in the nature. As it will be shown in the following chapters, these algorithms require to build the feature space to apply the algorithm.

Design and validate a damage detection and classification methodology based on bio-inspired systems.

The damage detection and classification problems were addressed by means of three methodologies, the self organized maps (SOM), fuzzy clustering and artificial immune systems (AIS). The first one is used in a methodology to detect and classify structural changes under temperature variations. This methodology was evaluated using an aluminium plate and a simplified aircraft skin panel. In both structures, the damages were detected and classified. The second and third methodologies were used to detect and classify damages in an aircraft skin panel. The published results using the bio-inspired algorithms are listed below:

1. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Structural damage assessment using an artificial immune system. In: Emerging Design Solutions in Structural Health Monitoring Systems, IGI-Global, 2015, doi: 10.4018/978-1-4666-8490-4.ch005.

2. **M. Anaya**, D.A. Tibaduiza, M.A. Torres-Arredondo, F. Pozo, M. Ruiz, L.E. Mujica, J. Rodellar and C.P. Fritzen. Data-driven methodology to detect and classify structural changes under temperature variations. Smart Materials and Structures, 23(4):045006, 2014, doi:10.1088/0964-1726/23/4/045006.

3. **M. Anaya**, D.A. Tibaduiza, F.Pozo. Detection and classification of structural changes using artificial immune systems and fuzzy clustering. International Journal on Bio-Inspired Computation, to appear.

4. **M. Anaya**, D.A. Tibaduiza, F.Pozo. A bioinspired methodology based on an artificial immune system for damage detection in structural health monitoring. Shock and Vibration (Special issue on Structural Dynamical Monitoring and Fault Diagnosis), 2015, article ID 648097, 15 pages, 2015, doi:10.1155/2015/648097.

5. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Data driven methodology based on artificial immune systems for damage detection. 7th European Workshop on Structural Health Monitoring. Nantes (France), July 8-11, 2014.

6. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Structural damage classification using artificial immune system and fuzzy clustering. 6th World Conference on Structural Control and Monitoring, Barcelona (Spain), July 15-18, 2014.

7. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Artificial immune systems for damage detection. 6th World Conference on Structural Control and Monitoring, Barcelona (Spain), July 15-18, 2014.

Design and validate a methodology for damage detection and classification under temperature changes.

A methodology for the detection and classification of damages under temperature variations was developed. The proposed scheme is partially based on a previous methodology developed by [3]. This methodology uses self-organizing maps (SOM) and a special unfolding to consider temperature variations and to produce most robust baselines. The results have been published in the following journal paper:

1. **M. Anaya**, D.A. Tibaduiza, M.A. Torres-Arredondo, F. Pozo, M. Ruiz, L.E. Mujica, J. Rodellar and C.P. Fritzen. Data-driven methodology to detect and classify structural changes under temperature variations. Smart Materials and Structures, 23(4):045006, 2014, doi:10.1088/0964-1726/23/4/045006.

2. **M. Anaya**, D.A. Tibaduiza, M.A. Torres-Arredondo, F. Pozo. Principal component analysis and self-organizing maps for damage detection and classification under temperature variations. 10th International Workshop on Structural Health Monitoring. Stanford University (Palo Alto), California, USA, 1-3 September, 2015.

Validate and adapt the proposed methodologies using experiments with small-scale structures or components.

Two small-scale structures were used to validate the proposed methodologies, the description of all of them is included in Chapter 4. These two structures correspond to a simplified aircraft skin panel made of carbon fibe reinforced plastic (CFRP) and an aluminium plate structure from the University of Siegen (Siegen, Germany).

Validate the methodologies using aircraft and aerospace structures in real-scale.

A real-scale structure was used to validate the methodologies. This structure corresponds to an aircraft skin panel located at the Universidad Politécnica de Madrid (Madrid, Spain). More details about this structure can be found in Chapter 4.

1.4 Theoretical background

The damage detection and classification methodologies presented in this thesis are based on data driven analysis. This means that the damage identification is developed by analyzing the data collected by several experiments from the structures under test. To perform this analysis, bio-inspired methodologies that are based on features extraction for pattern recognition have

been developed. More precisely, the developed methodologies apply several methods such as discrete wavelet transform (DWT), multiway principal component analysis (MPCA), damage indices, self-organizing maps (SOM) and artificial immune systems (AIS). In this sense, the key concepts that are used within the proposed methodologies are presented in the subsequent sections.

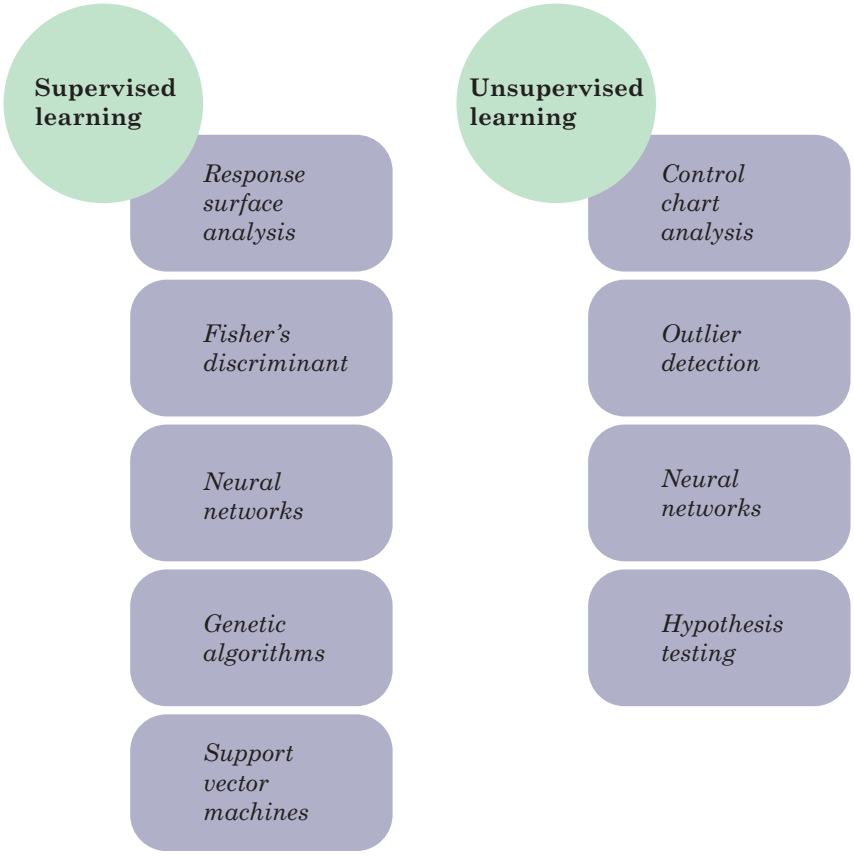


Figure 1.2: A state-of-the-art classification of the statistical models developed to enhance the SHM process

1.4.1 Bioinspired systems

The adaptation of the different living beings of the planet in harsh environments and the development of skills to solve the inherent problems in the interaction with the world of nature has resulted in the evolution of the species in order to survive and avoid their extinction. Some examples are the communication abilities, the reasoning, the physical structures design or the response of the body to external agents, among others [8].

Taking advantage of the fact that nature provides robust and efficient solutions to many different problems, more and more researchers on different areas work in the development of biologically inspired hardware and algorithms. The inspiration process is called *biomimetic* or *bioinspired* and aims to apply the developments in the field of biology to the engineering developments [9].

1.4.2 Natural immune systems

The human immune system is a complex and robust defense mechanism which is composed by a large network of specialized cells, tissues and organs. The system further includes an elevated number of sensors and a high processing capability. This system has demonstrated its effectiveness in the detection of foreign elements by protecting the organism against disease. The principal skills of the human immune system are:

- to distinguish between its own cells (self) and foreign cells (non-self).
- to recognize different invaders (called antigen) in order to ensure the protection of the body.
- to learn from specific antigen and adapt to them in order to improve the immune response to this kind of invader.

In general, when a foreign particle wants to access to the organism, it has to break several defense levels provided by the immune system that protects the organism, as shown in Figure 1.3. This levels are [10]:

- *External barriers*. These are the first and the major line of defense into the human body. These include elements such as the skin, the mucus secreted by the membranes, the tears, saliva and urine which present different physiological conditions that are harmful to the antigens as the temperature and the Ph level, among others. The response of these barriers is equal for any foreign invader [11].
- *Innate immune system*. This barrier refers to the defense mechanisms that are activated immediately or within a short lapse of time of an antigen's arrival in the body. The innate immune system operates when the first barrier has been broken. This system, in opposition to the adaptive immune system, is not adaptive [10].
- *Adaptive immune system*. This is the last defense level and reacts to the stimulus of foreign cells or antigens that evade both the external barriers and the innate immune defense [10]. Adaptive immunity creates some sort of memory that leads to an improved response to future encounters with this antigen.

With respect to different type of cells, the immune system includes cells born in the bone marrow that are usually called *white blood cells*, *leukocytes* or *leucocytes* [12]. Among the white blood cells, it is possible to highlight the T-cells and the B-cells. On one hand, the T-cells are so called since their maturation takes place in the thymus.Besides, this kind of cells have high mobility and can also be found in the blood and the lymph [13]. One can distinguish three types of T-cells:

- the T-helper cells, involved in the activation of B-cells;
- the T-killer cells that destroy the invaders; and finally
- the T-suppressor cells that avoid the allergic reactions [14].

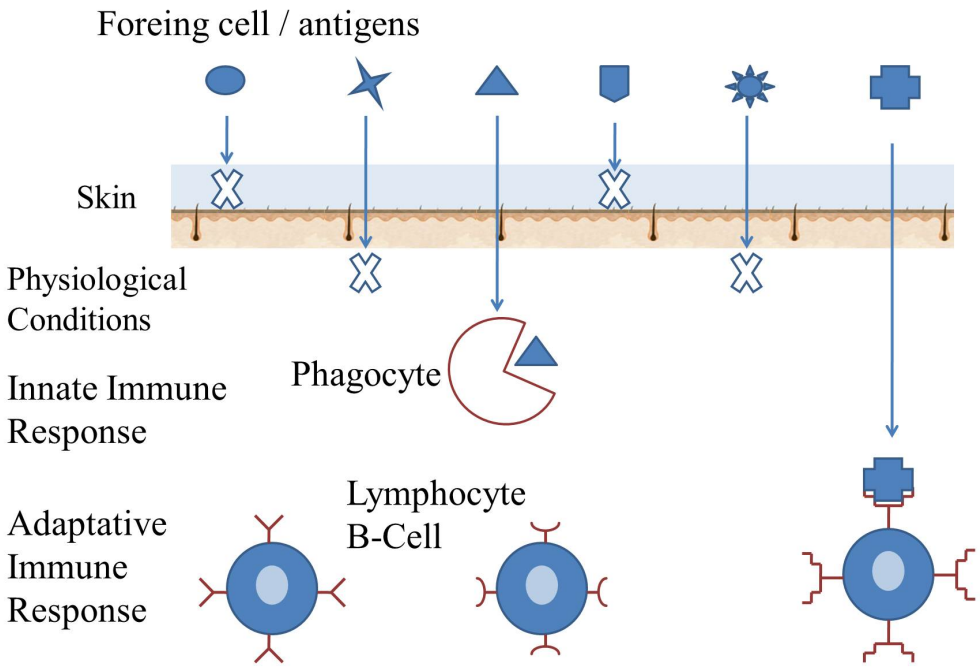


Figure 1.3: System Defense Barriers

On the other hand, the B-cells produce and secrete a special protein called antibody, which recognizes and binds the antigen. The responsibility of each B-cell is the production of a specific antibody. This protein is then used for signaling other cells what elements have to be removed from the body [13]. When the antigen passes over the first barrier of the immune system, the HIS performs the following steps to eliminate the invader [13]:

1. The specialized cells of the immune system, called *antigen presenting cells* (APCs) (macrophages, for instance). These cells active the immune response by ingesting the antigen and dividing it into simple substances known as antigenic peptides.

2. These peptides are joined to the molecules called *major histocompatibility complex* (MHC), inside of the macrophage, and the result passes to the immune cell surface.

3. The T-cells have receptor molecules able to identify and recognize different combinations of MHC-peptide. When the receptor molecule recognizes the combination, the T-cell is activated and sends a chemical signal to other immune cells.

4. The B-cells are activated by chemical signals and they initiate the recognition of the antigen in the bloodstream. This process is performed by the receptor molecules in the B-cells.

5. The mission of the B-cells –when they are activated– is to secrete antibodies to bind the antigens they find, and to neutralize and eliminate them from the body.

The T- and B-cells that have recognized the antigen proliferate and, some of them, become memory cells. These memory cells remain in the immune system to eliminate the same antigen –in the future– in a more effective manner [8, 13].

Three immunological principles are used in artificial immune systems [8, 13, 15]:

- *Immune network theory*. This theory was first introduced by Niels Jerne in 1974 and describes how the immune memory is built by means of the dynamic behaviour of the immune system cells. These cells can recognize by themselves, detect invaders, as well as interconnect between them to stabilize the network [10].

- *The negative selection*. The negative selection is a process that allows the identification and eradication of the cells that react to the own body cells. This ensures a convenient operation of the immune system since it is able to distinguish between foreign molecules and self-molecules thus avoiding autoimmune diseases. This process is similar to the maturation of T-cells carried out in the thymus [8].

- *The clonal selection*. This is a mechanism of the adaptive immune responses in which the cells of the system are adapted to identify an invader element [13]. Antibodies that are able to recognize or identify an antigen can proliferate. Those antibodies unable to recognize the antigens are eliminated. The new cells are clones of their parents and they are subjected to an adaptation process by mutation. From the new antibody set, the cells with the greatest affinity with respect to the primary antigen are selected as memory cells therefore excluding the rest.

1.4.3 Artificial Immune System

Artificial immune systems (AIS) are an adaptive and bio-inspired computational systems based on the processes and performance of the human immune system (HIS) and its properties –diversity, error tolerance, dynamic learning, adaptation, distributed computation and self-monitoring– [16]. Nowadays, these computational systems are used in several research areas such as pattern recognition [17], optimization [13], computer security [18], among others [19]. Table 1.1 presents the analogy between the natural and artificial immune systems applied to the field of structural health monitoring.

In the implementation of an artificial immune system, it is fundamental to bear in mind two important aspects:

- To define the role of the antigen (ag) and the antibody (ab) in the context of the application. Both are represented or coded in the same way. This representation is generally given by a vector of binary or real numbers [14].

- To define the mechanism that measures the degree of correspondence between an antigen and an antibody. This measure is usually related to the distance between them [8]. If both an antigen and an antibody are represented by $L-$dimensional arrays,

$$ab = (ab_1, \ldots, ab_L) \in \mathbb{R}^L,$$
$$ag = (ag_1, \ldots, ag_L) \in \mathbb{R}^L,$$

  the distance d between them can be computed using, for instance, the Euclidean distance (related to the $2-$norm) or the so-called Manhattan distance (related to the $1-$norm) as

Table 1.1: Analogy between the biological immune system and artificial immune system [15].

Biological Immune System	**Artificial Immune System in SHM**
antibodies	a detector of a specific pattern
antigens	structural health or damage condition
matured antibodies	database or information system for damage detection
recognition of antigens	identification of health and damage condition
process of mutation	training procedure
immune memory	memory cells

in equations (4.1) and (4.2), respectively [12]:

$$d(ab, ag) = \|ab - ag\|_2 = \sqrt{\sum_{i=1}^{L}(ab_i - ag_i)^2} \tag{1.1}$$

$$d(ab, ag) = \|ab - ag\|_1 = \sum_{i=1}^{L}|ab_i - ag_i| \tag{1.2}$$

However, when both the antibody and the antigen are represented by a sequence of binary values, the distance between them is obtained by means of the Hamming distance in equation (1.3):

$$d(ab, ag) = \sum_{i=1}^{L}\delta_i,\ \delta_i = \begin{cases} 1, & ab_i = ag_i \\ 0, & ab_i \neq ag_i \end{cases} \tag{1.3}$$

Finally, there exists the adaptation process of the molecules in the artificial immune system. This adaptation allows to include the dynamic of the system, for instance, the antibodies excitation, cloning all the excited antibodies and the interconnection between them. All these elements are adapted from the three immunologic principles previously introduced.

1.4.4 Discrete Wavelet Transform

The discrete wavelet transform (DWT) is a powerful tool specially used in areas dealing with the analysis of transient signals. This transform provides sufficient information for both the analysis and synthesis of the original signal with a significant reduction in the computational time cost compared with the analysis of the wavelet in continuous time domain normally called *continuous wavelet transform* (CWT) [20]. Roughly speaking, the characteristics of the given function at a specified time and scale are represented by the transformation coefficients. In other words, each wavelet coefficient represents time and frequency information of the signal. The DWT analysis can be performed by means of a fast, pyramidal algorithm related to a two-channel subband coding scheme using a special class of filters called quadrature mirror filters as proposed by Mallat [21]. In this algorithm, the signal is analyzed at different frequency bands with different resolution by decomposing the signal into approximation and detail coefficients. This is achieved by successive high-pass (HP) and low-pass (LP) filtering of the input

signal. The optimum number of level decompositions could be determined, for example, based on a minimum-entropy decomposition algorithm [22]. Figure 1.4 shows the structure of the decomposition of the signal and the corresponding frequency bandwidths for the details (D_i) and approximations (A_i) [21]. Each level is obtained by the decomposition of the signal with high and low pass filtering and a down-sampling of the input signal. The approximations represent the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. The frequency $f_{\max}$ represents the maximum frequency contained in the recorded signal and n is the decomposition level. Discrete wavelets are often associated with dilation equations and scaling functions. The basic scaling function can be defined as:

$$\gamma(i) = \sum_{m=0}^{j} c_m \gamma(2i - m), \tag{1.4}$$

where the values c_m, $m = 0, \ldots, j$ are the wavelet coefficients. These coefficients must satisfy certain conditions that are discussed in [23]. The scaling functions are then used to construct the corresponding wavelets $\varphi(i)$ from the following formula:

$$\phi(i) = \sum_{m=0}^{j} (-1)^m c_m \gamma(2i + m - j + 1). \tag{1.5}$$

The approximation and detail coefficients can be calculated as follows:

$$A_{n,k} = 2^{-n/2} \sum_{i=1}^{j} x(i) \gamma(2^{-n} j - k), \tag{1.6}$$

$$D_{n,k} = 2^{-n/2} \sum_{i=1}^{j} x(i) \varphi(2^{-n} j - k), \tag{1.7}$$

where γ is called the scaling function, j is the number of discrete points of the input signal, n and k are considered to be the scaling (dilation) index and the translation index, respectively. Each value of n allows analyzing a different resolution level of the input signal. Scaling functions are similar to wavelet functions except that they have only positive values and are designed to smooth the input signal [21, 24].

1.4.5 Multiway Principal Component Analysis (MPCA)

MPCA is a very common practice in multivariate statistical procedures for monitoring the progress of batch processes [25, 26] and it is based in the fact that the data are inherently correlated [27]. This is similar to the use of PCA on a large two-dimensional matrix constructed by unfolding the three-way data matrix [28]. In this work the original data are expressed in a matrix whose dimensions are I experiments, K time instants and J sensors. This three-way data matrix is decomposed into a large two-dimensional matrix X as shown in Figure 1.5.

1.4.6 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a classical method used in applied multivariate statistical analysis with the goal of dimensionality reduction and, more precisely, feature extraction and

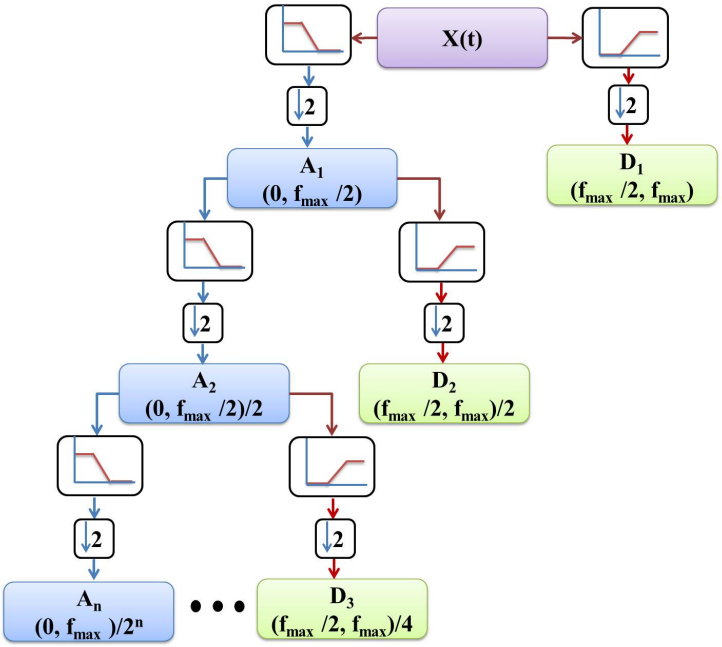


Figure 1.4: DWT Decomposition tree

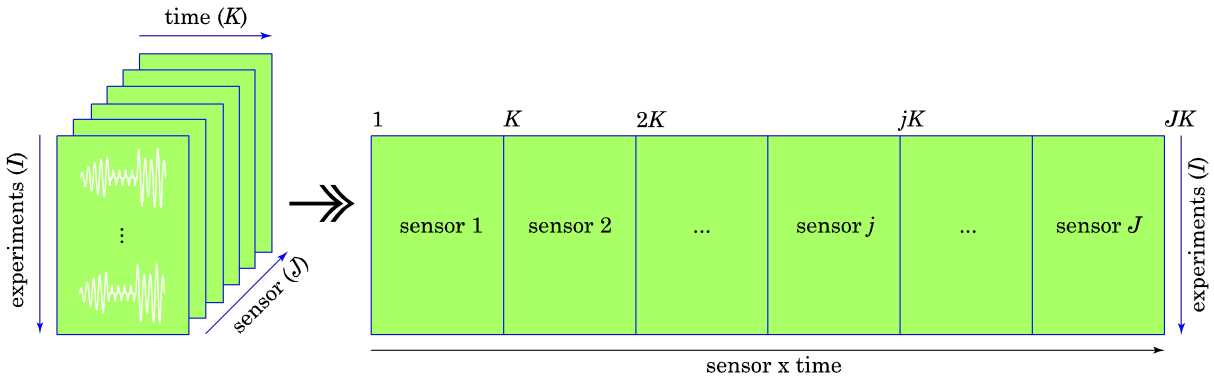


Figure 1.5: The collected data arranged in a three dimensional matrix is unfolded in a two dimensional matrix [3]

data reduction. It was developed by Karl Pearson in 1901 and integrated to the mathematical statistics in 1933 by Harold Hotelling [29]. The general idea in the use of PCA is to find a smaller set of variables with less redundancy [30]. To find these variables, the analysis includes the transformation of the current coordinate space to a new space to re-express the original data trying to filtering the noise and redundancies. These redundancies are measured by means of the correlation between the variables.

Matrix unfolding

The application of PCA starts –for each actuation phase– with the collected data arranged in a three dimensional matrix $n \times L \times N$. The matrix is subsequently unfolded –as illustrated in

Figure 1.5–in a two dimensional $n \times (N \cdot L)$ matrix as follows:

$$\mathbf{X} = \begin{pmatrix} x_{11}^1 & x_{12}^1 & \cdots & x_{1L}^1 & x_{11}^2 & \cdots & x_{1L}^2 & \cdots & x_{11}^N & \cdots & x_{1L}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1}^1 & x_{i2}^1 & \cdots & x_{iL}^1 & x_{i1}^2 & \cdots & x_{iL}^2 & \cdots & x_{i1}^N & \cdots & x_{iL}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1}^1 & x_{n2}^1 & \cdots & x_{nL}^1 & x_{n1}^2 & \cdots & x_{nL}^2 & \cdots & x_{n1}^N & \cdots & x_{nL}^N \end{pmatrix} \tag{1.8}$$

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

In other words, the objective is to find a linear transformation orthogonal matrix $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ that will be used to transform the original data matrix $\mathbf{X}$ according to the following matrix multiplication

$$\mathbf{T} = \mathbf{XP} \in \mathcal{M}_{n\times(N\cdot L)}(\mathbb{R}). \tag{1.9}$$

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 $\mathbf{PP}^{\mathrm{T}} = \mathbf{I}$, 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

$$\mathbf{X} = \mathbf{TP}^T.$$

Group scaling

Since the data in matrix $\mathbf{X}$ come from experimental trials and could have different magnitudes and scales, it is necessary to apply a preprocessing step to scale the data using the mean of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor.

More precisely, for $k = 1, 2, \ldots, N$ we define

$$\mu_j^k = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^k,\ j = 1,\ldots,L, \tag{1.10}$$

$$\mu^k = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} x_{ij}^k, \tag{1.11}$$

$$\sigma^k = \sqrt{\frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}(x_{ij}^k - \mu^k)^2}, \tag{1.12}$$

where μ_j^k is the mean of the n measures of sensor k at the time instant j; μ^k is the mean of all the measures of sensor k; and σ^k is the standard deviation of all the measures of sensor k.

Therefore, the elements x_{ij}^k of matrix $\mathbf{X}$ are scaled to define a new matrix $\check{\mathbf{X}}$ as

$$\begin{aligned} \check{x}_{ij}^k &:= \frac{x_{ij}^k - \mu_j^k}{\sigma^k}, \\ & i = 1, \ldots, n,\ j = 1, \ldots, L,\ k = 1, \ldots, N. \end{aligned} \tag{1.13}$$

When the data are normalized using equation (1.13), the scaling procedure is called *variable scaling* or *group scaling*.

For simplicity, and throughout the rest of the paper, the scaled matrix $\check{\mathbf{X}}$ is renamed as simply $\mathbf{X}$. The mean of each column vector in the scaled matrix $\mathbf{X}$ can be computed as

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} \check{x}_{ij}^k = \frac{1}{n}\sum_{i=1}^{n} \frac{x_{ij}^k - \mu_j^k}{\sigma^k} &= \frac{1}{n\sigma^k}\sum_{i=1}^{n} \left(x_{ij}^k - \mu_j^k\right) \\ &= \frac{1}{n\sigma^k}\left(\sum_{i=1}^{n} x_{ij}^k - n\mu_j^k\right) \\ &= \frac{1}{n\sigma^k}\left(n\mu_j^k - n\mu_j^k\right) = 0. \end{aligned}$$

Since the scaled matrix $\mathbf{X}$ is a *mean-centered* matrix, it is possible to calculate the covariance matrix as follows:

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R}). \tag{1.14}$$

The covariance matrix $\mathbf{C_X}$ is a $(N \cdot L) \times (N \cdot L)$ symmetric 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 follows:

$$\mathbf{C_X P} = \mathbf{P}\Lambda \tag{1.15}$$

where the columns of $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvectors of $\mathbf{C_X}$. The diagonal terms of matrix $\Lambda \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvalues $\lambda_i,\ i = 1, \ldots, N \cdot L$, of $\mathbf{C_X}$ whereas the off-diagonal terms are zero, that is,

$$\begin{aligned} &\Lambda_{ii} = \lambda_i,\ i = 1, \ldots, N \cdot L \\ &\Lambda_{ij} = 0,\ i, j = 1, \ldots, N \cdot L,\ i \neq j \end{aligned}$$

The eigenvectors $p_j,\ j = 1, \ldots, N \cdot L$, 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* or the *loading vectors* 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

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

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

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

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 $m-$dimensional space to get a data matrix $\hat{\mathbf{X}}$ as follows:

$$\hat{\mathbf{X}} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T \in \mathcal{M}_{n\times m}(\mathbb{R}). \tag{1.16}$$

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

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}, \tag{1.17}$$

or, equivalenty,

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T + \mathbf{E}. \tag{1.18}$$

The residual error matrix $\mathbf{E}$ describes the variability not represented by the data matrix $\hat{\mathbf{X}}$, and can be also expressed as

$$\mathbf{E} = \mathbf{X}(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T). \tag{1.19}$$

1.4.7 Damage detection indices based on PCA

Several damage detection indices based on PCA have been proposed and applied with excellent results in pattern recognition applications. In particular, two damage indices are commonly used: (i) the Q index (also known as SPE, *square prediction error*) and (ii) the Hotelling's T^2 index.

The Q index of the ith experimental trial x_i^T measures the magnitude of the vector $\tilde{x}_i^T := \tilde{\mathbf{X}}(i,:)$, that is, the events that are not explained by the model of principal components [31], and it is defined as follows:

$$Q_i = \tilde{\mathbf{X}}(i,:)\tilde{\mathbf{X}}(i,:)^T = x_i^T(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T)x_i. \tag{1.20}$$

The T^2 index of the ith experimental trial x_i^T is the weighted norm of the projected vector $\hat{t}_i^T := \hat{\mathbf{T}}(i,:) = x_i^T\hat{\mathbf{P}}$, that is, a measure of the variation of each sample within the PCA model and it is defined as follows:

$$T_i^2 = \sum_{j=1}^{\ell} \frac{\hat{t}_{i,j}^2}{\lambda_j} = \hat{t}_i^T\Lambda^{-1}\hat{t}_i = x_i^T(\hat{\mathbf{P}}\Lambda^{-1}\hat{\mathbf{P}}^T)x_i \tag{1.21}$$

1.4.8 Self-organizing maps (SOM)

The self-organizing map (SOM) is a kind of unsupervised neural network also known as Kohonen network [32] in honor to the professor Teuvo Kohonen. This neuronal network is specialized in visualization and analysis of high dimensional data and has the special property of

generating one organized map in the output based on the inputs, grouping input data with similar characteristics in clusters [33, 34]. To do that, the SOM internally organizes the data based on features and their abstractions from input data. In particular, these maps have been used in practical speech recognitions [35], robotics [36], control [37] and telecommunications [38], among others [39].

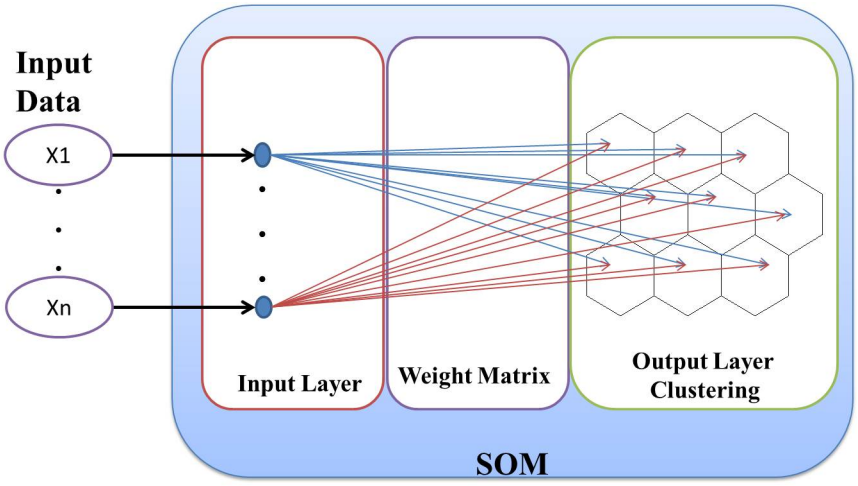


Figure 1.6: The self-organizing map (SOM)

The structure of the SOM is a single feed forward network and each node in the input layer is connected to all the output neurons [40]. The connection between the input layer and the output layer is defined by a weighting matrix (see Figure 1.6). In general, the inputs to the network are defined in the input layer and a weighting matrix relates each input with each cluster in the output layer.

The result of the SOM is an organised map which contains the final clustering of the input data. This information can be visualised by using the cluster map. The distance matrix techniques are widely used in the visualization of the cluster structure of the SOM. One of these techniques is the unified distance matrix, called U-matrix, which shows the separation between prototype vectors and neighbouring map units. The U-matrix is related with the linkage measure and can be visualized using a colour scale [40].

1.5 Cases studies

1.5.1 Aircraft skin panel

This structure corresponds to an aircraft skin panel and is one of the structures from the Universidad Politécnica de Madrid (Madrid, Spain). The skin panel is divided in small sections by means of stringers and ribs as it is shown in Figure 1.7. To test the proposed approaches, two sections of this structure were used. The dimensions of each section and the damage description are depicted in Figure 1.8. These sections were instrumented with 6 piezoelectric transducers

(PZT), two in the upper section, two in the lower section and two in the rib. The transducers dimensions are: diameter 26 mm and thickness 0.4 mm.

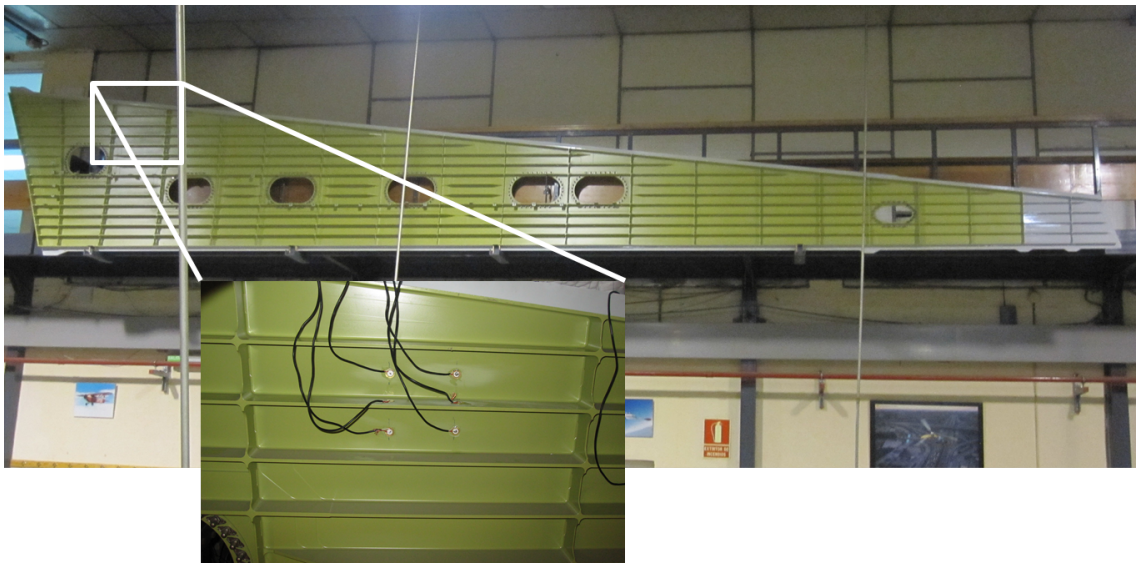
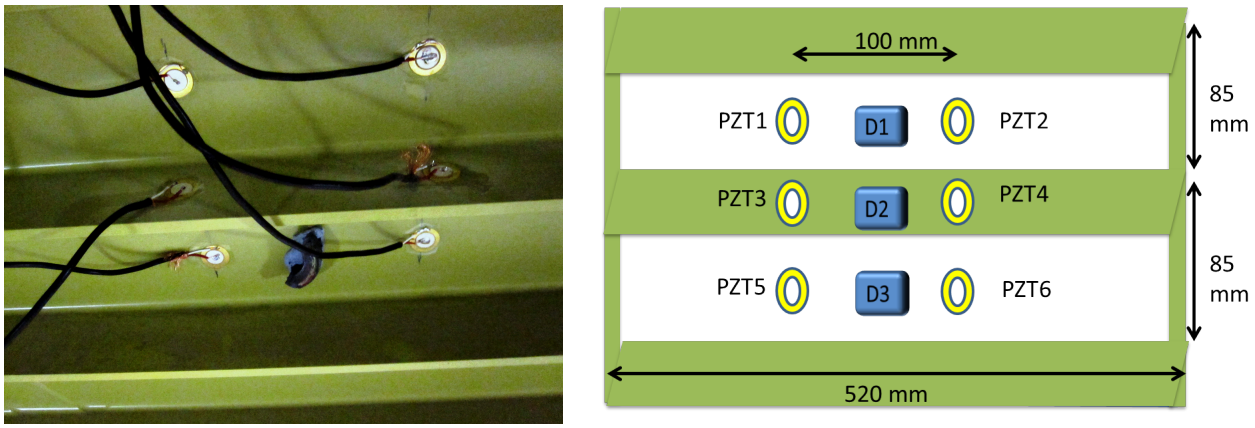


Figure 1.7: Sections tested with the PZT location.



(a) (b)

Figure 1.8: Damage description.

As excitation input, a BURST signal with 205 kHz as central frequency and nine peaks was used. Four different states of the structure were analyzed: the healthy structure and the structure with three different damages. Damages were simulated by adding a mass at three different positions (Figure 1.8), two of them on the skin and the other on the stringer. 100 experiments were performed and recorded: 25 with the undamaged structure and 25 more per damage. To ensure a good signal to noise ratio each signal was averaged 10 times. To apply and collect the signals to the PZT transducers, a chassis PXI 1033 from National Instruments ® was used. Due to the complexity and the size of this structure, a wideband power amplifier model 7602M of Krohn-Hite corporation is used to amplify the signal applied to the actuators.

1.5.2 Simplified aircraft skin panel

This structure corresponds to a simplified aircraft composite skin panel made of carbon fibre reinforced plastic (CFRP) located at the University of Siegen (Siegen, Germany). The structure is illustrated in Figure 1.9. The overall size of the plate is approximately 500 mm $\times$ 500 mm $\times$ 1.9 mm 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 $v_{FIBRE} = 60\%$, $E_1 = 142.6$GPa, $E_2 = 9.65$, $v_{12} = 0.334$, $v_{13} = 0.328$, $v_{23} = 0.54$ and $G_{12} = G_{13} = 6.0$GPa. The plate and the stringers consist of nine plies. All plies are aligned in the same direction. Damage on the tested composite plate was simulated by localized masses at different positions . To determine the carrier central frequency for the actuation signal in each structure, a frequency sweep was performed and the spectral content of each signal was analysed. The carrier frequencies were found to be 50 kHz. The carrier frequency was chosen to maximize the propagation efficiency. This type of excitation generates a dominant A mode that is propagated along the structure allowing a better interaction of the guided wave with the simulated damage.

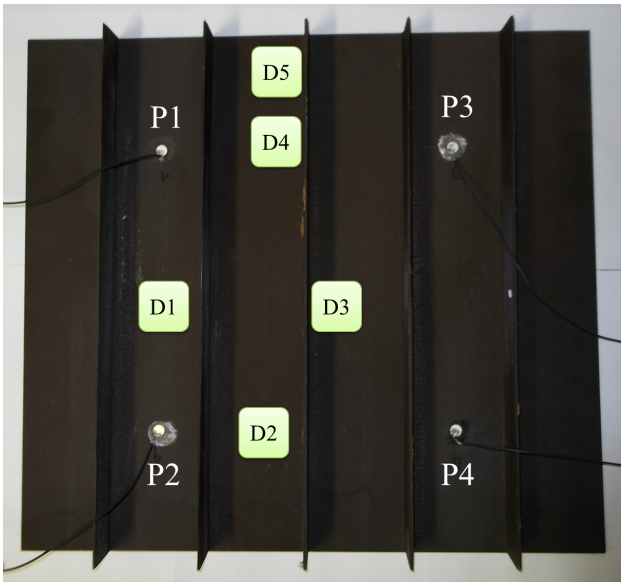
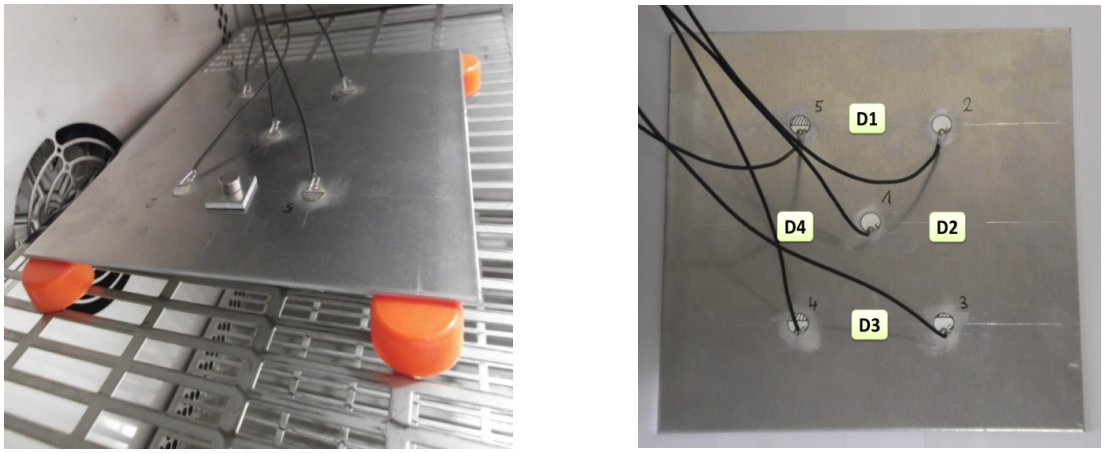


Figure 1.9: Simplified aircraft skin panel.

1.5.3 Aluminium plate with reversible damages and temperature variations

This structure corresponds to an aluminium plate with dimensions 200 mm $\times$ 200 mm instrumented with 5 PZT transducers (PIC-151). This structure was one of the structures tested at the University of Siegen during the research stage. Damages in the structure were simulated by adding masses at four different positions on the surface as it is shown in Figure 1.10.

To assess the structure, the Handyscopes HS3 and HS4 without pre-amplification were used. As excitation, a BURST signal of 50 kHz as central frequency was used and each collected experiment was averaged 100 times to ensure a good signal to noise ratio.



(a) (b)

Figure 1.10: Damages in the aluminium plate.

This structure was subjected to temperature changes. To perform these experiments, the structure was introduced in a oven with controlled temperature and data from the structure under six different temperatures (24, 30, 35, 40, 45 and 50 °C) for each structural state were collected. To avoid the contact between the metallic surfaces and eliminate the noise in the experiments, some plastic elements are used. For each temperature, 100 experiments were collected for each state.

1.5.4 Vertical axis wind generator ALEKO WGV75W

This experimental setup consists of a blade from a wind turbine which is instrumented with an active piezoelectric system that works in several actuation phases [41]. The wind turbine is a vertical axis wind generator *ALEKO WGV75W* which is designed to produce 50 W as nominal power and its maximum power is 75 W with 24 volts in direct current (see Figure 1.11). Some features for this wind turbine generator are: speed for nominal power: 10 meters per second; operation range: 3 m/s to 20 m/s; survival wind speed: 35 m/s, above this range damages can be appear, so for this reason SHM systems need to be ready to detect any change in the structure. The wind generator has a full weight of 10.5 kg, its rotor diameter is 560 mm, it has five blades made in aluminium alloy with a length of 745 mm. Figures 1.11 and 1.12 show the wind turbine and the blade.

The inspected blade was instrumented by means of 4 PZT transducers which are distributed over one face as it is shown in Figure 1.12. As excitation input, a *BURST* signal with 1 MHz as central frequency, 5 peaks and 8 volts of amplitude was used. Eight different states of the structure were analyzed: a healthy state and seven damages. Damages were simulated by adding a mass at seven different positions of the structure (see Figure 1.12). 40 experiments for healthy state and 20 experiments per damage were performed and collected. The acquisition modules are two *TiePie HS4* oscilloscopes and the arbitrary signal generator is a *TiePie HS5*.

Figure 1.11: Vertical axis wind generator ALEKO WGV75W

1.6 Research framework

The research group *Control, Dynamics and Applications (CoDAlab,* `codalab.ma3.upc.edu`*)* currently in the Department of Applied Mathematics at the Universitat Politècnica de Catalunya (UPC) has been working in structural health monitoring since 2009 through the projects *Intelligent aircraft structures: Development and validation of defect detection techniques based on pattern recognition* (2009-2011), *Intelligent structures: monitoring and damage detection systems with applications in aeronautics and offshore wind energy plants* (2011-2014, code number DPI2011-28033-C03-01) and *Development and validation of failure detection and localization systems and design of fault-tolerant control strategies with application in offshore wind energy plants (WinTurCoM)* (2014-2016, code number DPI2014-58427-C2-1-R) all of them funded by the CICYT *Ministerio de Economía y Competividad de España*. This thesis is linked to this research and makes use of bio-inspired methodologies and statistical methods for the detection and classification of damages in structures.

Currently, the author does not have any scholarship to perform her thesis. However, she had the support from the Spanish Ministry of Education through *Movilidad de estudiantes en programas de doctorado con mención hacia la excelencia* for her 3 months research stage at the

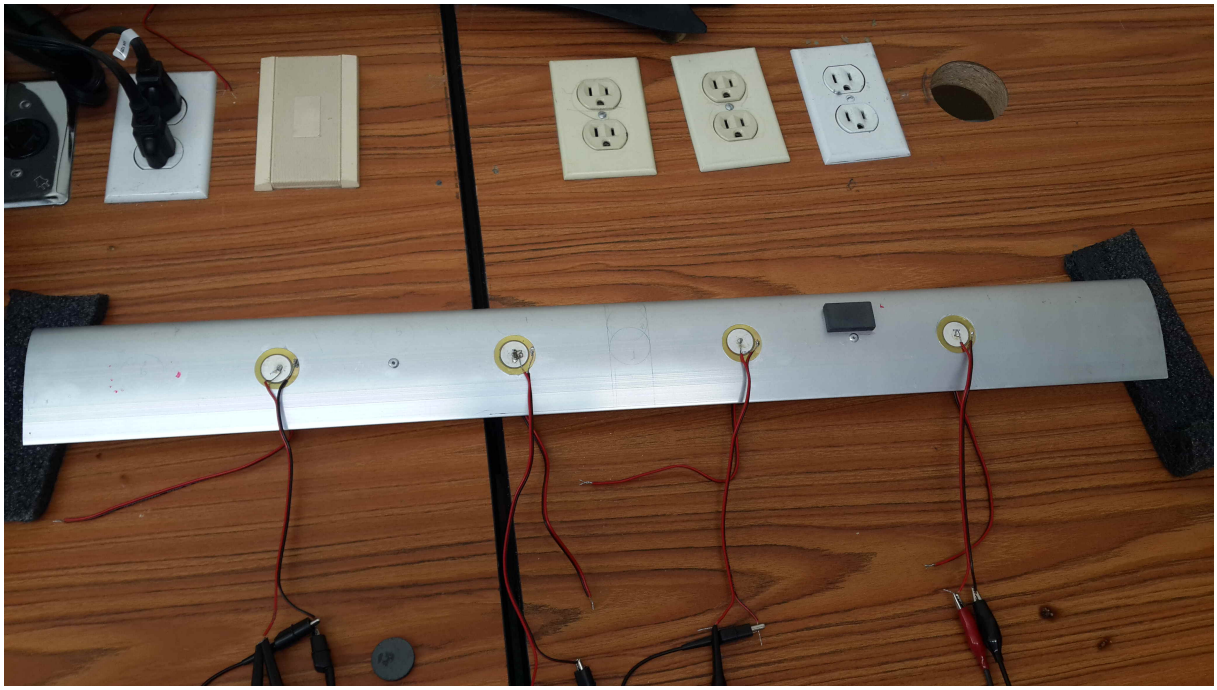


Figure 1.12: Instrumented Blade

University of Siegen (Germany) in 2012.

1.7 Organization

This thesis is organized in five chapters starting with this introduction where the main contribution, the objectives, the general results, a theoretical background, the case studies and the research framework of this thesis are included. Chapters 2 to 4 include the three journal papers published in relevant journals describing the bio-inspired methodologies and its results. Finally, Chapter 5 presents the conclusions and includes a discussion of the directions of the future work. In addition, Appendix A includes a complete list of the publications in conferences, journals and chapter books as a result of the development of this thesis and in collaboration with other related research groups.

Bibliography

[1] C. R. Farrar, S. W. Doeblng, and D. A. Nix. Vibration-based structural damage identification. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences*, 359(1778):131–149, 2001.

[2] S.W. Doebling, C.R. Farrar, and M.B. Prime. A summary review of vibration-based damage identification methods. *The Shock and Vibration Digest*, 30(2):91–105, 1998.

[3] D.A. Tibaduiza. *Design and validation of a structural health monitoring system for aeronautical structures*. PhD thesis, Department of Applied Mathematics III, Universitat Politcnica de Catalunya, 2013.

[4] H. Sohn. Effects of environmental and operational variability on structural health monitoring. *Phil. Trans. R. Soc. A*, 365:539–560, 2007.

[5] N. H. M. Kamrujjaman Serker and Z. S. Wu. structural health monitoring using static and dynamic strain data from long-gage distributed fbg sensor. In Okui Amin and Bhuiyan, editors, *IABSE-JSCE Joint Conference on Advances in Bridge Engineering-II*, pages 519–526, August 8-10 2010.

[6] A. Rytter. *Vibration Based Inspection of Civil Engineering Structures*. PhD thesis, Department of Building Technology and Structural Engineering. Aalborg University, Denmark, 1993.

[7] D. Balageas. *Structural Health Monitoring*, chapter Introduction to Structural Health Monitoring, pages 13–44. ISTE, 2006.

[8] L. N. de Castro and J. Timmis. *Artificial Neural Networks in Pattern Recognition*, chapter Artificial Immune Systems: A Novel Approach to Pattern Recognition, pages 67–84. University of Paisley, 2002.

[9] A.K. Eroglu, Z. Erden, and A. Erden. Bioinspired conceptual design (bicd) approach for hybrid bioinspired robot design process. In *Mechatronics (ICM), 2011 IEEE International Conference on. 2011 , Page(s): 905- 910. Digital Object Identifier: 10.1109/ICMECH.2011.5971243.*, 2011.

[10] N. Cruz Cortès. *Sistema Inmune Artificial para Solucionar Problemas de Optimizacin*. PhD thesis, Centro de Investigacin y de Estudios Avanzados del Instituto Politècnico Nacional, 2004.

[11] P. J. Delves, S. J. Martin, D. R. Burton, and I. M. Roitt. *Roitt's essential immunology*. Wiley-Blackwell, 2011.

[12] L. N. de Castro and F. J. Von Zuben. Artificial immune systems: Part i basic theory and applications. Technical report, School of Computing and Electrical Engineering, State University of Campinas, Brazil, No. DCA-RT 01/99, December 1999.

[13] D. Trejo Pérez. Optimizacin global en espacios restringidos mediante un sistema inmune artificial. Master's thesis, Centro de Investigacin y de Estudios Avanzados del Instituto Politécnico Nacional-Mexixo D.F., 2005.

[14] U. Aickelin and D. Dasgupta. *Search Metodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, chapter Artificial Immune System, pages 375–399. Springer US, 2005.

[15] W. Xiao. *Structural Health Monitoring and Fault Diagnosis based on Artificial Immune System*. PhD thesis, Worcester Polytechnic Institute, 2012.

[16] S. A. Hofmeyr and S. Forrest. Architecture for an artificial immune system. *Evolutionary Computation*, 8(4):443–473, 2000.

[17] J. L. Alexandrino and E. C. de Barros Carvalho Filho. Investigation of a new artificial immune system model applied to pattern recognition. In *Hybrid Intelligent Systems, 2006. HIS'06. Sixth International Conference on. IEEE, 2006.*, 2006.

[18] T. Gong. Artificial immune system based on normal model and immune learning. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, 2008.

[19] L. Hong. Artificial immune system for anomaly detection. In *Knowledge Acquisition and Modeling Workshop, 2008. KAM Workshop 2008. IEEE International Symposium on*, 2008.

[20] K. Worden, W.J. Staszewski, and J.J. Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111, 2011.

[21] S.G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[22] R.R. Coifman and M.V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, 1992.

[23] D.E. Newland. *Random vibration, spectral and wavelet analysis*. New York, NY: Longman, Harlow and John Wiley, 1993.

[24] S. Mallat. *A Wavelet Tour of Signal Processing. 2 ed.* San Diego: Academic Press, 1997.

[25] S. Wold, P. Geladi, K. Esbensen, and J. Ohman. Multiway principal component and pls analysis. *Journal of Chemometrics*, 1:41–56, 1987.

[26] P. Nomikos and J.F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):1361–1375, aug 1994.

[27] G. Cherry and S.J. Qin. Multiblock principal component analysis based on a combined index for semiconductors fault detection and diagnosis. *IEEE Transactions on semiconductors manufacturing.*, Vol. 19, No. 2:159 – 172, 2006.

[28] M. Ruiz, G. Sin, X. Berjaga, J. Colprim, S. Puig, and J. Colomer. Multivariate principal component analysis and case-based reasoning for monitoring, fault detection and diagnosis in a wwtp. *Water science and technology*, , vol. 64, no8:1661–1667, 2011.

[29] B. Mnassri, E. El Adel, and M. Ouladsine. Fault localization using principal component analysis based on a new contribution to the squared prediction error. In *16th Mediterranean Conference on Control and Automation Congress Centre, ajaccio, France. June 25-27*, 2008.

[30] A. Hyvarinen, J. Kahunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, INC, 2001.

[31] L.E. Mujica, J. Rodellar, A. Güemes, and L. López-Diez. Pca based measures: Q-statistics and t^2-statistics for assessing damages in structure. In T. Uhl, W. Ostachowicz, and J. Holnicki-Szulc, editors, *Proceeding of Fourth European Workshop on Structural Health Monitoring*, pages 1088–1095. DEStech Publications, Inc., 2008.

[32] T. Kohonen. The self-organizing map. *Proceeding of the IEEE*, 78, No. 9(9), 1990.

[33] M.A. Torres, D.A. Tibaduiza, L.E. Mujica, J. Rodellar, and C.P. Fritzen. Data-driven multivariate algorithms for damage detection and classification: Evaluation and comparison. *International Journal on Structural Health Monitoring*, 13(1):19–32, January 2014.

[34] M.A. Torres, I. Buethe, D.A. Tibaduiza, J. Rodellar, and C.P. Fritzen. Damage detection and classification in pipework using acousto-ultrasonics and probabilistic non-linear modelling. In *Workshop on Civil Structural Health Monitoring (CSHM-4), Berlin-Germany*, 2012.

[35] N. Kasabov, D. Nikovski, and E. Peev. Speech recognition based on kohonen self-organizing feature maps and hybrid connectionist systems. In *In Proceedings of First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems*, 1993.

[36] J.A. Walter and K.J. Schulten. Implementation of self-organizing neural networks for visuo-motor control of an industrial robot. *IEEE Transactions on Neural Networks*, Vol.4, No.1:86–96, 1993.

[37] Y. Rongfeng, W. Gaolin, Y. Yong, and X. Dianguo. The self-organizing map and fuzzy control for sensorless induction motor speed control. *Natural Computation (ICNC), 2010 Sixth International Conference on*, vol.3:1406–1409, 2010.

[38] B. Ghimire, S. Sinanovic, H. Haas, and G. Auer. Self-organised interference mitigation in wireless networks using busy bursts. In *Applied Sciences in Biomedical and Communication Technologies. ISABEL 2009. 2nd International Symposium on*, 2009.

[39] T. Kohonen and T. Honkela. Essentials of the self-organizing map. *Neural Networks*, 37:52–65, 2013.

[40] V. K. Pachghare, P. Kulkarni, and D.M Nikam. Intrusion detection system using self organizing maps. In *Intelligent Agent Multi-Agent Systems, 2009. IAMA 2009. International Conference on*, 2009.

[41] M. Anaya, D.A. Tibaduiza, and F. Pozo. A bioinspired methodology based on an artificial immune system for damage detection in structural health monitoring. *Shock and Vibration, Article ID 648097, in press*, 2015.

Chapter 2

Data-driven methodology for damage detection and classification under temperature changes

Authors: Maribel Anaya, Diego A. Tibaduiza, Miguel A. Torres-Arredondo, Francesc Pozo, Magda Ruiz, Luis E. Mujica, José Rodellar, Claus-Peter Fritzen
Published in:
Journal: Smart Materials and Structures, 23(4):045006, 2014.
Publisher: iopscience
Online ISSN: 1361-665X
Print ISSN: 0964-1726
2014 Impact Factor: 2.502 (Q1, 8/56 in the area of Instruments and instrumentation)

This chapter is a true copy of the original paper published in the journal where the only changes are performed to fit the page setup.

Data-driven methodology for damage detection and classification under temperature changes

This paper presents a methodology for the detection and classification of damages under different temperature scenarios using a statistical data-driven modelling approach by means of a distributed piezoelectric active sensor network at different actuation phases. An initial baseline pattern for each actuation phase for the healthy structure is built by applying Multiway Principal Component Analysis (MPCA) to wavelet approximation coefficients calculated using the Discrete Wavelet Transform (DWT) from ultrasonic signals which are collected during several experiments. In addition, experiments are performed with the structure in different states (simulated damages), pre-processed and projected into the different baseline patterns for each actuator. Some of these projections and Squared Prediction Errors (SPE) are used as input feature vectors to a Self-Organizing Map (SOM) which is trained and validated in order to build a final pattern with the aim of providing an insight into the classified states. The methodology is tested 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 the kind of damage or the temperature in both structures.

Keywords: Damage Classification, Damage Index, Discrete Wavelet Transform (DWT), Principal Component Analysis (PCA), Self-Organizing Maps (SOM), Structural Health Monitoring (SHM), Temperature Effects.

2.1 Introduction

Structural Health Monitoring (SHM) is an important discipline which attempts to assess the proper performance of a structure. To achieve this objective, SHM systems make use of sensors permanently installed in the structure for inspecting and defining its current state based on the analysis of structural responses. As a result, the collected structural responses are analysed and compared with baseline patterns in order to detect abnormal characteristics and define the structural integrity. The obtained information can be used to define whether the structure can operate under safe and reliable conditions.

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 which 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 data collected. However, to ensure the reliability of the analysis performed to the signals collected in several experiments, it is necessary to consider the variation in the environmental and operational conditions, the proper functioning of the sensors, actuators and hardware used to inspect the structure.

Among the big quantity of damage types that can be found in the normal service of a structure, the following categories can be distinguished [1, 2]:

- Gradual damage such as fatigue, corrosion and aging.
- Sudden and predictable damage coming from aircraft landing and planned explosions in confinement vessels.
- Sudden and unpredictable damage originating from foreign-object-impact, earthquakes and wind loads.

At the same time, these different kinds of damage can also be classified depending on their severity in three big groups [3]:

- Light damage: This corresponds to the initial stage of a damage, which can be relatively easily-repairable and is not dangerous for the normal operation of the structure.
- Moderate damage: In comparison with the previous one, this damage requires major repairs and needs to be evaluated more carefully in order to define if the structure can operate in normal conditions.
- Severe damage: This type of damage unlike the previous damages requires big reparations or the replacement of the structure.

For a structure in service, the variability in its dynamic properties can be a result of time-varying environmental and operational conditions [4]. Some of the environmental conditions to consider are humidity, wind loads, temperature and pressure, among others. Operational conditions include loading conditions, operational speed and mass loading [4]. The design of methodologies to improve the damage identification including environmental and operational conditions is currently an area of active research interest. This interest is motivated by the fact that new designs in civil, aeronautics and astronautics include the use of more complex structures subjected to variable environmental and operational conditions. To assess the structural integrity it is necessary to have a reliable continuous monitoring that allows avoiding false alarms resulting from these variable conditions. Since a structure in normal operating conditions undergoes temperature variations and different forces by the environmental conditions, it is necessary to understand the effects of such conditions and test the developed methodologies under these conditions. In this sense, several works from different authors can be found in literature. For instance, in the work by Fritzen et al. [5] an existing subspace-based identification method is modified including temperature compensation for damage diagnosis. In this work, the detection is focused on the detection of small delaminations in composite materials and the excitation is performed by random noise to collect the system's vibration response at different temperatures. Konstantinidis et al. [6] applied optimal baseline subtraction (OBS) to investigate its robustness under temperature changes for damage localization by using an aluminium plate, piezoelectric discs and temperature changes in the range of $22^{\circ}\mathrm{C}$ to $32^{\circ}\mathrm{C}$. Sohn [4], presented a review of the environmental and operational variation effects on real structures focused on data normalization progresses. Moll et al. [7] studied the compensation of environmental influences using a simulation model and a laboratory structure. Raghavan and Cesnik [8] reported studies for the selection of suitable bonding agent for piezoelectric transducers on aluminium plates.

These studies were conducted in the temperature range of $20°\mathrm{C}$ to $150°\mathrm{C}$ and the identified thermally sensitive variables in the experiments were used to model the experimentally observed changes under temperature variation. Lu and Michaels [9] used ROC curves in order to find selective features which are sensitive to damage but insensitive to the applied surface wetting. Deraemaeker et al. [10] used factor analysis to treat the environmental changes and statistical process control for damage detection. In the study, a numerical example of a bridge subjected to environmental changes is used. More recently, Mujica et al. [11] studied the use of PCA, extended PCA, T^2-statistic and Q-statistic to detect and distinguish damages in structures under varying operational and environmental conditions. Croxford et al. [12] evaluated two different methods to compensate the temperature effect, namely optimal baseline selection (OBS) and baseline signal stretch (BSS). A combination of these methods was proposed to improve the effective temperature compensation. The methodology was tested in two independent long-term experiments using both fundamental modes of propagation (A0 mode and S0 mode). In 2011, Kraemer et al. [13] proposed an approach based on Artificial Neural Networks (ANN) using Self Organizing Maps (SOM) in order to compensate the temperature effects on different features obtained from measured time data of a Carbon Fiber Reinforced Polymer plate (CFRP). Torres and Fritzen [14] presented theoretical developments, numerical and experimental results on the effects of variable temperature and operation conditions (EOC) on wave propagation in composite materials. Increase in time-of-flight, effects of surface wetting and changes in sensor response magnitude with temperature were analyzed and discussed. Recently, Dodson and Inman [15] studied the physics of thermal effects on Lamb waves in plates. This work included the discussion and experimental verification in a thin plate of the thermal sensitivity of dispersion curves.

In the present paper, the authors suggest the use of Discrete Wavelet Transform (DWT), Multiway Principal Component Analysis (MPCA) and Self-Organizing Maps (SOM) to detect and classify damages in structures using signals collected from acousto-ultrasonics tests by considering temperature changes. The methodology 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 (by adding a mass at different positions of the structure) and two damages in a sensor (25% and 50% breakage). On the other hand, six states are studied in the second structure: these states include the undamaged structure and 5 simulated damages 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. In difference with previous works by the authors this work includes a study about the temperature effects in the damage detection and classification methodology. Also, it is proposed the use of a type unfolding to consider the data from all the temperatures evaluated in order to define a robust baseline for detection and classification damage. In this work, the term "robust baselines" is considered to define baselines obtained from different temperatures. Finally, it is also included the use of the U-matrix surface as a tool for visualizing the results. In this way, the U-matrix surface as an output of the SOM allows validating the results obtained by the cluster-map. The paper is organised as follows: Section 2.2 presents the experimental setup followed by the basic theoretical background on the concepts of Discrete Wavelet Transform, PCA and Self-Organizing Maps in the section 2.3. The methodologies for the damage detection and classification are presented in the Section 2.4. The results of the application of the methodology to the different

specimens are presented in the section 2.5. Finally, conclusions are drawn in Section 2.6.

2.2 Experimental setup

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 2.1. Six damages have been simulated on the structure. The first four damages are simulated by placing magnets on both sides of the structure by using a mass at different positions on the structure and the other two damages correspond to a broken sensor with a different percentage of breakage. More precisely, damage 5 refers to a 25% breakage and damage 6 corresponds to a 50% breakage. The location of the damages is shown in Figure 2.1.

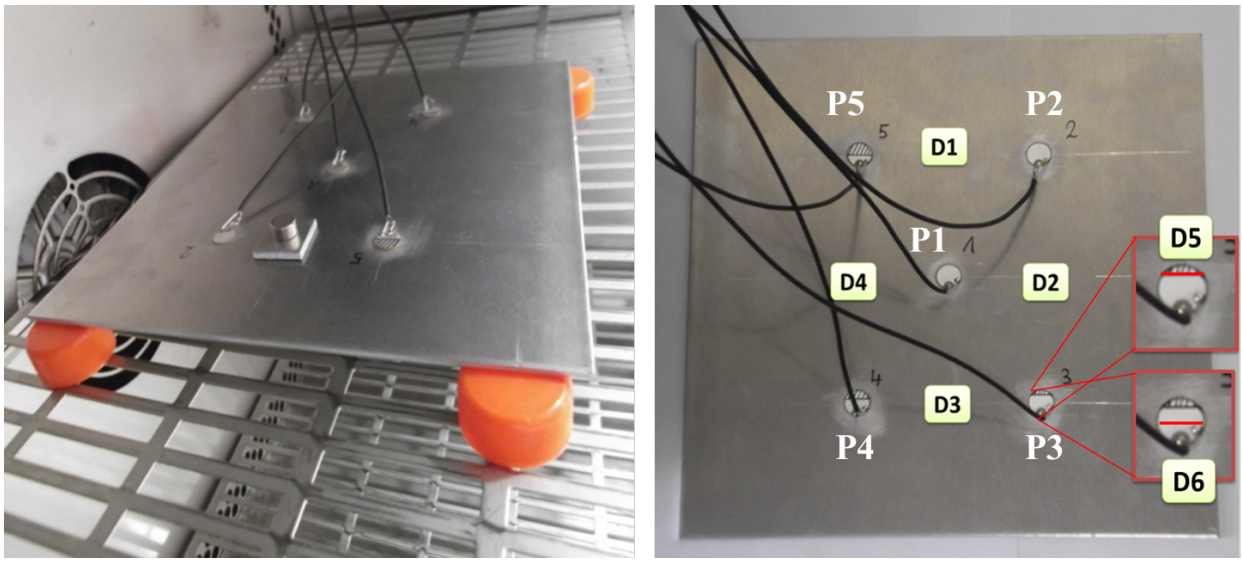


Figure 2.1: Aluminium plate and damage description.

To inspect the structure, a 12V Hanning windowed cosine train signal with 5 cycles and 50 KHz as central frequency was used. This 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 (24°C, 30°C, 35°C, 40°C, 45°C and 50°C) for each structural state were collected. To avoid the contact between the metallic surfaces and eliminate the noise in the experiments, plastic elements are used. 100 experiments were collected for each state and for each temperature.

The second structure is a simplified aircraft composite skin panel made of carbon fibre reinforced plastic (CFRP).The structure is depicted in Figure 2.2.. The overall size of the plate is approximately $500 \times 500 \times 1.9$ mm 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 V_{Fibre} = 60%, E_1=142.6GPa, E_2=9.65GPa, v_{12}=0.334, v_{13}= 0.328, v_{23} = 0.54 and G_{12} = G_{13} = 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 2.2. outlines the coordinates for the simulated damage on the composite skin panel. The excitation voltage signal is a 12V Hanning windowed cosine train signal with 5 cycles; 120 experiments were recorded per sensor-actuator configuration at five different temperatures (35°C, 45°C, 55°C, 65°C and 75°C). To determine the carrier central frequency for the actuation signal in each structure, a frequency sweep was performed and the spectral content of each signal was analyzed. The carrier frequencies were found to be 50 kHz. The carrier frequency was chosen to maximize the propagation efficiency. This type of excitation generates a dominant A0 mode that is propagated along the structure allowing a better interaction of the guided wave with the simulated damage.

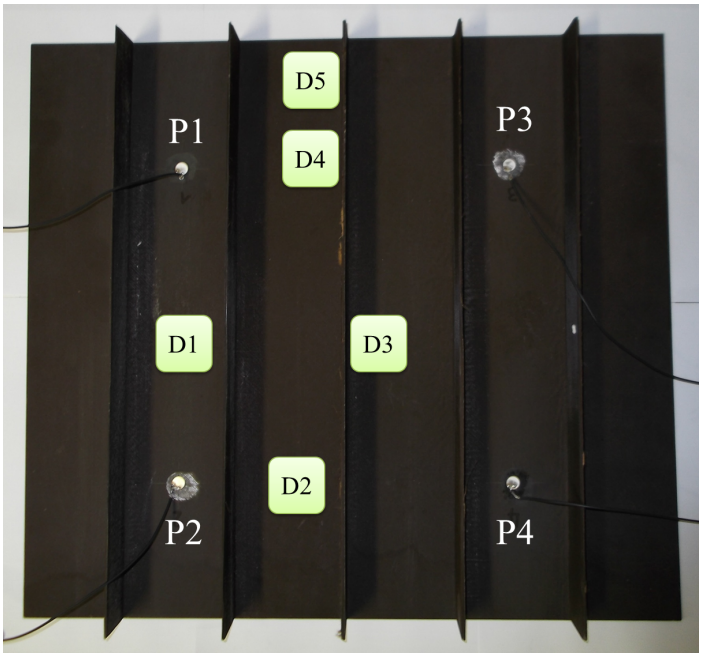


Figure 2.2: Simplified Aircraft Composite Skin Panel: Setup and Damage Positions.

2.3 Theoretical background

The damage detection and classification methodology presented in this work is based on the use of data-driven analysis. The analysis is performed by pre-processing the data collected from the structures using Discrete Wavelet Transform and applying Multiway Principal Component Analysis and damage indices as pattern recognition techniques. Finally the results obtained in the statistical modelling with PCA and the damage indices are included in a classifier to perform the final analysis. As classifier, Self-Organizing Maps are used. Details about how these steps are applied will be explained in the section 2.4. However, a brief introduction to each method used in this methodology is included in this section. More details can be consulted in the references included in each method.

2.3.1 Discrete Wavelet Transform

The discrete wavelet transform (DWT) is a powerful tool specially used in areas dealing with the analysis of transient signals. This transform provides sufficient information for both the analysis and synthesis of the original signal with a significant reduction in the computational time cost compared with the analysis of the wavelet in continuous time domain normally called Continuous Wavelet Transform (CWT) [16]. Roughly speaking, the characteristics of the given function at a specified time and scale are represented by the transformation coefficients. In other words, each wavelet coefficient represents time and frequency information of the signal. The DWT analysis can be performed by means of a fast, pyramidal algorithm related to a two-channel subband coding scheme using a special class of filters called quadrature mirror filters as proposed by Mallat [17]. In this algorithm, the signal is analyzed at different frequency bands with different resolution by decomposing the signal into approximation and detail coefficients. This is achieved by successive high-pass (HP) and low-pass (LP) filtering of the input signal. The optimum number of level decompositions could be determined, for example, based on a minimum-entropy decomposition algorithm [18]. Figure 2.3 shows the structure of the decomposition of the signal and the corresponding frequency bandwidths for the details (Di) and approximations (Ai) [17]. Each level is obtained by the decomposition of the signal with high and low pass filtering and a down-sampling of the input signal. The approximations represent the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. The frequency fmax represents the maximum frequency contained in the recorded signal and n is the decomposition level. Discrete wavelets are often associated with dilation equations and scaling functions. The basic scaling function can be defined as:

$$\gamma(i) = \sum_{m=0}^{j} c_m \gamma(2i - m), \tag{2.1}$$

where the values c_m, $m = 0, \ldots, j$ are the wavelet coefficients. These coefficients must satisfy certain conditions which are discussed in [19]. The scaling functions are then used to construct the corresponding wavelets $\phi(i)$ from the following formula:

$$\phi(i) = \sum_{m=0}^{j} (-1)^m c_m \gamma(2i + m - j + 1). \tag{2.2}$$

The approximation and detail coefficients can be calculated as follows:

$$A_{n,k} = 2^{-n/2} \sum_{i=1}^{j} x(i) \gamma(2^{-n} j - k) \tag{2.3}$$

$$D_{n,k} = 2^{-n/2} \sum_{i=1}^{j} x(i) \varphi(2^{-n} j - k) \tag{2.4}$$

where γ is called the scaling function, j is the number of discrete points of the input signal, n and k are considered to be the scaling (dilation) index and the translation index, respectively. Each value of n allows analyzing a different resolution level of the input signal. Scaling functions are similar to wavelet functions except that they have only positive values and are designed to smooth the input signal [20, 17].

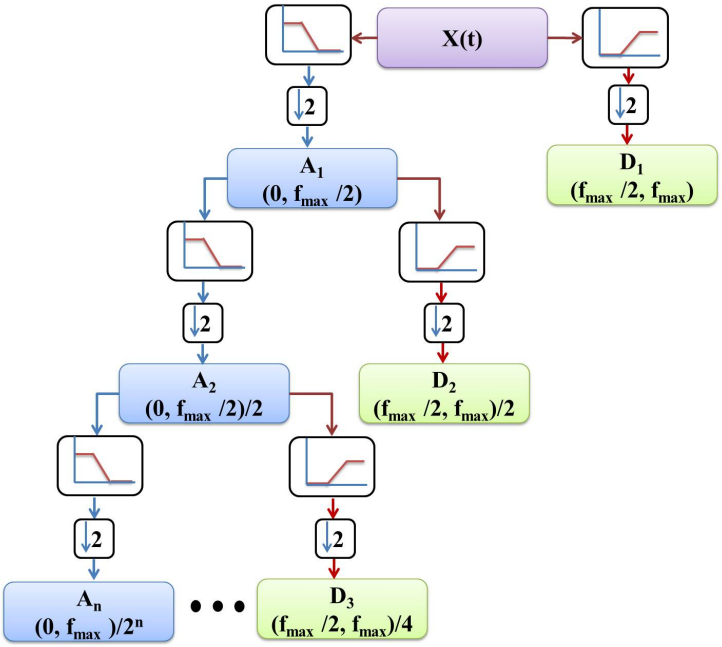


Figure 2.3: DWT Decomposition Tree.

2.3.2 Multiway Principal Component Analysis (MPCA)

MPCA is a very common practice in multivariate statistical procedures for monitoring the progress of batch processes [21, 22] and it is based in the fact that the data are inherently correlated [23]. This is similar to the use of PCA on a large two-dimensional matrix constructed by unfolding the three-way data matrix [24]. In this work the original data are expressed in a matrix whose dimensions are I experiments, K time instants and J sensors. This three-way data matrix is decomposed into a large two-dimensional matrix $\mathbf{X}$ as shown in Figure 2.4.

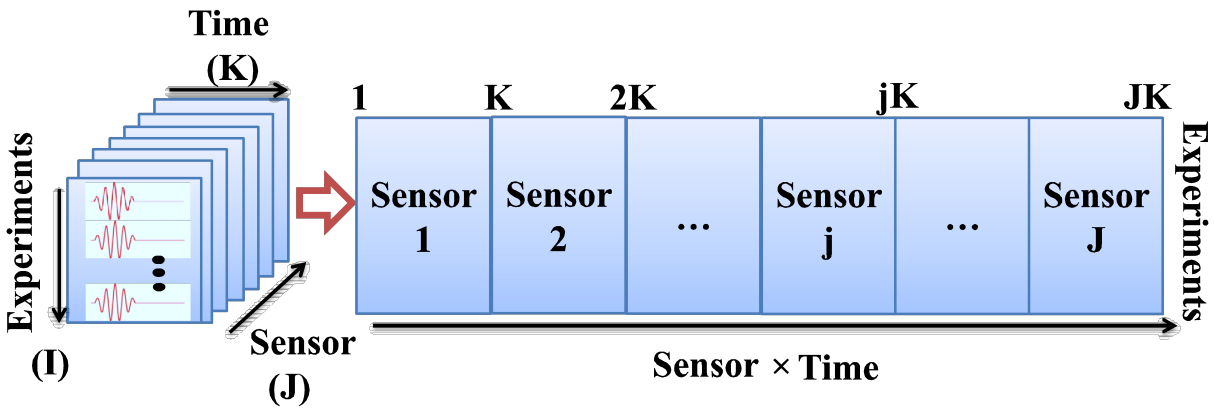


Figure 2.4: Unfolding the original data.

2.3.3 Principal Component Analysis

Principal Component Analysis (PCA) is a useful technique of multivariable and megavariate analysis [25] that provides arguments to reduce a complex data set to a lower dimension and reveals some hidden and simplified structure/patterns that often underlie it. The goal of PCA is to discern which information are more important in the system by determining a new space (coordinates) to re-express the original data based on the variance-covariance structure. 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. In order to develop a PCA model, it is necessary to arrange the collected data in a matrix $\mathbf{X}$. This $n \times m$ matrix contains information from m sensors and n experiments [26]. Since physical variables and sensors have different magnitudes and scales, each data-point is scaled using the mean and the standard deviation of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor. Once the variables are scaled, the covariance matrix $\mathbf{C_X}$ is calculated as follows:

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}. \tag{2.5}$$

This 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 follows:

$$\mathbf{C_X}\widetilde{\mathbf{P}} = \widetilde{\mathbf{P}}\mathbf{\Lambda}, \tag{2.6}$$

where the eigenvectors of $\mathbf{C_X}$ are the columns of $\widetilde{\mathbf{P}}$, and the eigenvalues are the diagonal terms of $\mathbf{\Lambda}$ (the off-diagonal terms are zero). Columns of matrix $\widetilde{\mathbf{P}}$ are sorted according to the eigenvalues by descending order and they are called principal components of the data set or loading vectors. The eigenvectors with the highest eigenvalue represents the most important pattern in the data with the largest quantity of information. Choosing only a reduced number $r < m$ of principal components, those corresponding to the largest eigenvalues, the reduced transformation matrix could be imagined as a model for the structure. In this way, the new matrix $\mathbf{P}$ ($\widetilde{\mathbf{P}}$ sorted and reduced) can be called as PCA model. Geometrically, the transformed data matrix $\mathbf{T}$ (score matrix) represents the projection of the original data over the direction of the principal components $\mathbf{P}$:

$$\mathbf{T} = \mathbf{XP}. \tag{2.7}$$

In the full dimension case (using $\widetilde{\mathbf{P}}$), this projection is invertible (since $\widetilde{\mathbf{P}}\ \widetilde{\mathbf{P}}^T$=$\mathbf{I}$) and the original data can be recovered as $\mathbf{X} = \mathbf{T}\widetilde{\mathbf{P}}^T$. In the reduced case (using $\mathbf{P}$), with the given $\mathbf{T}$, it is not possible to fully recover $\mathbf{X}$, but $\mathbf{T}$ can be projected back onto the original m-dimensional space and obtain another data matrix as follows:

$$\widehat{\mathbf{X}} = \mathbf{TP}^T = (\mathbf{XP})\mathbf{P}^T. \tag{2.8}$$

Therefore, the residual data matrix (the error when not all the principal components are used) can be defined as the difference between the original data and the projected back:

$$\mathbf{E} = \mathbf{X} - \widehat{\mathbf{X}} = \mathbf{X}(\mathbf{I} - \mathbf{PP}^T). \tag{2.9}$$

There are several kinds of indices that can give information about the accuracy of the model and/or the adjustment of each experiment to the model. Two well-known indices are commonly used to this aim: the Q-statistic (or SPE-index) and the Hotelling's T^2-statistic (D-index) [27, 28, 26, 29]. In a previous work by the authors [30], it was demonstrated that the use of the SPE-index and the scores allow a good classification, since the T^2-statistic is a measure calculated from the scores, and its use in the methodology in combination with the scores introduces redundancies to the analysis. The former is based on analyzing the residual data matrix to represent the variability of the data projection within the residual subspace. Denoting $\mathbf{e}_i$ as the i-th row of the matrix $\mathbf{E}$, the Q-statistic for each experiment can be defined as its squared norm as follows:

$$Q_i = \mathbf{e}_i \mathbf{e}_i^T = \mathbf{x}_i(\mathbf{I} - \mathbf{P}\mathbf{P}^T)\mathbf{x}_i^T. \tag{2.10}$$

2.3.4 Self-Organizing Map (SOM)

The Self Organizing Map (SOM) is a kind of unsupervised neural network also known as Kohonen network [31] in honour to professor Teuvo Kohohenen. This neuronal network is specialized in visualization and analysis of high dimensional data and has the special property of generating one organized map in the output based on the inputs, grouping input data with similar characteristics in clusters [32, 33]. To do that, the SOM internally organizes the data based on features and their abstractions from input data. In particular, these maps have been used in practical speech recognitions [34], robotics [35], control [36] and telecommunications [37], among others [38].

The structure of the SOM is a single feed forward network and each node in the input layer is connected to all the output neurons [39]. The connection between the input layer and the output layer is defined by a weighting matrix as shown in Figure 2.5. In general, the inputs to the network are defined in the input layer and a weighting matrix relates each input with each cluster in the output layer.

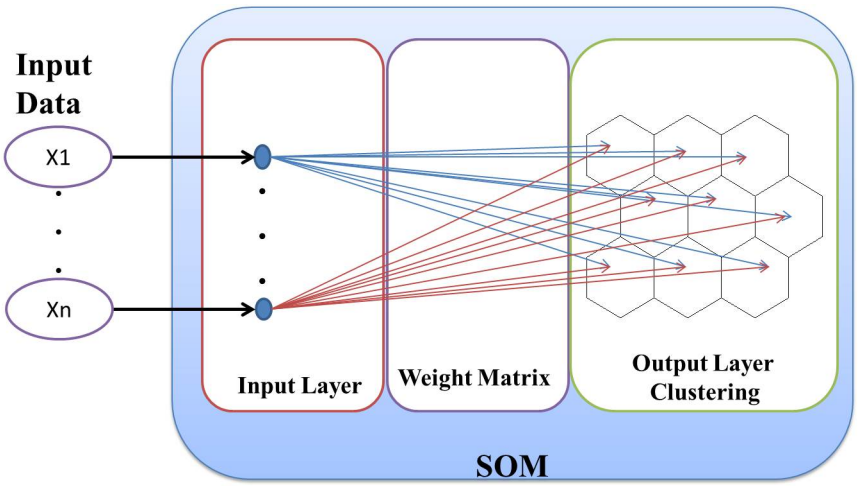


Figure 2.5: The Self-organized map.

The result of the SOM is an organised map which contains the final clustering of the input

data. This information can be visualised by using the cluster map. The distance matrix techniques are widely used in the visualization of the cluster structure of the SOM. One of these techniques is the unified distance matrix, called U-matrix, which shows the separation between prototype vectors and neighbouring map units. The U-matrix is related with the linkage measure and can be visualized using a colour scale [40].

2.4 Damage detection and classification

In this paper, a methodology is proposed for damage detection and classification, which is partly based on a methodology previously developed by the authors [30],[3]. 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. In this case, a robust baseline is defined as a baseline with the data from the healthy structure at different temperatures. The multi-actuator piezoelectric system consists of several piezoelectric transducers attached to the surface of the structure under test working in several actuation phases. In each actuation phase, one lead zirconium titanate (PZT) transducer is selected as actuator and a signal is applied to produce Lamb waves in the structure. 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 order to obtain the approximation coefficient. 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. This was accomplished by means of different trial and error tests by evaluating different mother wavelets and levels of decomposition so that the signal could be properly reconstructed from the calculated wavelet coefficients. The chosen 'db8' wavelet is an orthogonal wavelet with the advantage of avoiding phase shifts and allowing exact reconstruction of the signal. The index number refers to the number of coefficients. The number of vanishing moments for each wavelet is equal to half the number of coefficients, so the 'db8' has 4 vanishing moments. A vanishing moment confines the ability of the wavelet to represent polynomial behaviour [41, 20]. Special attention was paid to selection of the optimum decomposition level in order to avoid removing important information that could be related to some of the modes of propagation contained in the signal. The optimum number of level decompositions is determined based on both a minimum-entropy decomposition algorithm and systematic trials [18]. The entropy-based methods perform a systematic comparison of the signal and noise using the Shannon information theory [20]. It has been shown that a pure noise signal has the largest entropy while a systematic signal has zero entropy [42].

Figures 2.6, 2.7 and 2.8 show 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 (IK), 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 (IKJ) or in a two-dimensional unfolded matrix (IJK), where data from each

sensor are located besides the other sensors as in the Figure 2.4. This is a very common practice in multivariate statistical procedures for monitoring the progress of batch processes [21, 22]. To manage the use of different temperatures, in this paper the data (approximation coefficient) from each temperature are unfolded and organized in a matrix as shown in Figure 2.6 in order to obtain a matrix by each actuator phase with the information from all the sensors at different temperatures.

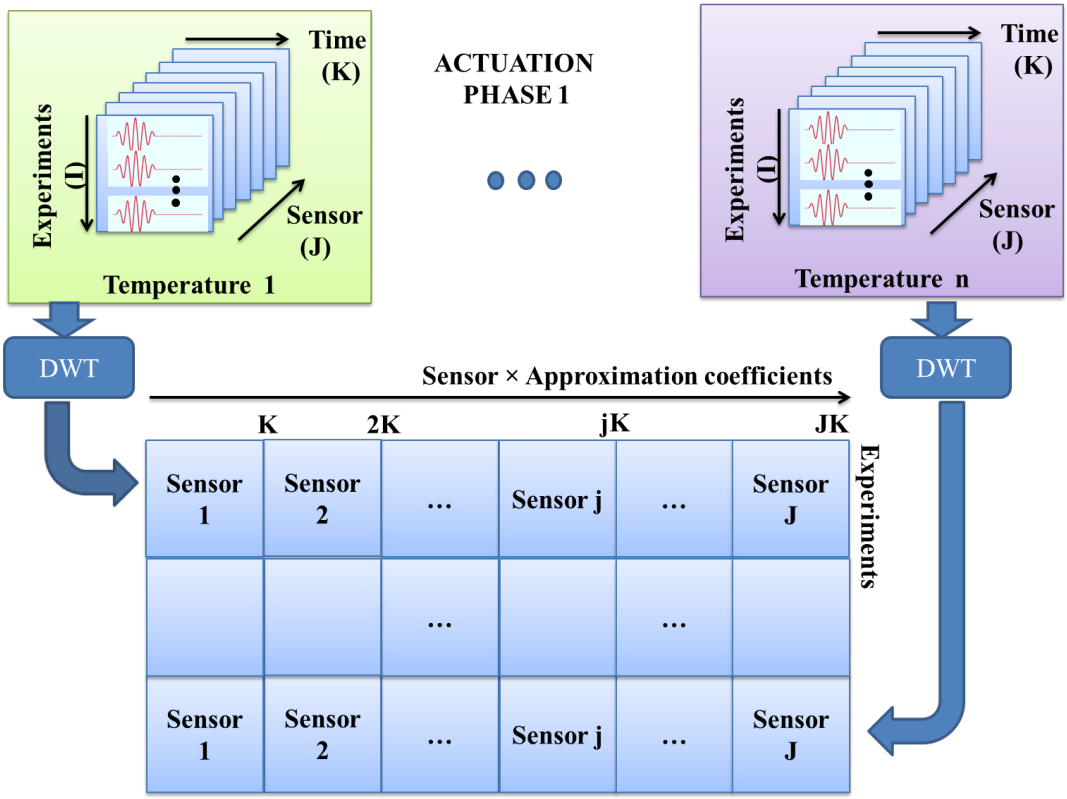


Figure 2.6: Unfolding the collected data in the actuation phase 1 using n temperatures.

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) [43]. 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. This way, the data are reduced by selecting a reduced number of components according to the cumulative percent variance (CPV) approach. The CPV approach guarantees the minimal model dimension that captures the desired variance.

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. 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 (equation (2.7.)), and the scores and the SPE-index ($Q - index$) are obtained, as represented in Figure 2.7.

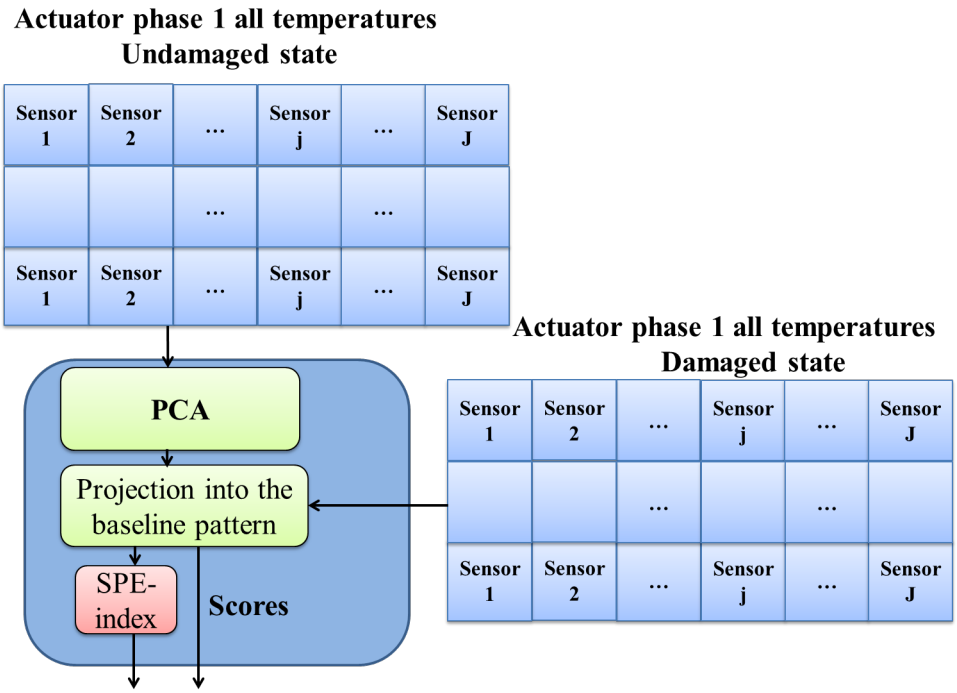


Figure 2.7: Calculation of the baseline pattern using PCA and calculation of scores and SPE-index in the actuator phase 1.

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 2.8. 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 [30]. 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.

The proposed methodology is indeed a qualitative approach since a quantitative assessment of damage is not provided in the presented work.

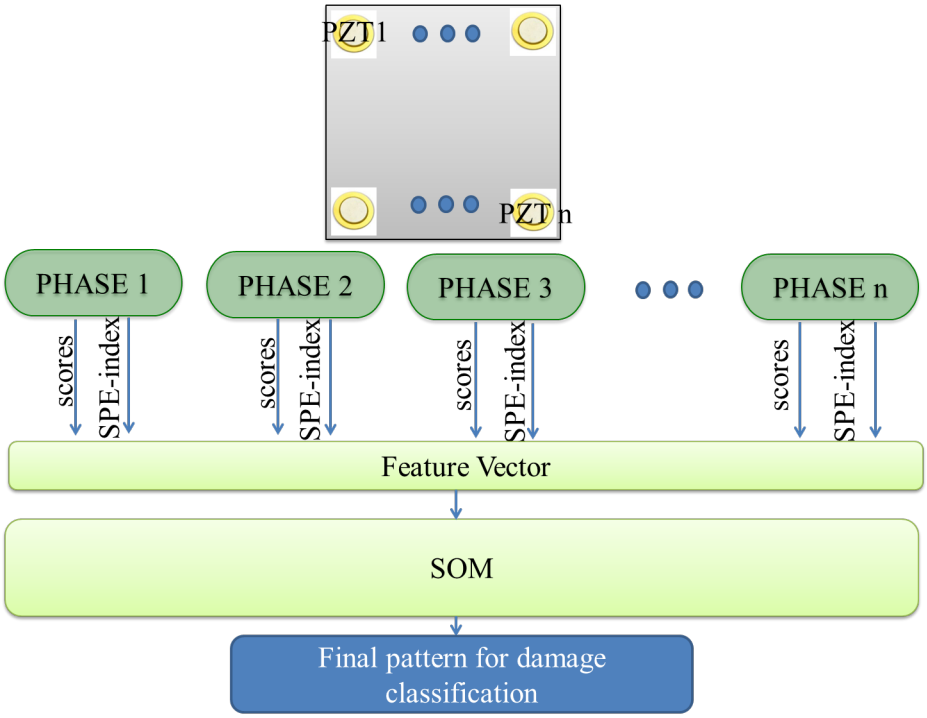


Figure 2.8: Final pattern for damage classification.

2.5 Analysis and discussion of the results

2.5.1 Aluminium plate

The following three studies were conducted in order to verify and determine the approach sensitivity and the influence of the temperature variations in the detection and classification of damage:

- Classification of baselines at different temperatures
- Damage detection and classification at each temperature level
- Damage detection and classification using all temperatures

Classification of baselines at different temperatures

Firstly, the classification of the different baselines at the different temperatures is evaluated. In particular, data from the structure at six different temperatures ($24°C$, $30°C$, $35°C$, $40°C$, $45°C$ and $50°C$) were used to evaluate the approach.

Following the scheme described in Section 2.4, the wavelet approximation coefficients from the data of the healthy structure at different temperatures are computed. These approximation coefficients are organized as in Figure 2.4 and then the baselines (PCA models) are built. To determine the number of principal components that will be used to project the new data, the variance captured by each principal component was computed for each temperature in order to determine the number of principal component. This analysis is fundamental in order to ensure

that enough variance is captured by the model, therefore, allowing an optimal reduction. The distribution of the retained variance along the principal components in two of these actuator phases is depicted in Figure 2.9. Only the components with a significant variance value are shown. The components with the highest variance represent variables or factors that describe major trends in a confusing data set. Similar results are obtained in the other actuation phases. From this analysis, the first ten principal components are selected and used to define the PCA model by each actuator phase. This number of components considers more than 90% of variance in the statistical modelling.

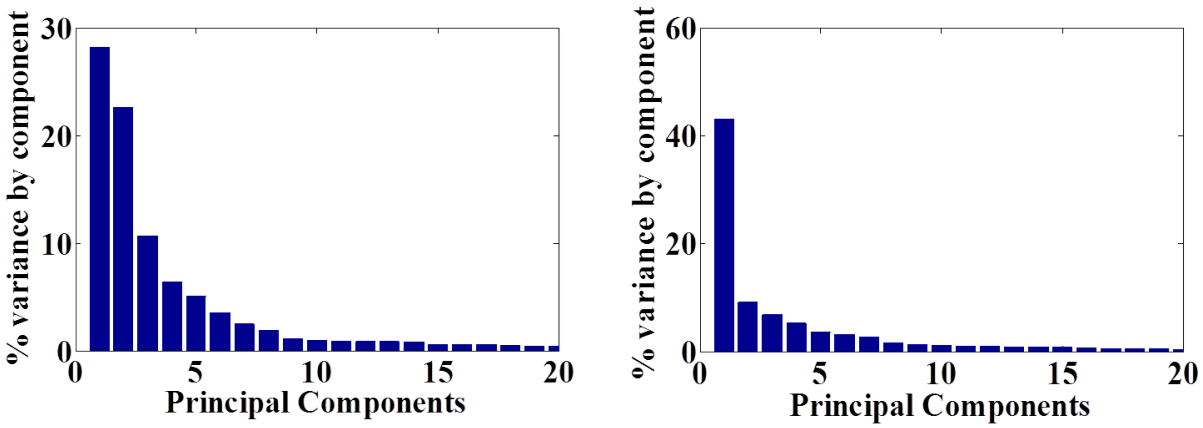


Figure 2.9: Percent of variance in the aluminium plate in (a) actuation phase 1 and (b) actuation phase 5.

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 as follows:

$$F_i = [\text{score}_i^k \quad Q_i^k], \tag{2.11}$$

where score $= [\text{score}_1, \ldots, \text{score}_n]$ with n principal components, k corresponds to the number of actuator phase and i is the number of the experiment.

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 to normalize the feature vector. Normalization type *histD* is a discrete histogram equalization that sorts the values of the feature vector and replaces each value by its ordinal number [44]. Finally, it scales the values in such a way that they lie between the range [0,1]. 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 30×10 is obtained to train the SOM.

The results of the training are presented by means of both the cluster map, in Figure 2.10, and the U-matrix surface in Figure 2.11. The analysis of these plots allows one to find differences between the data from the healthy structure at different temperatures. In this case six clusters seem to have been well identified in the cluster map and in the U-matrix surface. More precisely, the observations of the cluster map shows that the data from the structure when the

temperature is $24°C$ are better organized. This fact can be inferred by reviewing the number of output neurons occupied in the cluster map. In the U-matrix surface the boundaries represent the separation between the clusters and the colour shows the magnitude of the differences between the clusters. In this sense, higher value colours according to the colour scale imply large separations. According to the colours in the boundaries, it can be seen that there are not significant differences between the first four temperatures. However, this difference is remarkable with respect to the last two temperatures. Another important issue is related to the management of the outliers. In the cluster map presented in Figure 2.10, there is an outlier that corresponds to the structure at $45°C$. In this case, it is worth noting that the U-matrix surface (Figure 2.11) allows the isolation of the outlier by the separation between the different areas. This separation represents the different zones delimited by boundaries, where the colours with low magnitude according to the colour scale depict these zones. The colours with high magnitude into these boundaries can be interpreted as the zones with a high separation between the clusters or the zones where the outliers appear. The result of this first study demonstrates the influence of the temperature despite the fact that the data comes from the healthy structure. These differences are not so evident with respect to the first temperatures.

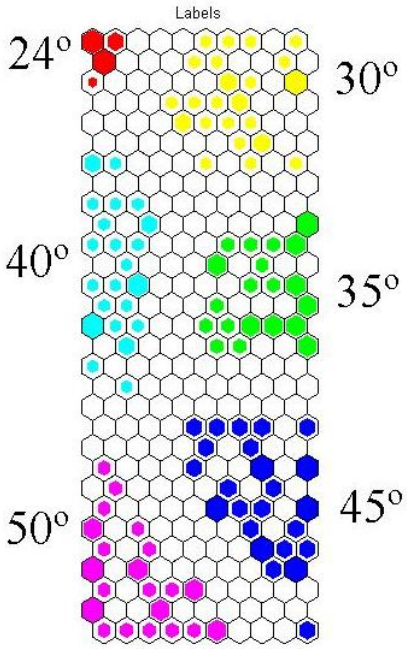


Figure 2.10: Classification of the different baselines at different temperatures using 10 scores and the SPE-index using the cluster map.

Damage detection and classification by each temperature

In the second study, the maps were trained using the data from seven different states (healthy state and six damages). The normalization, the hexagonal lattice and the cluster size were defined as in the previous study. Four damages were simulated in the tested plate by adding a mass at different locations on the surface of the structure. In addition, two real damages in

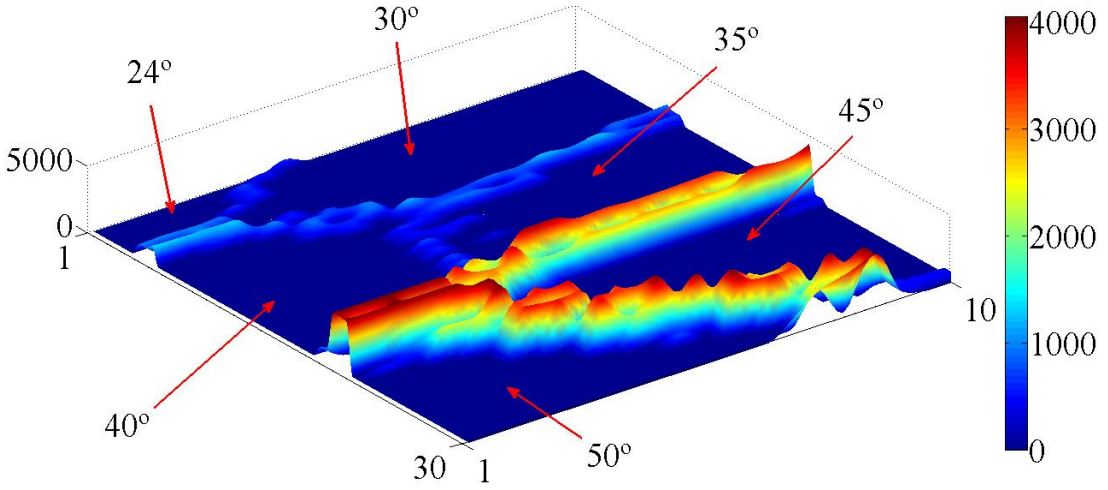


Figure 2.11: Classification of the different baselines at different temperatures using 10 scores and the SPE-index using the U-matrix surface.

a sensor were performed. These damages correspond to a breakage at 25% and 50% of the sensor 3. Figures 2.12 to 2.17 show the classification results when the structure is exposed to a temperature of $24°C$, $40°C$ and $50°C$.
When the structure is exposed to a temperature of $24°C$, seven zones seem to have been well identified, as can be seen in Figures 2.12 and 2.13. The boundaries between the clusters in the U-matrix surface show that there is a clear separation between all the states. Moreover, and according to the cluster map, there are no outliers.

Figures 2.14 and 2.15 show the results when the structure is exposed to a temperature of $40°C$. As it can be seen, there are seven clusters well identified in the cluster map and the U-matrix surface. As in the previous case, there is no presence of outliers. However, the damage distribution across the map is different.
The results when the structure is exposed to $50°C$ (Figures 2.16 and 2.17) show that all the states are clearly classified. The lowest boundary is presented between the data from the undamaged state and damage 2 in the U-matrix surface. If the results of the classification are compared with respect to all the temperatures, it can be concluded that the damages have been correctly classified. Moreover, the way the outliers have been treated does not imply the creation of a new cluster. The undamaged state in all the temperatures is clearly separated from the rest of the damages although the number of output neurons in the cluster map is different for each temperature.

The results in the previous two studies have shown that the proposed methodology applied to data coming from dynamic responses at different temperatures allows a proper final classification in the cluster map and the U-matrix surface. However, these results change from one temperature to another and the distribution at each temperature is different. Therefore, these results suggest that changes in the temperature during the inspection can probably lead to false results in the final classification.

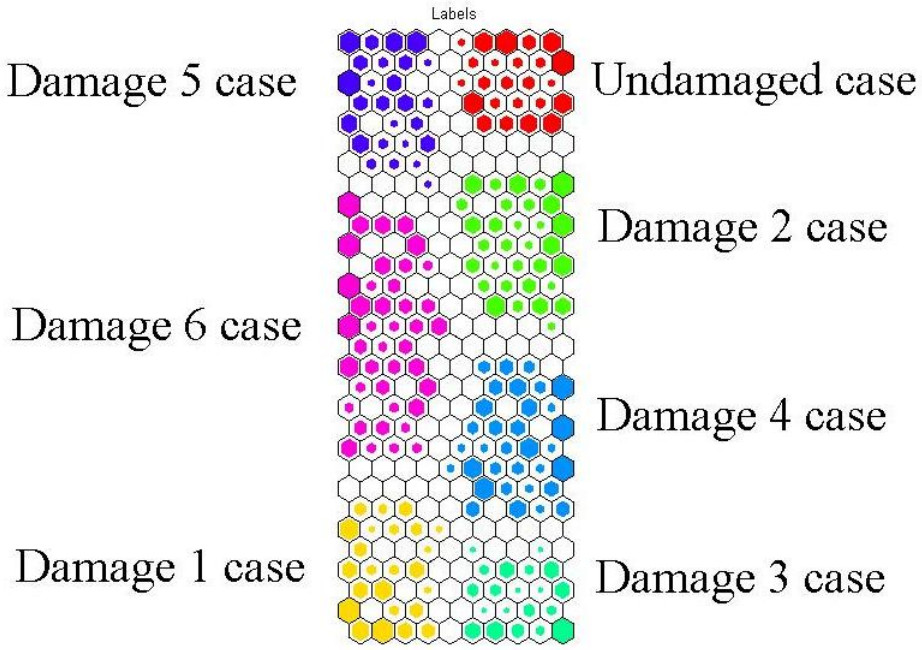


Figure 2.12: Classification of the different states at 24°C using 10 scores and the SPE-index using the cluster map.

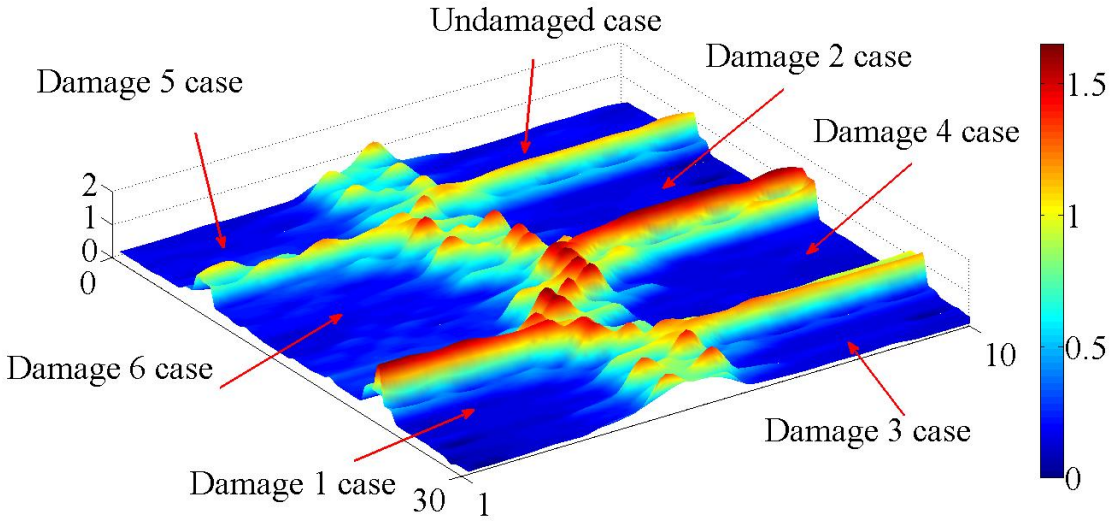


Figure 2.13: Classification of the different states at 24°C using 10 scores and the SPE-index using the U-matrix surface.

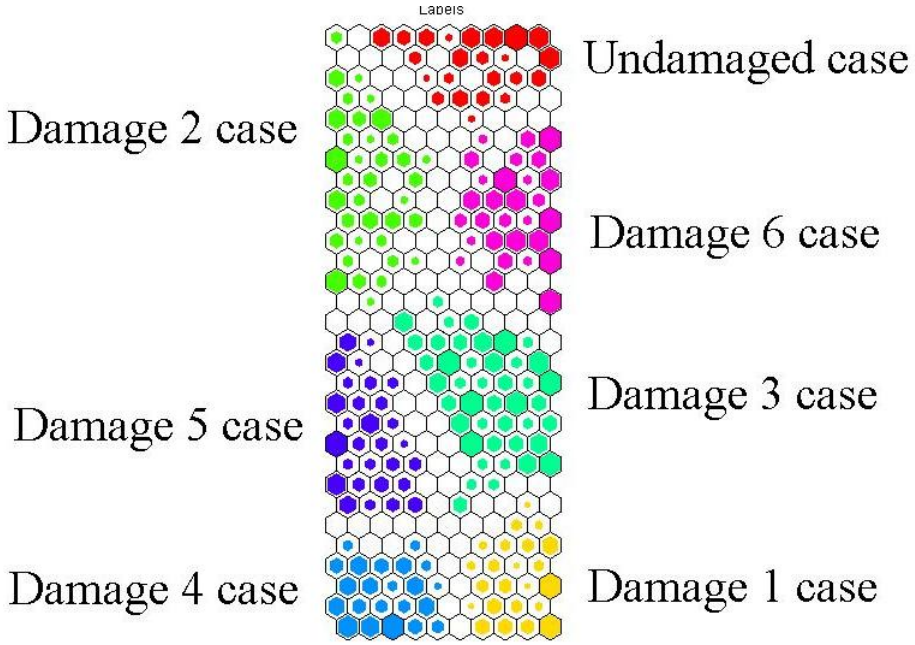


Figure 2.14: Classification of the different states at 40°C using 10 scores and the SPE-index using the cluster map.

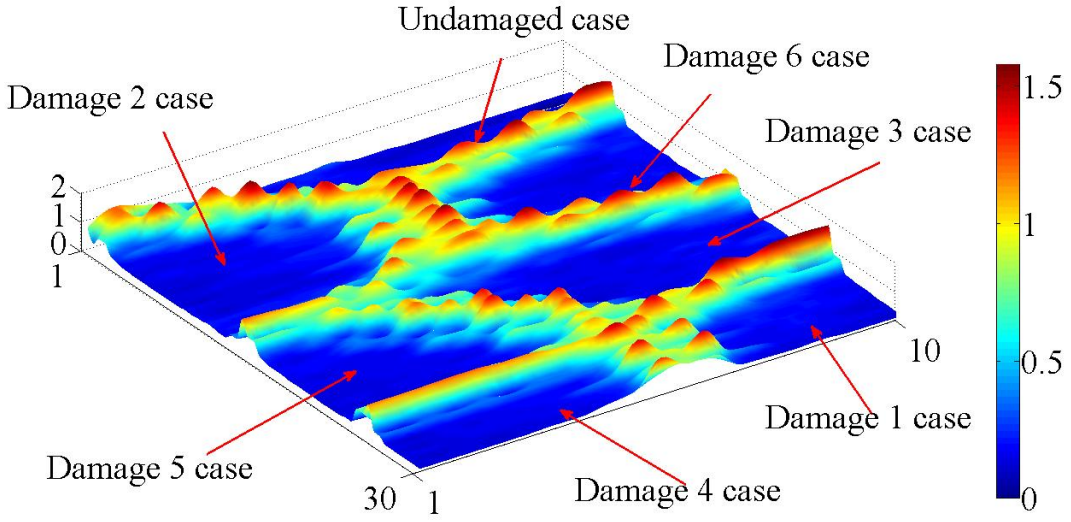


Figure 2.15: Classification of the different states at 40°C using 10 scores and the SPE-index using the U-matrix surface.

Damage detection and classification using all temperatures

The third study with the aluminium plate is aimed to solve the question about how it is possible to improve the final classification pattern by considering the data covering different temperature ranges. As a possible solution, in order to improve the robustness of the methodology to temperature changes, the data from the different temperatures are organized and processed as it has

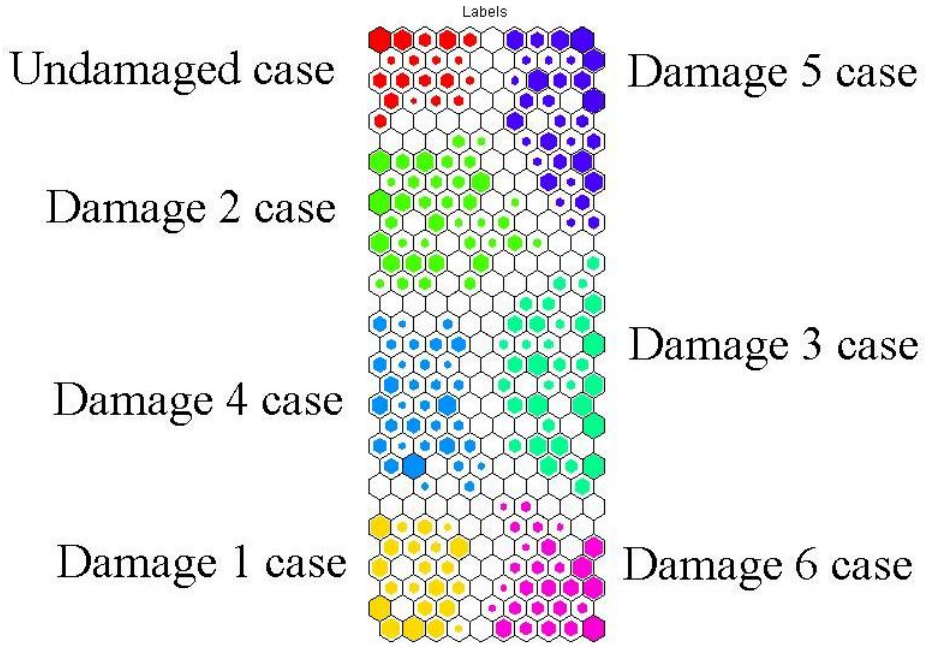


Figure 2.16: Classification of the different states at 50°C using 10 scores and the SPE-index using the cluster map.

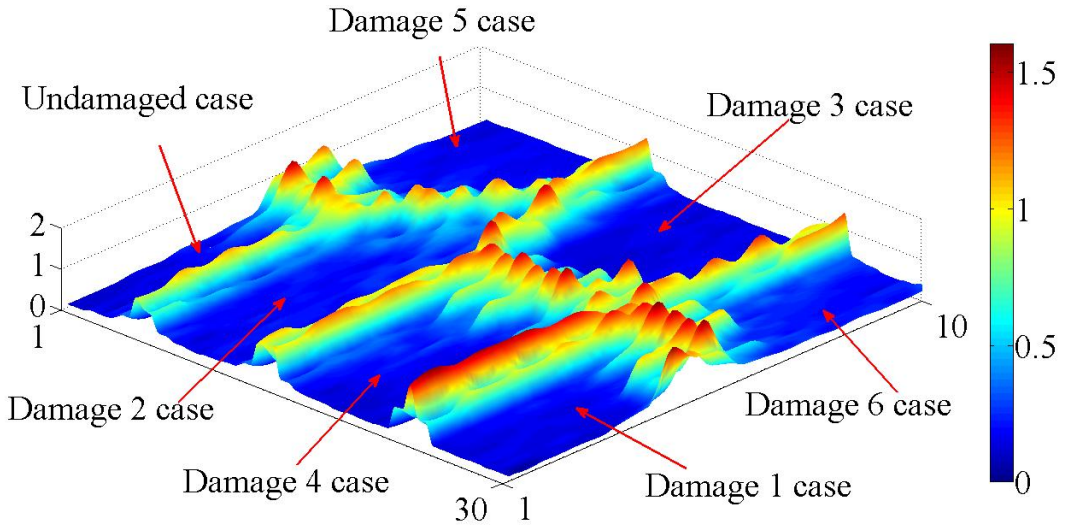


Figure 2.17: Classification of the different states at 50°C using 10 scores and the SPE-index using the U-matrix surface.

been explained in Section 2.4. In the previous cases, we have just used ten scores. However, the use of data from all the temperatures increases the number of scores that have to be considered. 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 30×30. 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.

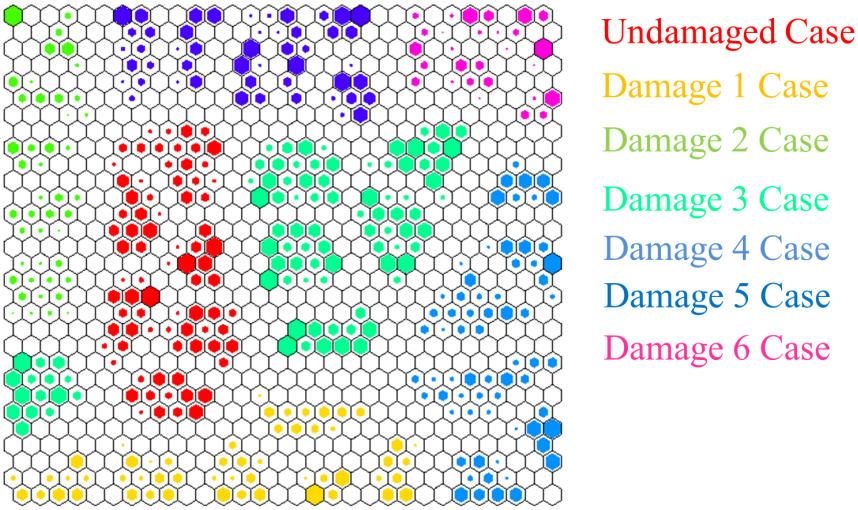


Figure 2.18: Classification of the different states using all temperatures, 150 scores, SPE-index, normalization type *histD* and map size 30 x 30 in the cluster map.

The results are presented in the Figures 2.18 and 2.19 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.

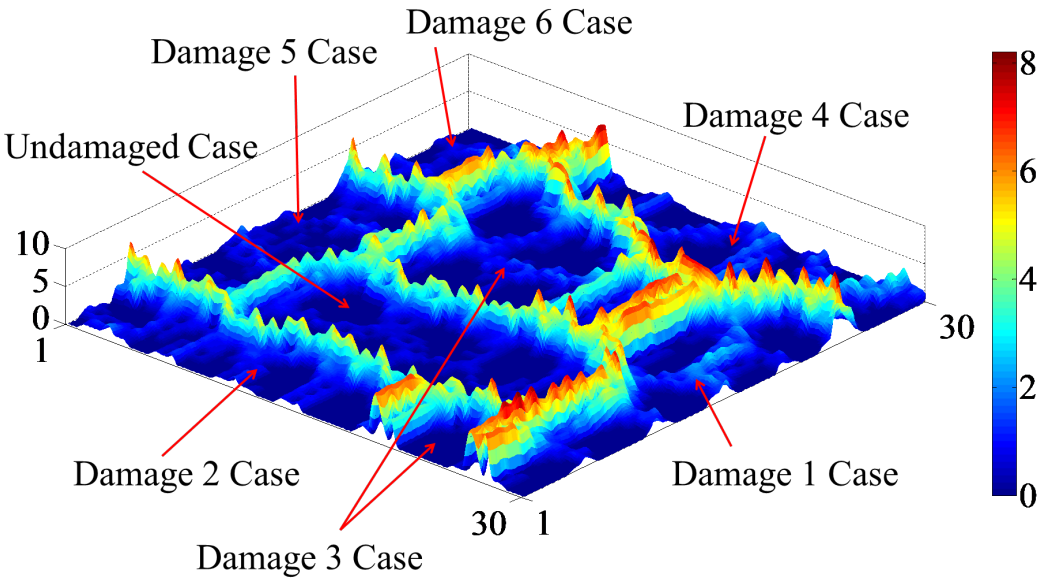


Figure 2.19: Classification of the different states using all temperatures, 150 scores, SPE-index, normalization type *histD* and map size 30 x 30 in the U-matrix surface.

2.5.2 Multi-layered composite plate

The evaluation of the damage detection and classification methodology using this specimen was performed by means of the following three studies:

- Damage detection and classification at different temperature levels.
- Damage detection and classification of different damages at different temperatures when the baseline is built using the data from all temperatures.
- Damage detection and classification of the same damage at different temperatures when the baseline is built using the data from all the temperatures.

To define the optimal set of parameters to configure the map in each study, several SOMs are again trained and validated. As a consequence, a defined number of scores and the SPE-index are used by each actuation phase to define the feature vector. To train the SOM, a normalization type *histD* is applied to the feature vector, and a hexagonal lattice with a flat sheet shape and a cluster size of 30×10 are used.

Damage detection and classification at different temperatures

Results from the first study are presented in Figures 2.20-2.25. These figures illustrate the results when the plate is exposed to temperatures of $35°\mathrm{C}$, $55°\mathrm{C}$ and $75°\mathrm{C}$. Similar results are obtained with the rest of the temperatures, but the figures are omitted for space reasons.
Figures 2.20 and 2.21 show the results obtained by means of the cluster map and the U-matrix surface at $35°\mathrm{C}$, respectively. These results show that six structural states are well identified. As it can be shown in the U-matrix surface (Figure 2.21), there is a clear separation between the undamaged state and the rest of damages at this temperature.

Figures 2.22 and 2.23 show the results when the multi-layered composite plate is exposed to a temperature of $55°\mathrm{C}$. As it has been discussed previously, the cluster map and the U-matrix surface present six clusters well defined. The biggest boundary can also be found between the undamaged state and the rest of damages. However, in this case, the boundary between the damage 4 and damage 2 is less evident.

Figures 2.24 and 2.25 present the results obtained when the temperature is $75°\mathrm{C}$. As well as the previous cases, the undamaged state is clearly separated from the damaged cases.

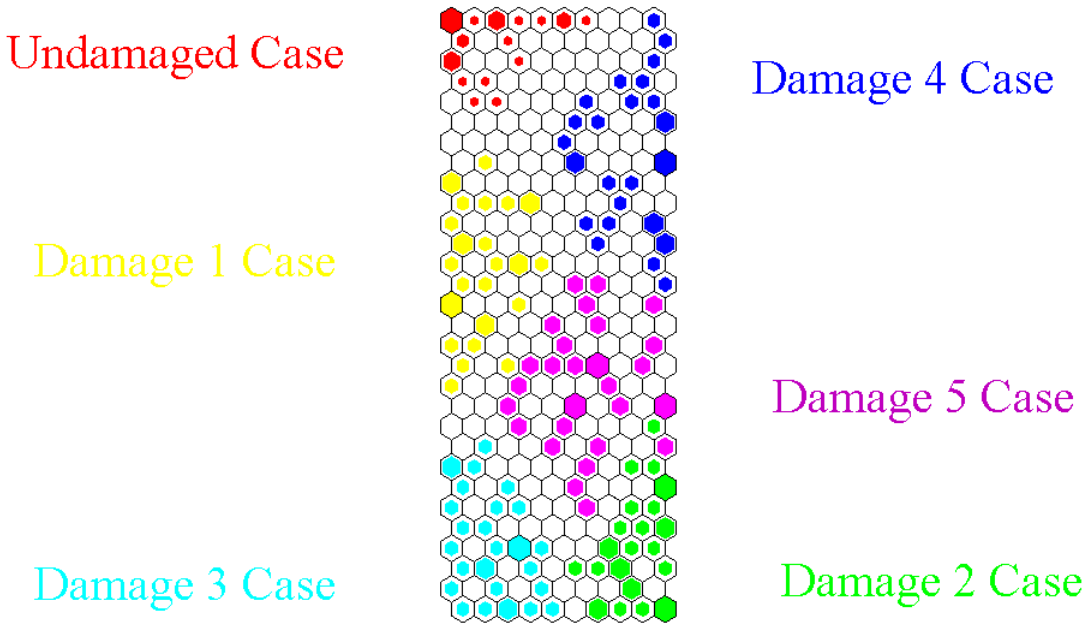


Figure 2.20: Classification of the different states at 35°C using 5 scores and the SPE-index.

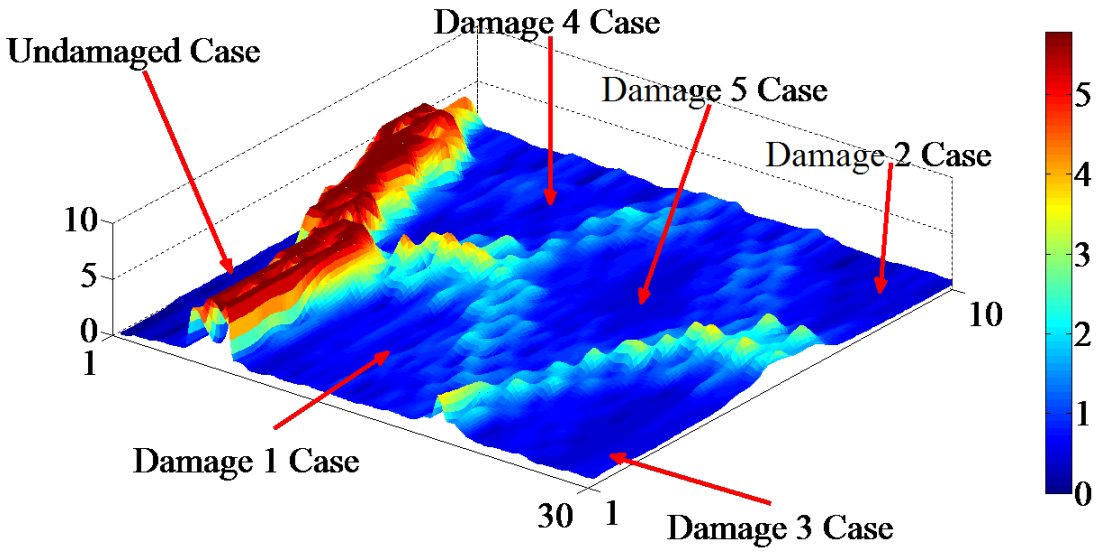


Figure 2.21: Classification of the different states at 35°C using 5 scores and the SPE-index.

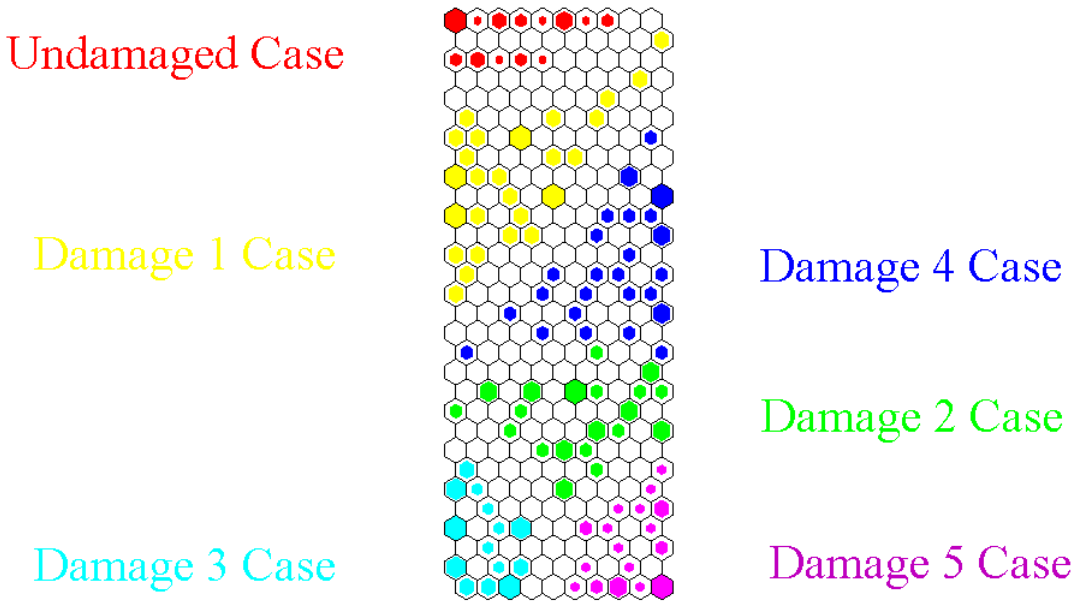


Figure 2.22: Classification of the different states at 55°C using 9 scores and the SPE-index using the cluster map.

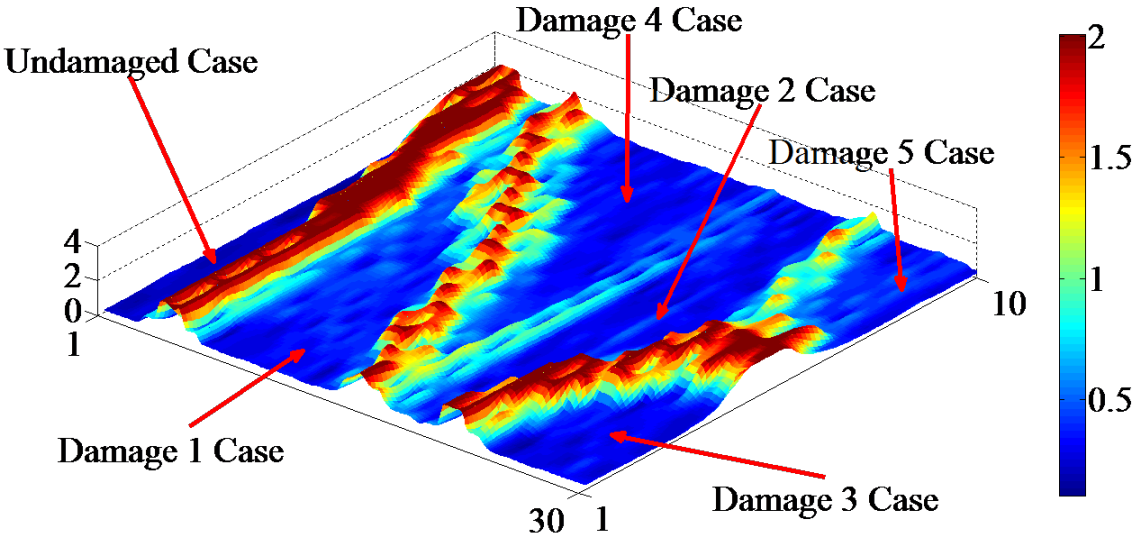


Figure 2.23: Classification of the different states at 55°C using 9 scores and the SPE-index using the U-matrix surface.

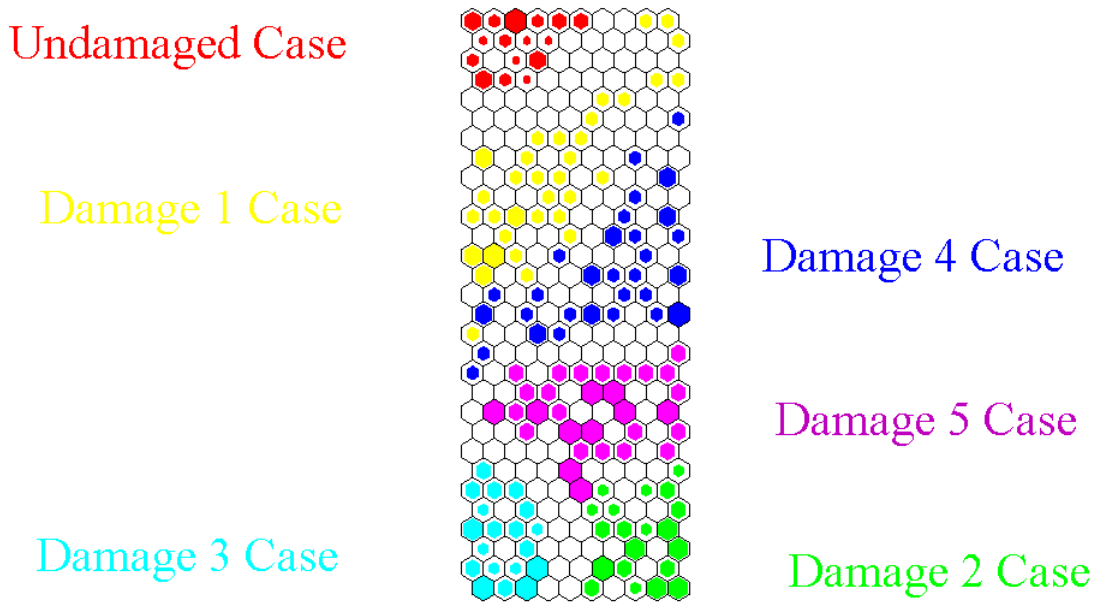


Figure 2.24: Classification of the different states at 75°C using 11 scores and the SPE-index in the cluster map.

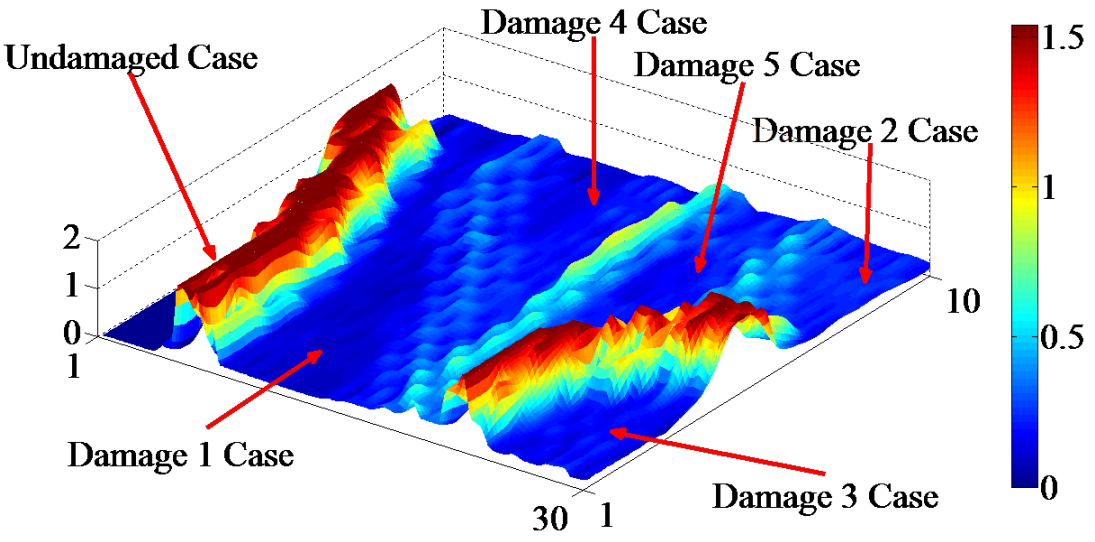


Figure 2.25: Classification of the different states at 75°C using 11 scores and the SPE-index in the U-matrix surface.

Damage detection and classification of different damages at different temperatures when the baseline is built using the data from all temperatures.

Results from the second study are depicted in Figures 2.26 and 2.27. 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 $35°C$; damage 2 is the damage 2 at $45°C$; damage 3 is the damage 3 at $55°C$; damage 4 is the damage 4 at $65°C$; and damage 5 is the damage 5 at $75°C$. The feature vector is formed by 30 scores and the SPE-index of each actuation phase. 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.

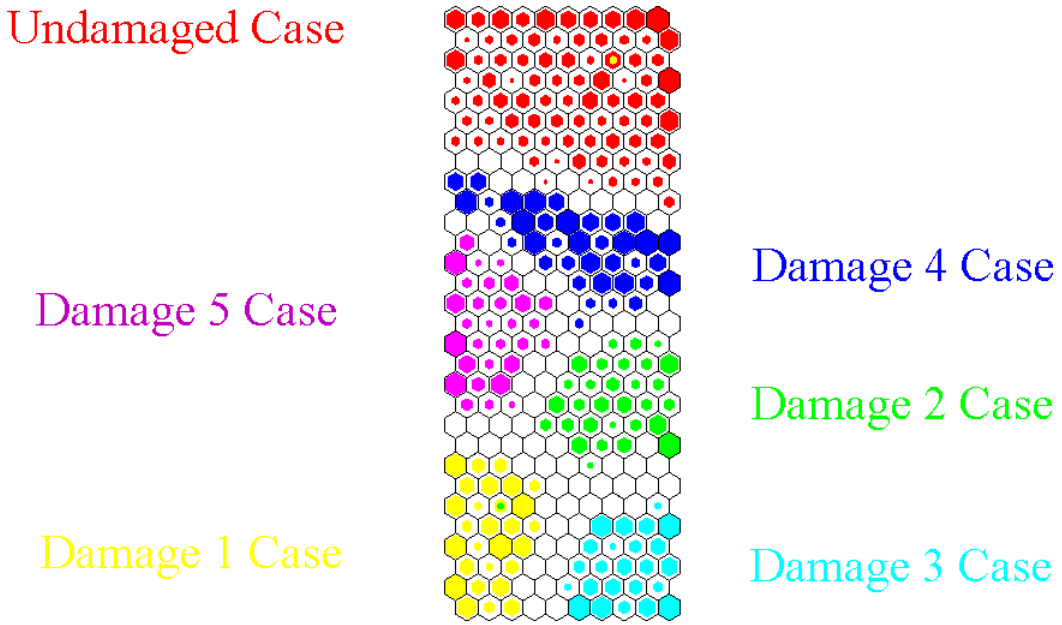


Figure 2.26: Classification using the Baseline with all temperatures and Damages at different temperatures and Positions.

Damage detection and classification of the same damage at different temperatures when the baseline is built using the data from all the temperatures.

Figures 2.28 and 2.29 show the results from the third study. In this case, the baseline is built using the data from all the temperatures. The objective is now to classify the same damage (damage 1) at different temperatures. The feature vector is composed by 30 scores and the SPE-index of each actuation phase. The results in the cluster map and the U-matrix surface show that it is possible to perform the damage detection because there is a clear distinction between the healthy state and the damaged states. Nevertheless, the clusters with damaged states cannot be clearly separated, and some outliers are present in the cluster map. The last

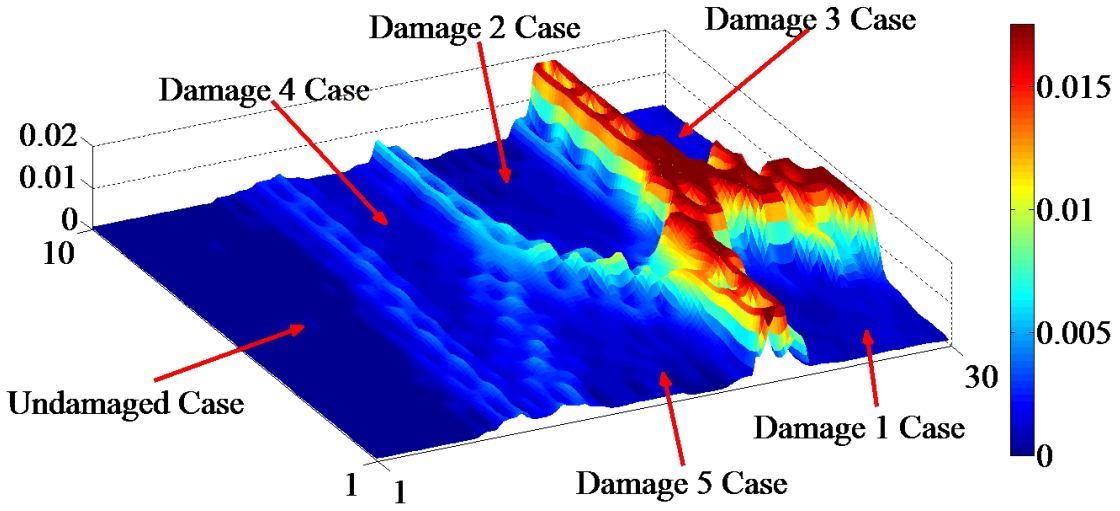


Figure 2.27: Classification using the Baseline with all temperatures and Damages at different temperatures and Positions.

remark can be validated using the U-matrix surface since its boundaries do not have a significant value. As a conclusion, a proper classification of all damages from this case is not possible.

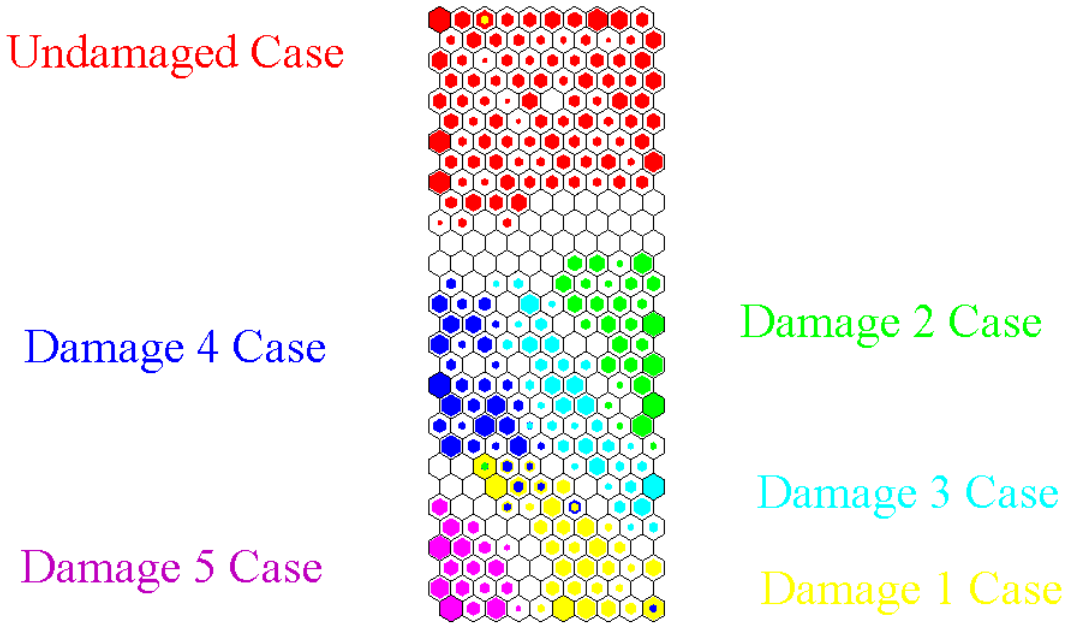


Figure 2.28: Classification using the Baseline with all temperatures and Damage 1 as damage at all temperatures in the cluster map.

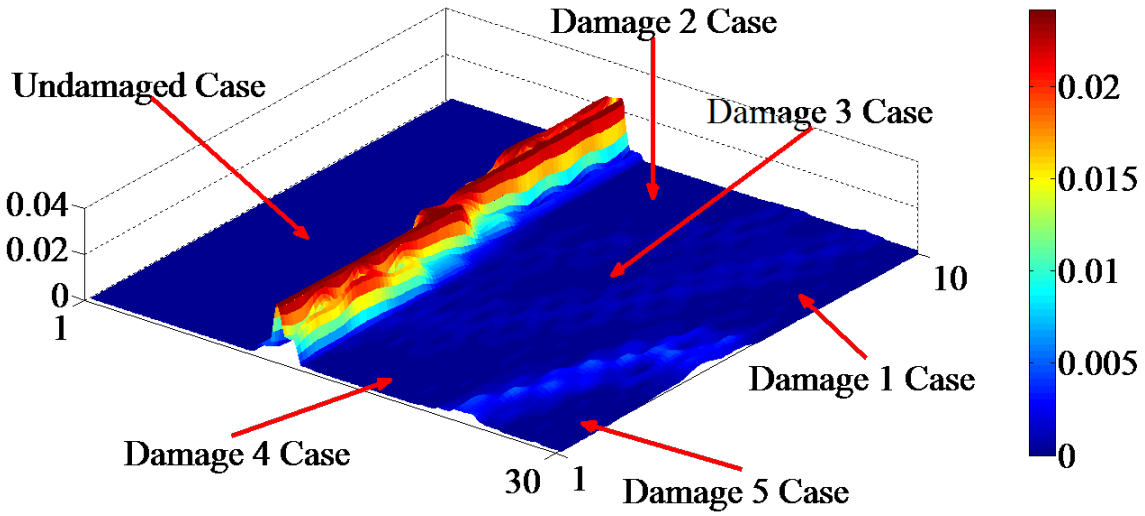


Figure 2.29: Classification using the Baseline with all temperatures and Damage 1 as damage at all temperatures in the U-matrix surface.

2.6 Conclusions

An automatic method was proposed for the analysis of ultrasonic signals for the purpose of damage detection and classification under temperature variations. 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. The first study indicated that the use of the approach enables to detect and classify damages in the structure and the breakage at different percentage in the sensor when the data is collected under the same temperature conditions. In the specific cases presented, the pre-processing with DWT, the use of the scores and the SPE-index by each phase and the data fusion by means of a Self-Organizing Map presented the best result in the detection and classification in both structures.

The use of the robust baselines as those presented in this paper allowed in all the cases the damage detection in spite of temperature changes and 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.

In the classification problem, the use of the robust baselines allowed the classification in most of the cases with the exception of the same damage at different temperatures in the simplified aircraft composite skin panel where the use of scores and the SPE-index do not allow the classification. In this specific case, additional studies need to be performed to obtain a better classification of all the damaged states.

Finally, it is significant to highlight that the presented approach solves the problem of evaluating all the phases to define the existence of damage. This is especially relevant for large structures instrumented with several PZT transducers. Now, the solution implies only the evaluation of the cluster map or the U-matrix surface obtained by data fusion.

2.7 Acknowledgment

The authors would like to express their gratitude to the Research School on Multi Modal Sensor Systems for Environmental Exploration (MOSES) and the Centre for Sensor Systems (ZESS) for sponsoring the research presented herein. The authors would like to further thank the support from the Spanish Ministry of Education through the grant “Movilidad de estudiantes en programas de doctorado con mención hacia la excelencia” which supported Maribel Anaya and Diego Tibaduiza’s placement in Siegen and to the Spanish Ministry of Science and Innovation through the coordinated project DPI2011-28033-C03-01. Finally, the authors also thank Mr. Gerhard Dietrich, Ms. Inka Buethe and Mr. R. Singh from the University of Siegen for they support during the execution of the experimental phase.

Bibliography

[1] S.W. Doebling, C.R. Farrar, and M.B. Prime. A summary review of vibration-based damage identification methods. *The Shock and Vibration Digest*, 30(2):91–105, 1998.

[2] C. R. Farrar, S. W. Doeblng, and D. A. Nix. Vibration-based structural damage identification. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences*, 359(1778):131–149, 2001.

[3] D.A. Tibaduiza. *Design and validation of a structural health monitoring system for aeronautical structures*. PhD thesis, Department of Applied Mathematics III, Universitat Politcnica de Catalunya, 2013.

[4] H. Sohn. Effects of environmental and operational variability on structural health monitoring. *Phil. Trans. R. Soc. A*, 365:539–560, 2007.

[5] C.-P. Fritzen, G. Mengelkamp, and A. Guemes. Elimination of temperature effects on damage detection within a smart structure concept. In *4th International Workshop on Structural Health Monitoring, Stanford, CA: DEStech Publications, Inc., pp. 1530-1538*, 2003.

[6] G. Konstantinidis, P.D. Wilcox, and B.W. Drinkwater. An investigation into the temperature stability of a guided wave structural health monitoring system using permanently attached sensors. *IEEE Sensors Journal*, 7(5):905–912, 2007.

[7] J. Moll, P. Kraemer, and C.-P. Fritzen. Compensation of environmental influences for damage detection using classification techniques. In *4th European Workshop on Structural Health Monitoring, Krakow, Poland: DEStech*, 2008.

[8] A. Raghavan and C.E.S. Cesnik. Effects of elevated temperature on guided-wave structural health monitoring. *Journal of Intelligent Material Systems and Structures*, 19:1383–1398, 2008.

[9] Y. Lu and J.E. Michaels. Discriminating damage from surface wetting via feature analysis for ultrasonic structural health monitoring systems. *Review of Progress in Quantitative Nondestructive Evaluation*, 27(1):1420–1427, 2008.

[10] A. Deraemaeker, E. Reynders, G. De Roeck, and J. Kullaa. Vibration-based structural health monitoring using output-only measurements under changing environment. *Mechanical Systems and Signal Processing*, 22:34–56, 2008.

[11] L.E. Mujica, J. Rodellar, J. Vehí, K. Worden, and W. Staszewski. Extended pca visualization of system damage features under environment.mental and operational variations. In *Smart Structures and Materials & Nondestructive Evaluation and Health Monitoring. San Diego, California, USA.*, March 2009.

[12] A.J. Croxford, J. Moll, P.D. Wilcox, and J.E. Michaels. Efficient temperature compensation strategies for guided wave structural health monitoring. *Ultrasonics*, 50(4-5):517–528, 2010.

[13] P. Kraemer, I. Buethe, and C.-P. Fritzen. Damage detection under changing operational and environmental conditions using self organizing maps. In *In Proceedings of SMART 11 , Saarbruecken, Germany*, 2011.

[14] M.A. Torres Arredondo and C.-P. Fritzen. Ultrasonic guided wave dispersive characteristics in composite structures under variable temperature and operational conditions. In *Proceedings of the 6th European Workshop in Structural Health Monitoring, EWSHM 2012, Dresden, Germany, pp. 261-268.*, 2012.

[15] J.C. Dodson and D.J. Inman. Thermal sensitivity of lamb waves for structural health monitoring applications. *Ultrasonics*, 53:677–685, 2013.

[16] K. Worden, W.J. Staszewski, and J.J. Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111, 2011.

[17] S.G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[18] R.R. Coifman and M.V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, 1992.

[19] D.E. Newland. *Random vibration, spectral and wavelet analysis*. New York, NY: Longman, Harlow and John Wiley, 1993.

[20] S. Mallat. *A Wavelet Tour of Signal Processing. 2 ed.* San Diego: Academic Press, 1997.

[21] P. Nomikos and J.F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):1361–1375, aug 1994.

[22] S. Wold, P. Geladi, K. Esbensen, and J. Ohman. Multiway principal component and pls analysis. *Journal of Chemometrics*, 1:41–56, 1987.

[23] G. Cherry and S.J. Qin. Multiblock principal component analysis based on a combined index for semiconductors fault detection and diagnosis. *IEEE Transactions on semiconductors manufacturing.*, Vol. 19, No. 2:159 – 172, 2006.

[24] M. Ruiz, G. Sin, X. Berjaga, J. Colprim, S. Puig, and J. Colomer. Multivariate principal component analysis and case-based reasoning for monitoring, fault detection and diagnosis in a wwtp. *Water science and technology*, , vol. 64, no8:1661–1667, 2011.

[25] I.T. Jolliffe. *Principal Component Analysis*. Springer, 2002.

[26] L.E. Mujica, J. Rodellar, A. Fernandez, and A. Güemes. Q-statistic and t^2-statistic pca-based for damage assessment in structures. *Structural Health Monitoring. An international Journal*, 10, No. 5:539–553, 2011.

[27] C. Alcala and S.J. Qin. Unified analysis of diagnosis methods for process monitoring. In *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes. Barcelona, Spain.*, 2009.

[28] G. Li, S.J. Qin, Y. Ji, and D. Zhou. Reconstruction based fault prognosis for continuous processes. In *7 th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Process. Barcelona-Spain.*, 2009.

[29] D.A. Tibaduiza, L.E. Mujica, and J. Rodellar. Comparison of several methods for damage localization using indices and contributions based on pca. *Journal of Physics: Conference Series*, 305(1), 2011.

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

[31] T. Kohonen. The self-organizing maps. In *Proceedings of the IEEE, VOL. 78, NO.9*, 1990.

[32] M.A. Torres, I. Buethe, D.A. Tibaduiza, J. Rodellar, and C.P. Fritzen. Damage detection and classification in pipework using acousto-ultrasonics and probabilistic non-linear modelling. In *Workshop on Civil Structural Health Monitoring (CSHM-4), Berlin-Germany*, 2012.

[33] M.A. Torres, D.A. Tibaduiza, L.E. Mujica, J. Rodellar, and C.P. Fritzen. Damage assessment in a stiffened composite panel using non-linear data-driven modelling and ultrasonic guided waves. In *4 th International Symposium on NDT in Aerospace. Ausburg, Germany*, 2012.

[34] N. Kasabov, D. Nikovski, and E. Peev. Speech recognition based on kohonen self-organizing feature maps and hybrid connectionist systems. In *In Proceedings of First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems*, 1993.

[35] J.A. Walter and K.J. Schulten. Implementation of self-organizing neural networks for visuo-motor control of an industrial robot. *IEEE Transactions on Neural Networks*, Vol.4, No.1:86–96, 1993.

[36] Y. Rongfeng, W. Gaolin, Y. Yong, and X. Dianguo. The self-organizing map and fuzzy control for sensorless induction motor speed control. *Natural Computation (ICNC), 2010 Sixth International Conference on*, vol.3:1406–1409, 2010.

[37] B. Ghimire, S. Sinanovic, H. Haas, and G. Auer. Self-organised interference mitigation in wireless networks using busy bursts. In *Applied Sciences in Biomedical and Communication Technologies. ISABEL 2009. 2nd International Symposium on*, 2009.

[38] T. Kohonen and T. Honkela. Essentials of the self-organizing map. *Neural Networks*, 37:52–65, 2013.

[39] V. K. Pachghare, P. Kulkarni, and D.M Nikam. Intrusion detection system using self organizing maps. In *Intelligent Agent Multi-Agent Systems, 2009. IAMA 2009. International Conference on*, 2009.

[40] J. Vesanto and E.Alhoniemi. Clustering of the self-organizing map. In *Neural Networks IEEE Transactions on , vol.11, no.3, pp.586,600*, 2000.

[41] G. Kalogiannakis and D.V. Hemelrijck. Classification of wavelet decomposed ae signals based on parameter-less self-organized mappig. *International Journal of Materials and Product Technology*, 41(1-4):89–104, 2011.

[42] M.M.R. Taha, A. Noureldin, J.L. Lucero, and T.J. Baca. Wavelet transform for structural health monitoring: A compendium of uses and features. *Structural Health Monitoring*, 5(3):267–295, 2006.

[43] J. Westerhuis, T. Kourti, and J. MacGregor. Comparing alternative approaches for multivariate statistical analysis of batch process data. *Journal of Chemometrics*, 13:397–413, 1999.

[44] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas. *SOM Toolbox for Matlab 5*. Helsinki University of Technology: Helsinki, Finland, 2000.

Chapter 3

A bio-inspired methodology based on an artificial immune system for damage detection in structural health monitoring

Authors: Maribel Anaya, Diego A. Tibaduiza, Francesc Pozo
Published in:
Journal: Shock and Vibration (Special Issue on Structural Dynamical Monitoring and Fault Diagnosis), 2015, article ID 648097, 15 pages, 2015, doi:10.1155/2015/648097
Publisher: Hindawi
Online ISSN: 1875-9203
Print ISSN: 1070-9622
2014 Impact Factor: 0.722 (Q3, 20/31 in the area of Acoustics)

This chapter is a true copy of the original paper published in the journal where the only changes are performed to fit the page setup.

A bio-inspired methodology based on an artificial immune system for damage detection in SHM

Among all the aspects that are linked to a structural health monitoring (SHM) system, algorithms, strategies or methods for damage detection are currently playing an important role for improving the operational reliability of critical structures in several industrial sectors. This paper introduces a bio-inspired strategy for the detection of structural changes using an artificial immune system (AIS) an a statistical data-driven modeling approach by means of a distributed piezoelectric active sensor network at different actuation phases. Damage detection and classification of structural changes using ultrasonic signals is traditionally performed using methods based on the time of flight. The approach followed in this paper is a data-based approach based on AIS, where sensor data-fusion, feature extraction and pattern recognition are evaluated. One of the key advantages of the proposed methodology is that the need to develop and validate a mathematical model is eliminated. The proposed methodology is applied, tested and validated with data collected from two sections of an aircraft skin panel. The results show that the presented methodology is able to accurately detect damages.

Keywords: Artificial immune systems; principal component analysis; damage indices.

3.1 Introduction

Structural health monitoring (SHM) is a discipline that makes use of sensors permanently attached to a structure together with different software analysis developments in order to detect damages and assess the proper performance of structures. An SHM system traditionally includes continuous monitoring, data processing algorithms and pattern recognition techniques for a robust analysis. Different methodologies have been developed in the last years in the field of SHM. However, with the use of bio-inspired algorithms, promising results have been obtained, mainly due to its adaptive, distributed and autonomous features.

This work presents a damage detection methodology that is mainly based on an artificial immune system (AIS) as a pattern recognition technique and affinity plots to discriminate the different structural states of the structure. This methodology is applied to the collected data by a piezoelectric system. The artificial immune system has been proposed and used in several applications. However, in structural health monitoring, this methodology is relatively new. A brief state-of-the-art in chronological order is presented in the next lines highlighting the most representative works with respect to AIS in structural health monitoring.

The use of non-destructive inspection methods (NDT) has proved to be a very useful tool for damage detection tasks. However, in some situations where it is impossible to manually inspect a structure –as in the inspection of large-scale structures–, the use of automated methods present significant advantages. Some of these advantages can be summarized as follows: (i) continuous monitoring, since the sensors are permanently attached to the structure; (ii) early damage

detection; (iii) damage identification, among others. In this sense, structural health monitoring (SHM) extends the limits of the NDT methods by including the use of data processing algorithms, pattern recognition and continuous monitoring because the sensors are permanently attached to the structure. This is one the reasons why the development of improvements in data processing algorithms is a current demand. The contribution of the present work is the development of a methodology for data-driven damage classification using a bio-inspired algorithm, which is applied to data that comes from a piezoelectric system. More precisely, this work uses an artificial immune system that allows the use of this methodology as a pattern recognition approach. The use of artificial immune systems (AIS) is relatively new in the literature and, compared with the application of other approaches in SHM, there are still a reduced number of works. In the next lines we briefly compile in chronological order the most representative works in the use of AIS.

In 2003 Costa *et al.* [1] developed three module algorithms called T-module, B-module and D-module. These algorithms are based on immunologic principles to detect anomalous situations in a squirrel-cage motor induction. The T-module distinguishes between self and non-self conditions, the B-module analyzes the occurrence of both cells (self and non-self) and finally the D-module is similar to a T-module but with a reduced space. In this work, the normal operation condition of the machine (self) is represented by the frequency spectrum, that can include or not include harmonics.

In 2007, Da Silva *et al.* [2] presented a damage detection algorithm applying an auto-regressive model and auto-regressive model with exogenous input (AR-ARX). This algorithm is based on the structural vibration response measurements and the residual error as damage sensitive index. Data compression is used by means of principal component analysis (PCA) and the fuzzy c-means clustering method is used to quantify the damage sensitive index. In this paper, the authors used a benchmark problem with several damage patterns to test the algorithm. As the main result, a structural diagnosis was obtained with high correlation with the actual state of the structure. Later on 2008, Da Silva *et al.* [3] developed a strategy to perform structural health monitoring. This strategy included three different phases: (i) the use of principal component analysis to reduce the dimensionality of the time series data; (ii) the design of an autoregressive-moving-average (ARMA) model using data from the healthy structure under several environmental and operational conditions; and finally (iii) the identification of the state of the structure through a fuzzy clustering approach. In this paper, the authors compared the performance of two fuzzy algorithms, fuzzy c-means (FCM) and Gustafson-Kessel (GK) algorithms. The proposed strategy was applied to data from a benchmark structure at Los Alamos National Laboratory. The work showed that the GK algorithm outperforms the FCM algorithm, because the first algorithm considers an adaptive distance norm and allows clusters with several geometrical distributions.

Also in 2008, Zhang *et al.* [4] used a clonal selection algorithm to solve a combinatorial optimization problem called *sensor optimization*. This problem consists in choosing an appropriate distribution of a set of sensors in a structure to detect impacts. To test the algorithm, the authors used a composite plate instrumented with 17 lead zirconium titanate (PZT) transducers.

De Moura *et al.* [5] presented a fuzzy based meta-model to detect damages in a flat structure under corrosion conditions. This work considers data obtained from a SHM approach based on electro-mechanic impedance. Chen [6], in 2010, applied an agent-based artificial immune system for adaptative damage detection. In the approach, a group of agents is used as immune cells (B-cells) patrolling over a distributed sensor network installed in the structure. The damage

diagnosis is based on the analysis of structural dynamic response data. Each mobile agent inspects the structure using agent based cooperation protocols. In 2010, Tan *et al.* [7] presented a damage detection algorithm based on fuzzy clustering and support vector machines (SVM). In this work, as a first step, the wavelet packet transform is used to decompose the accelerator data from the structure and extract the energy of each wavelet component. Consequently, this energy is used as a damage index. In further steps, damages are classified by means of fuzzy clustering. As a final phase, damages are identified using a vector machine. The numerical example illustrated in this work shows that the proposed method is able to identify the damage from the spatial truss structure. In 2011, Chen and Zang [8] presented an algorithm based on immune network theory and hierarchical clustering algorithms. Chilengue *et al.* [9] presented and artificial immune system (AIS) approach to detect and diagnose failures in the stator and rotor circuits of an induction machine. In the approach, the dynamic of the machine is compared before and after the fault condition. Similarly, the alpha-beta ($\alpha\beta\gamma$) transformation (also known as the Clarke transformation) was applied to the stator current to obtain a characteristic pattern of the machine that is finally applied to the pattern recognition algorithm.

In 2012, Zhou *et al.* [10], inspired by Chen's work, developed a damage classifier in structures based on the immune principle of clonal selection. Using evolution algorithms and the immune learning, a high quality memory cell is generated in the classifier generated able to recognize several damage patterns. In 2012, Xiao [11] developed a structural health monitoring and fault diagnosis system based on artificial immune system. In this approach, the antigen represents the structural state (health or damage), whereas the antibody represents database information to identify a damage pattern. In this work, the feature space is formed by natural frequencies and modal shapes collected by simulation of the structure in free vibration and seismic response. Quite recently, Liu *et al.* [12], in 2014, proposed a structural damage detection method using semi-supervised fuzzy c-means clustering method, wavelet packet decomposition and data fusion. This method is applied to detect damage in a four-level benchmark model. The data that was used include 11 damage pattern and 9 samples per damage. The method uses a Daubechies wavelet filter and 6 decompositions levels. According to the results, the method can achieve a reasonable detection performance. Huang *et al.* [13], in 2014, proposed an automatic methodology to know the status of a machine. The introduced method includes a semi-supervised fuzzy-based method to detect the faults or anomalies in the machine and to classify the unknown faults. The authors described two steps for the learning procedure: (i) a fuzzy c-means clustering to get candidates of labels (fuzzy centers); and (ii) a label matching by filtering out of the unreasonable labels candidates. The proposed method is validated in a roller bearing test top diagnose the state of the machine.

Compared with the works previously reviewed, the methodology described on the current work presents a new point of view, since this uses an artificial immune system (AIS) and some damage indices to define feature vectors which represents the structure under different conditions by allowing that the damage detection process can be understood as a pattern recognition approach. More precisely, damage detection and classification using ultrasonic signals have been traditionally performed using methods based on the time of flight. The approach followed in this paper, that complements and completes the initial work by Anaya *et al.* [14], is rather different because is a data-based approach based on AIS (artificial immune system), where sensor data-fusion, feature extraction and pattern recognition are evaluated. A clear major advantage of the methodology is that the development and validation of a mathematical model is not needed. Additionally, and in contrast to standard Lamb waves-based methods, there is no necessity of

directly analyzing the complex time-domain traces containing overlapping, multi-modal and frequency dispersive wave propagation which distorts the signals and makes difficult their analysis. However, using the proposed methodology, it is not possible to provide a multi-damage detection able to identify several occurring damages independently unless the model baselines are built with the structural responses that have interacted with previously detected and existing damage.

This paper is organized as follows. Section 3.2 describes the theoretical background that includes basic concepts about the methods and elements used in the methodology. Section 4.3 includes the damage detection methodology followed by the description about the experimental setup and the experimental results in Section 3.4. Finally, some conclusions are drawn in section 3.5.

3.2 General framework

The current work is based on data-driven analysis. This means that the damage detection will be developed by analyzing and interpreting the data collected in several experiments from the structures under to diagnose. To perform this analysis, a bio-inspired methodology based on features extraction for pattern recognition is developed. For the sake of clarity, basic concepts and fundamentals about the methods that will be used are presented in the following subsections.

3.2.1 Bio-inspired systems

The adaptation of the different living beings of the planet in harsh environments and the development of skills to solve the inherent problems in the interaction with the world of nature has resulted in the evolution of the species in order to survive and avoid their extinction. Some examples are the communication abilities, the reasoning, the physical structures design or the response of the body to external agents, among others [15].

Taking advantage of the fact that nature provides robust and efficient solutions to many different problems, more and more researchers on different areas work in the development of biologically inspired hardware and algorithms. The inspiration process is called *biomimetic* or *bioinspired* and aims to apply the developments in the field of biology to the engineering developments [16].

3.2.2 Natural immune systems

The human immune system (HIS) is a complex and robust defense mechanism composed by a large network of specialized cells, tissues and organs. The system further includes an elevated number of sensors and a high processing capability. The human immune system has proved its effectiveness in the detection of foreign elements by protecting the organism against disease. The principal skills of the human immune system are:

- To discriminate between its own cells (self) and foreign cells (non-self).
- To recognize different invaders (called antigens) in order to ensure the protection of the body.

- To learn from specific antigens and adapt to them in order to improve the immune response to this kind of invader.

In general, when a foreign particle wants to gain access to the organism, it has to break several defense levels provided by the immune system that protects the organism. The idea of several defense levels is illustrated in Figure 3.1. These levels can be summarized as follows [17]:

- *External barriers*. These are the first and the major line of defense into the human body. This level can include elements such as the skin, the mucus secreted by the membranes, the tears, the saliva and the urine. All of these elements present different physiological conditions that are harmful to the antigens, as the temperature or the pH level, among others. The response of these barriers is equal for any foreign invader [18].

- *Innate immune system*. This barrier refers to the defense mechanisms that are activated immediately or within a short lapse of time of an antigen's arrival in the body. The innate immune system operates when the first barrier has been broken This system, in opposition to the adaptive immune system, is not adaptive[17].

- *Adaptive immune system*. This is the last defense level and reacts to the stimulus of foreign cells or antigens that evade both the external barriers and the innate immune defense [17]. Adaptive immunity creates some sort of memory that leads to an improved response to future encounters with this antigen.

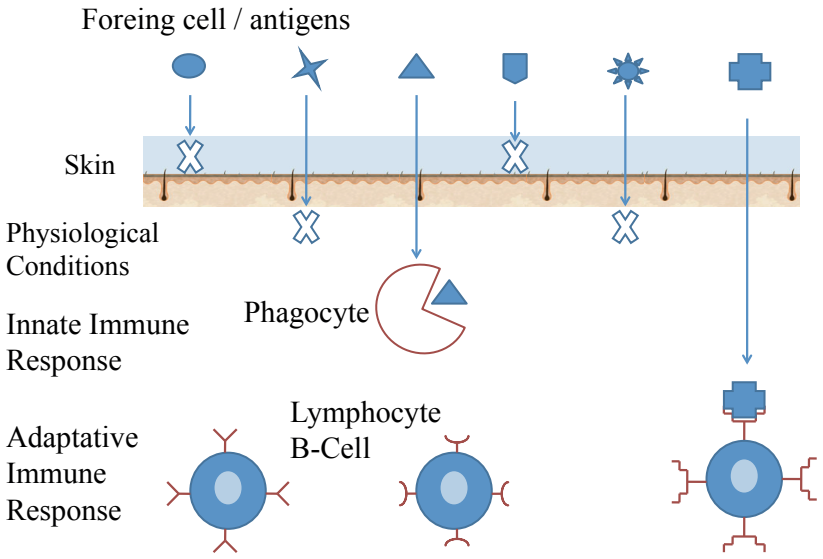


Figure 3.1: Schematic representation of a system's defense barriers.

With respect to different type of cells, the immune system includes cells born in the bone marrow that are usually called *white blood cells*, *leukocytes* or *leucocytes* [19]. Among the white blood cells, it is possible to highlight the T-cells and the B-cells. On one hand, the T-cells are so called since their maturation takes place in the thymus.Besides, this kind of cells have high mobility and can also be found in the blood and the lymph [20]. One can distinguish three types of T-cells:

- the T-helper cells, involved in the activation of B-cells;
- the T-killer cells that destroy the invaders; and finally
- the T-suppressor cells that avoid the allergic reactions [21].

On the other hand, the B-cells produce and secrete a special protein called antibody, which recognizes and binds the antigen. The responsibility of each B-cell is the production of a specific antibody. This protein is then used for signaling other cells what elements have to be removed from the body [20]. When the antigen passes over the first barrier of the immune system, the HIS performs the following steps to eliminate the invader [20]:

1. The specialized cells of the immune system, called *antigen presenting cells* (APCs) (macrophages, for instance). These cells active the immune response by ingesting the antigen and dividing it into simple substances known as antigenic peptides.
2. These peptides are joined to the molecules called *major histocompatibility complex* (MHC), inside of the macrophage, and the result passes to the immune cell surface.
3. The T-cells have receptor molecules able to identify and recognize different combinations of MHC-peptide. When the receptor molecule recognizes the combination, the T-cell is activated and sends a chemical signal to other immune cells.
4. The B-cells are activated by chemical signals and they initiate the recognition of the antigen in the bloodstream. This process is performed by the receptor molecules in the B-cells.
5. The mission of the B-cells –when they are activated– is to secrete antibodies to bind the antigens they find, and to neutralize and eliminate them from the body.

The T- and B-cells that have recognized the antigen proliferate and, some of them, become memory cells. These memory cells remain in the immune system to eliminate the same antigen –in the future– in a more effective manner [15, 20].

Three immunological principles are used in artificial immune systems [11, 15, 20]:

- *Immune network theory*. This theory was first introduced by Niels Jerne in 1974 and describes how the immune memory is built by means of the dynamic behaviour of the immune system cells. These cells can recognize by themselves, detect invaders, as well as interconnect between them to stabilize the network [17].
- *The negative selection*. The negative selection is a process that allows the identification and eradication of the cells that react to the own body cells. This ensures a convenient operation of the immune system since it is able to distinguish between foreign molecules and self-molecules thus avoiding autoimmune diseases. This process is similar to the maturation of T-cells carried out in the thymus [15].
- *The clonal selection*. This is a mechanism of the adaptive immune responses in which the cells of the system are adapted to identify an invader element [20]. Antibodies that are able to recognize or identify an antigen can proliferate. Those antibodies unable to recognize the antigens are eliminated. The new cells are clones of their parents and they

Table 3.1: Analogy between the biological immune system and artificial immune system [11].

Biological Immune System	Artificial Immune System in SHM
antibodies	a detector of a specific pattern
antigens	structural health or damage condition
matured antibodies	database or information system for damage detection
recognition of antigens	identification of health and damage condition
process of mutation	training procedure
immune memory	memory cells

are subjected to an adaptation process by mutation. From the new antibody set, the cells with the greatest affinity with respect to the primary antigen are selected as memory cells therefore excluding the rest.

3.2.3 Artificial immune systems

Artificial immune systems (AIS) are an adaptive and bio-inspired computational systems based on the processes and performance of the human immune system (HIS) and its properties –diversity, error tolerance, dynamic learning, adaptation, distributed computation and self-monitoring– [22, 23]. Nowadays, these computational systems are used in several research areas such as pattern recognition [16], optimization [20, 24], computer security [25], among others [26]. Table 3.1 presents the analogy between the natural and artificial immune systems applied to the field of structural health monitoring.

In the implementation of an artificial immune system, it is fundamental to bear in mind two important aspects:

- To define the role of the antigen (ag) and the antibody (ab) in the context of the application. Both are represented or coded in the same way. This representation is generally given by a vector of binary or real numbers [21].

- To define the mechanism that measures the degree of correspondence between an antigen and an antibody. This measure is usually related to the distance between them [15]. If both an antigen and an antibody are represented by $L-$dimensional arrays,

$$ab \in \mathbb{R}^L,$$
$$ag \in \mathbb{R}^L,$$

the distance d between them can be computed using, for instance, the Euclidean distance (related to the $2-$norm) or the so-called Manhattan distance (related to the $1-$norm) as in equations (3.1) and (3.2), respectively [19]:

$$d(ab, ag) = \|ab - ag\|_2 = \sqrt{\sum_{i=1}^{L}(ab_i - ag_i)^2} \tag{3.1}$$

$$d(ab, ag) = \|ab - ag\|_1 = \sum_{i=1}^{L}|ab_i - ag_i| \tag{3.2}$$

Finally, there exists the adaptation process of the molecules in the artificial immune system. This adaptation allows to include the dynamic of the system, for instance, the antibodies excitation, cloning all the excited antibodies and the interconnection between them. All these elements are adapted from the three immunologic principles previously introduced.

3.2.4 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a classical method used in applied multivariate statistical analysis with the goal of dimensionality reduction and, more precisely, feature extraction and data reduction. It was developed by Karl Pearson in 1901 and integrated to the mathematical statistics in 1933 by Harold Hotelling [27]. The general idea in the use of PCA is to find a smaller set of variables with less redundancy [28]. To find these variables, the analysis includes the transformation of the current coordinate space to a new space to re-express the original data trying to filtering the noise and redundancies. These redundancies are measured by means of the correlation between the variables.

Matrix unfolding

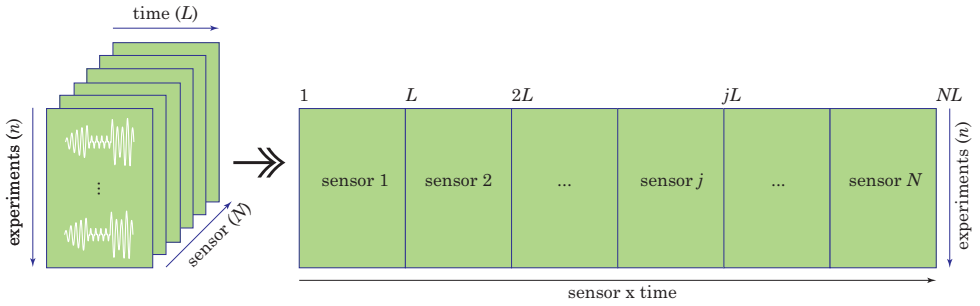


Figure 3.2: The collected data arranged in a three dimensional matrix is unfolded in a two dimensional matrix [29].

The application of PCA starts –for each actuation phase– with the collected data arranged in a three dimensional matrix $n \times L \times N$. The matrix is subsequently unfolded –as illustrated in Figure 3.2–in a two dimensional $n \times (N \cdot L)$ matrix as follows:

$$\mathbf{X} = \begin{pmatrix} x_{11}^1 & x_{12}^1 & \cdots & x_{1L}^1 & x_{11}^2 & \cdots & x_{1L}^2 & \cdots & x_{11}^N & \cdots & x_{1L}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1}^1 & x_{i2}^1 & \cdots & x_{iL}^1 & x_{i1}^2 & \cdots & x_{iL}^2 & \cdots & x_{i1}^N & \cdots & x_{iL}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1}^1 & x_{n2}^1 & \cdots & x_{nL}^1 & x_{n1}^2 & \cdots & x_{nL}^2 & \cdots & x_{n1}^N & \cdots & x_{nL}^N \end{pmatrix} \tag{3.3}$$

Matrix $\mathbf{X} \in \mathcal{M}_{n\times(N\cdot L)}(\mathbb{R})$ –where $\mathcal{M}_{n\times(N\cdot L)}(\mathbb{R})$ is the vector space of $n \times (N \cdot L)$ matrices over $\mathbb{R}$– contains data from N sensors at L discretization instants and n experimental trials [30]. Consequently, each row vector $x_i^T = \mathbf{X}(i,:) \in \mathbb{R}^{N\cdot L}$, $i = 1, \ldots, n$ represents, for a specific experimental trial, the measurements from all the sensors. Equivalently, each column vector

$\mathbf{X}(:,j) \in \mathbb{R}^n,\ j = 1,\ldots,N \cdot L$ represents measurements from one sensor in the whole set of experimental trials.

In other words, the objective is to find a linear transformation orthogonal matrix $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ that will be used to transform the original data matrix $\mathbf{X}$ according to the following matrix multiplication

$$\mathbf{T} = \mathbf{XP} \in \mathcal{M}_{n\times(N\cdot L)}(\mathbb{R}). \tag{3.4}$$

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 $\mathbf{PP}^{\mathrm{T}} = \mathbf{I}$, 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

$$\mathbf{X} = \mathbf{TP}^T.$$

Group scaling

Since the data in matrix $\mathbf{X}$ come from experimental trials and could have different magnitudes and scales, it is necessary to apply a preprocessing step to scale the data using the mean of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor [30].

More precisely, for $k = 1, 2, \ldots, N$ we define

$$\mu_j^k = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^k,\ j = 1,\ldots,L, \tag{3.5}$$

$$\mu^k = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} x_{ij}^k, \tag{3.6}$$

$$\sigma^k = \sqrt{\frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}(x_{ij}^k - \mu^k)^2}, \tag{3.7}$$

where μ_j^k is the mean of the n measures of sensor k at the time instant j; μ^k is the mean of all the measures of sensor k; and σ^k is the standard deviation of all the measures of sensor k. Therefore, the elements x_{ij}^k of matrix $\mathbf{X}$ are scaled to define a new matrix $\check{\mathbf{X}}$ as

$$\begin{aligned} \check{x}_{ij}^k &:= \frac{x_{ij}^k - \mu_j^k}{\sigma^k}, \\ & i = 1,\ldots,n,\ j = 1,\ldots,L,\ k = 1,\ldots,N. \end{aligned} \tag{3.8}$$

When the data are normalized using equation (3.8), the scaling procedure is called *variable scaling* or *group scaling* [30]. According to former studies of the authors [29, 31, 32], group scaling presents a better performance than other kind of normalizations. The reason is that group scaling considers changes between sensors and does not process them independently. Further discussion on this issue can be found in [30, 33].

For simplicity, and throughout the rest of the paper, the scaled matrix $\check{\mathbf{X}}$ is renamed as simply $\mathbf{X}$. The mean of each column vector in the scaled matrix $\mathbf{X}$ can be computed as

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\check{x}_{ij}^{k} &= \frac{1}{n}\sum_{i=1}^{n}\frac{x_{ij}^{k}-\mu_{j}^{k}}{\sigma^{k}} = \frac{1}{n\sigma^{k}}\sum_{i=1}^{n}\left(x_{ij}^{k}-\mu_{j}^{k}\right)\\ &= \frac{1}{n\sigma^{k}}\left(\sum_{i=1}^{n}x_{ij}^{k}-n\mu_{j}^{k}\right)\\ &= \frac{1}{n\sigma^{k}}\left(n\mu_{j}^{k}-n\mu_{j}^{k}\right)=0.\end{aligned}$$

Since the scaled matrix $\mathbf{X}$ is a *mean-centered* matrix, it is possible to calculate the covariance matrix as follows:

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}^{T}\mathbf{X} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R}). \tag{3.9}$$

The covariance matrix $\mathbf{C_X}$ is a $(N\cdot L)\times(N\cdot L)$ symmetric 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 follows:

$$\mathbf{C_X P} = \mathbf{P}\Lambda \tag{3.10}$$

where the columns of $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvectors of $\mathbf{C_X}$. The diagonal terms of matrix $\Lambda \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvalues $\lambda_i,\ i=1,\ldots,N\cdot L$, of $\mathbf{C_X}$ whereas the off-diagonal terms are zero, that is,

$$\begin{aligned}&\Lambda_{ii} = \lambda_i,\ i=1,\ldots,N\cdot L\\ &\Lambda_{ij} = 0,\ i,j=1,\ldots,N\cdot L,\ i\neq j\end{aligned}$$

The eigenvectors $p_j,\ j=1,\ldots,N\cdot L$, 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* or the *loading vectors* 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

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

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

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

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 $m-$dimensional space to get a data matrix $\hat{\mathbf{X}}$ as follows:

$$\hat{\mathbf{X}} = \hat{\mathbf{T}}\hat{\mathbf{P}}^{T} \in \mathcal{M}_{n\times m}(\mathbb{R}). \tag{3.11}$$

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

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}, \tag{3.12}$$

or, equivalenty,

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T + \mathbf{E}. \tag{3.13}$$

The residual error matrix $\mathbf{E}$ describes the variability not represented by the data matrix $\hat{\mathbf{X}}$, and can be also expressed as

$$\mathbf{E} = \mathbf{X}(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T). \tag{3.14}$$

Even though 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 [34].

3.2.5 Damage detection indices based on PCA

Several damage detection indices based on PCA have been proposed and applied with excellent results in pattern recognition applications. In particular, two damage indices are commonly used: (i) the Q index (also known as SPE, *square prediction error*) and (ii) the Hotelling's T^2 index.

The Q index of the ith experimental trial x_i^T measures the magnitude of the vector $\tilde{x}_i^T := \tilde{\mathbf{X}}(i,:)$, that is, the events that are not explained by the model of principal components [35], and it is defined as follows:

$$Q_i = \tilde{\mathbf{X}}(i,:)\tilde{\mathbf{X}}(i,:)^T = x_i^T(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T)x_i. \tag{3.15}$$

The T^2 index of the ith experimental trial x_i^T is the weighted norm of the projected vector $\hat{t}_i^T := \hat{\mathbf{T}}(i,:) = x_i^T\hat{\mathbf{P}}$, that is, a measure of the variation of each sample within the PCA model and it is defined as follows:

$$T_i^2 = \sum_{j=1}^{\ell} \frac{\hat{t}_{i,j}^2}{\lambda_j} = \hat{t}_i^T \Lambda^{-1} \hat{t}_i = x_i^T(\hat{\mathbf{P}}\Lambda^{-1}\hat{\mathbf{P}}^T)x_i \tag{3.16}$$

3.3 Damage detection methodology

The damaged detection methodology that we present in this paper involves an active piezoelectric system to inspect the structure. This active system consists of several piezoelectric transducers (lead zirconium titanate, PZT) distributed on different positions of the structure and working both as actuators or sensors in different actuation phases. Each PZT is able to produce a mechanical vibration if some electrical excitation is applied (actuator mode). Besides, the PZT are able to detect time varying mechanical response data (sensor mode). In each phase of the experimental stage, just one PZT is used a the actuator (exciting the structure). Then, the propagated signal through the structure is collected by using the rest of PZT, which are used as

sensors. This procedure is repeated in as many actuation phases as the number of PZT on the structure.

To determine the presence of damage in the structure, the data from each actuation phase will be used in the proposed artificial immune system. The proposed methodology is performed in three steps: (i) data pre-processing and feature extraction; (ii) training process and; (iii) testing. More precisely, in the first step the collected data is organized, pre-processed and dimensionally reduced –by means principal component analysis– to obtain relevant information. The damage indices in equations (3.15)-(3.16) are used to define the feature vectors. The training step includes the evolution of the data to *generate* good representatives for each pattern, damage or structural condition. A good accuracy in the damage detection using AIS depends on a good training. Finally, the testing step includes new data to evaluate the training step and the knowledge of the current state of the structure.

3.3.1 Data pre-processing, PCA modeling and feature extraction

For each different phase (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 $\mathbf{X}$ is defined and arranged as in equation (3.3) in Section 3.2.4 and scaled as stated in Section 3.2.4. PCA modeling basically consists of computing the projection matrix $\mathbf{P}$ for each phase as in equation (3.5). Matrix $\mathbf{P}$, renamed $\mathbf{P}_{\text{model}}$, provides an improved and dimensionally limited representation of the original data $\mathbf{X}$. The number of principal components retained at each different phase account for at least 90% of the cumulative variance.

Subsequently, the data from different structural states are projected into each PCA model in order to obtain the scores and calculate the damage detection indices T^2 and Q as in Sections 3.2.4 and 3.2.5. This way, for each experiment, a two dimensional *feature vector*

$$f^i = (T_i^2, Q_i)^T \in \mathbb{R}^2,\ i = 1, \ldots, \nu \tag{3.17}$$

is defined, where ν is the total number of experiments. The feature vector could include more components, as the scores, for instance. Several test were then performed in this sense with the combination of scores and damage indices. However, the results indicated that the single use of T^2 and Q lead to the best results. One of the reasons about the use of the damages indices can be found in [35]. In this paper, Rodellar *et al.* showed that the use of scores is not sufficient for damage detection when two scores do not account for a high cumulative variance. This result implies that it is necessary to use another type of measurement or statistic to obtain an accurate discrimination of the presence of damage in a structure.

To keep the affinity values within the range $[0, 1]$, the norm of the feature vectors $f_i,\ i = 1, \ldots, \nu$ is normalized to the unit circle. The normalization process uses the maximum norm of the feature vectors, that is,

$$M := \max_{i=1,\ldots,\nu} \|f^i\|,$$

where

$$\|f^i\| = \sqrt{\left(f_1^i\right)^2 + \left(f_2^i\right)^2},$$

and therefore the normalized feature vector f^i_{norm} of $f^i = (T_i^2, Q_i)^T$ is as follows

$$f^i_{\text{norm}} = \left(\frac{T_i^2}{M}, \frac{Q_i}{M} \right).$$

Since all the feature vectors are located within a unit circle, the Euclidean distance between any feature vectors is less than or equal to 2. The *healthy data set* (HDS) is defined as

$$HDS = \bigcup_{i=1}^{\nu} \{f^i\}. \tag{3.18}$$

3.3.2 Training step

This step can be modified according to different goals. For instance, in the most basic case in damage identification –the detection–, the training only needs to consider the feature vectors that come from data of the healthy structure. However, in a more complex analysis –the classification, for instance–, the training process must include the feature vectors of data coming from the structure in different and known structural states. The steps to perform the training in the basic case are summarized as follows:

Figure 3.3: Random selection of the antibody (AB_{training}) and antigen (AG_{training}) training sets.

- Randomly select $2k \in \mathbb{N},\ 2k < \nu$, feature vectors. The remaining $\nu - 2k$ feature vectors will be used in the testing process. This set of $2k$ feature vectors is divided in two subsets of the same size k, the *antibody training set* (AB_{training}) and the *antigen training set* (AG_{training}). This step is represented in Figure 3.3.
- Compute the affinity between the antibodies and antigens of the AB_{training} and AG_{training} sets, respectively. The affinity between an antibody and an antigen is defined as

$$\mathit{aff}(ab, ag) := 1 - \frac{1}{2} d(ab, ag),$$

where $d(ab, ag)$ is the distance –defined in equation (3.1)– between the feature vectors of ab and ag, respectively. Since the Euclidean distance between any feature vectors is less than or equal to 2, their affinity lies within the range $[0, 1]$.

- Evolve the antibodies. The evolution of the antibodies is performed when these are stimulated by an invading antigen invader and it consists on the mutation of the antibody. The mutation is performed by mutating the feature vectors of the cloned antibodies as shown in equation (3.19)

$$ab_{\text{evolved}} = ab + MV \cdot \phi, \tag{3.19}$$

where ab_{evolved} is the mutated antibody and MV represents the mutation value –a value used to indicate the mutation degree of the feature vector of an antibody–. In the present implementation, the mutation value is defined as in equation (3.20)

$$MV = 1 - CV, \tag{3.20}$$

where CV is the *clonal value* –a value that measures the response of an artificial B-cell to an antigen– and is equal to the affinity between the antibody and the stimulating antigen. The vector

$$\phi = (\phi_1, \phi_2)^T \in \mathbb{R}^2$$

in equation (3.19) is a randomly generated vector. Each element $\phi_i,\ i = 1, 2$ of the random vector is a normally distributed random variable with mean zero and standard deviation $\sigma = 0.5$.

The mutated antibody feature vectors must lie within the unit circle. Therefore, the norm of the feature vector for each mutated antibody is immediately checked after the mutation according to the following procedure:

  - if $\|ab_{\text{evolved}}\| \leq 1$, then no normalization is performed.
  - if $\|ab_{\text{evolved}}\| > 1$, then

$$ab_{\text{evolved}} = (\|ab\| + \mathcal{U} \cdot (1 - \|ab\|)) \cdot \frac{ab_{\text{evolved}}}{\|ab_{\text{evolved}}\|},$$

    where $\mathcal{U}$ is a uniform random function with a value within the range of $[0, 1]$.

The norm of the mutated antibody is greater than the norm of the original antibody and less than 1.

The *clonal rate* (CR) is an integer value used to control the number of antibody clones allowed. The number of clones (NC) is defined in equation (3.21)

$$NC = \lfloor CR \cdot CV \rfloor, \tag{3.21}$$

where $\lfloor \cdot \rfloor$ is the floor function. In this paper the value of CR is 8.

The highest affinity antibody is chosen as the candidate memory cell for possible updating of memory cell set.

- Define the threshold. A threshold T_h is defined in order to update the memory cell set to improve the representation quality of memory cells for the healthy state of the structure. This threshold is defined as a weighted affinity of the two elements in the healthy data set (HDS) in equation (3.18) with the maximum Euclidean distance. That is,

$$\Delta = \max_{i,j=1,\dots,n} \|f^i - f^j\| \tag{3.22}$$

$$\delta = \frac{7}{25}\Delta \tag{3.23}$$

$$T_h = 1 - \frac{1}{2}\delta \tag{3.24}$$

Then a comparison between the candidate memory cell and all the elements in the healthy data set (HDS) is performed through the affinity. If the affinities are greater than or equal to the threshold, the candidate memory cell becomes memory cell of the healthy state of the structure. Otherwise, the candidate memory cell is eliminated. The main outcome of this step is the *memory cell set* of the healthy state ($MCSH$) of the structure. This algorithmic training process is represented in Figure 3.4.

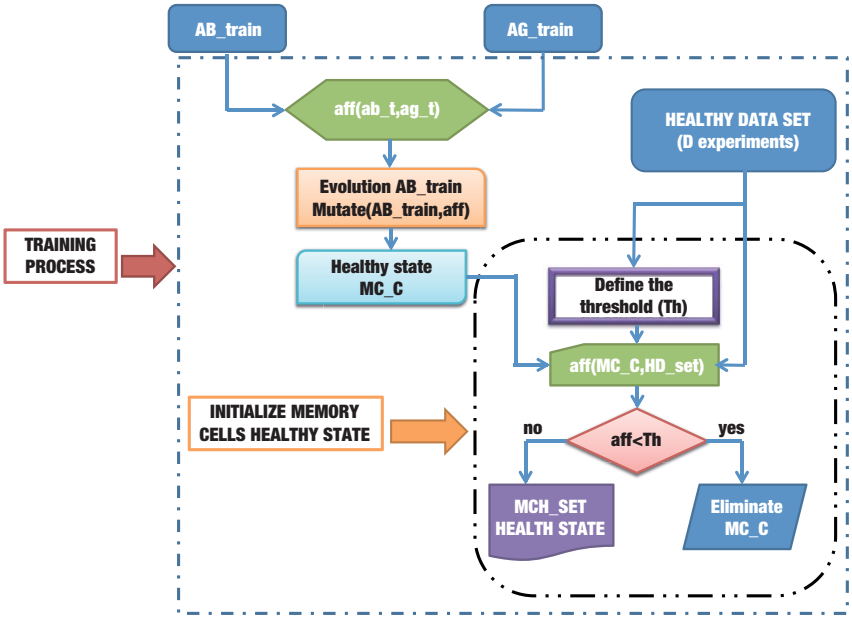


Figure 3.4: Training process in an artificial immune system applied (AIS) to structural health monitoring (SHM).

3.3.3 Testing step

The damage detection algorithm is finally illustrated in Figure 5. The damage detection is based on the affinity values between the elements in the memory cell set of the healthy state ($MCSH$) –acting as antibodies– and the data coming from the structure to test (TD, test data) –acting as antigens–. A *detection threshold* (DT_h) is defined in equation (3.25) for this purpose,

$$DT_h = \min_{\substack{ab \in MCSH \\ i \in \{1,\dots,\nu\}}} \mathit{aff}(ab, f^i), \tag{3.25}$$

that is, the minimum affinity between the elements in the memory cell set of healthy state (MCSH) and the elements in the healthy data set (HDS).

When the affinity is less than the threshold DT_h, we say that the data has been collected from a damaged state of the structure. Otherwise, the data comes from an undamaged structure.

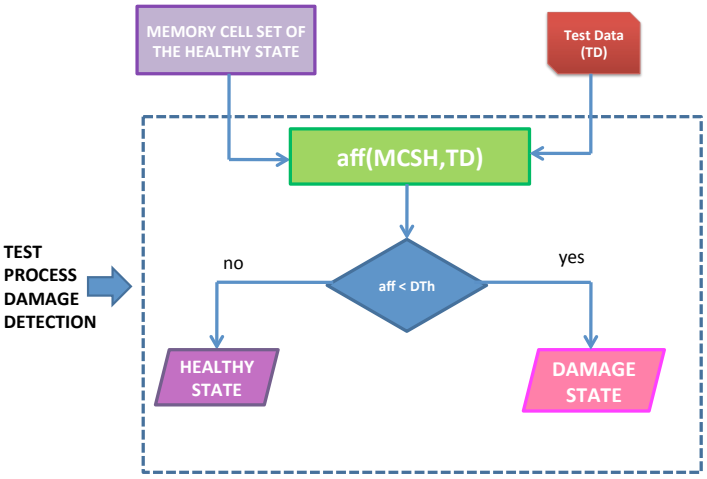


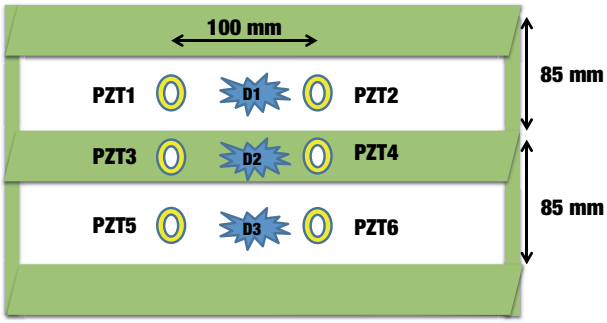
Figure 3.5: Damage detection process

3.4 Experimental setup and experimental results

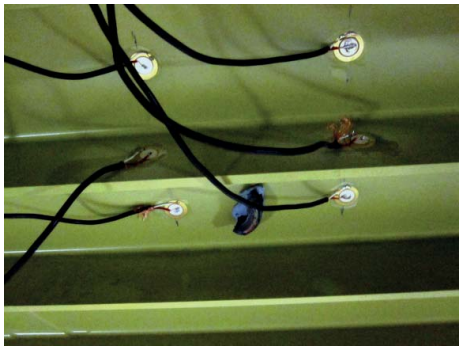
3.4.1 Experimental setup

To test the proposed methodology, data from an aircraft skin panel is used. The structure is divided in small sections by means of stringers and ribs as shown in Figure 3.6. To validate the proposed methodology, two sections of this structure were used. The dimensions of each section and the damage description are depicted in Figure 3.7. These sections were instrumented with 6 PZT transducers: two in the upper section; two in the lower section; and two in the rib. The transducers dimensions are: 26 mm diameter and 0.4 mm thickness.

Figure 3.6: Aircraft skin panel



(a) a.



(b) b.

Figure 3.7: Damage description.

3.4.2 Experimental results

As said in Section 3.3.1, the experiments are performed in 6 independent phases: (i) piezoelectric transducer 1 (PZT1) is configured as actuator and the rest of PZTs as sensors; (ii) PZT2 as actuator; (iii) PZT3 as actuator; (iv) PZT4 as actuator; (v) PZT5 as actuator; and (vi) PZT6 as actuator.

To apply the proposed methodology, and for each phase, the collected data is arranged in a matrix as in equation (3.3) in Section 3.2.4. With this *unfolded* data, the PCA model $\mathbf{P}$ is built as explained in Sections 3.2.4 and 3.3.1 using data from the healthy structure. In Figure 3.8 the amount of variance accounted for by each principal component is illustrated, for phases 1, 3, 5 and 6.

For each actuator phase, the number of principal components adopted varies since the principal components retained must account for at least 90% of the cumulative variance. Although there is not an accurate criterion to state a percentage of cumulative variance to be retained for a good representation, a high percentage can ensure that most of the variability is incorporated into the statistical model.

Figure 3.9 show the projections onto the two first principal components of several experiments that come from the undamaged and damaged structure under consideration. It can be clearly observed that no separation of damaged/undamaged can be determined using the scatter plot. These are then two motivating depictions in the sense that with the proposed methodology we will be able to both detect damage in the structure as well as classify it.

Damage detection

After the baseline modeling, the data coming from the structure to be diagnosed is projected onto the PCA model. Then, for each experiment, the feature vector in equation (3.17) formed by the two damage indices T^2 and Q is defined.

The ability of the proposed method to detect damages in the structure is illustrated in Figures 3.10 to 3.15. In these figures, the affinity of a memory cell from the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose is depicted. The 25 first experiments correspond to data that come from the undamaged structure, while the remainder 75 experiments come from the damaged structure. More precisely, experiments 25 to 50 correspond to damage 1 ($D1$), experiments 51 to 75 to damage 2 ($D2$) and experiments

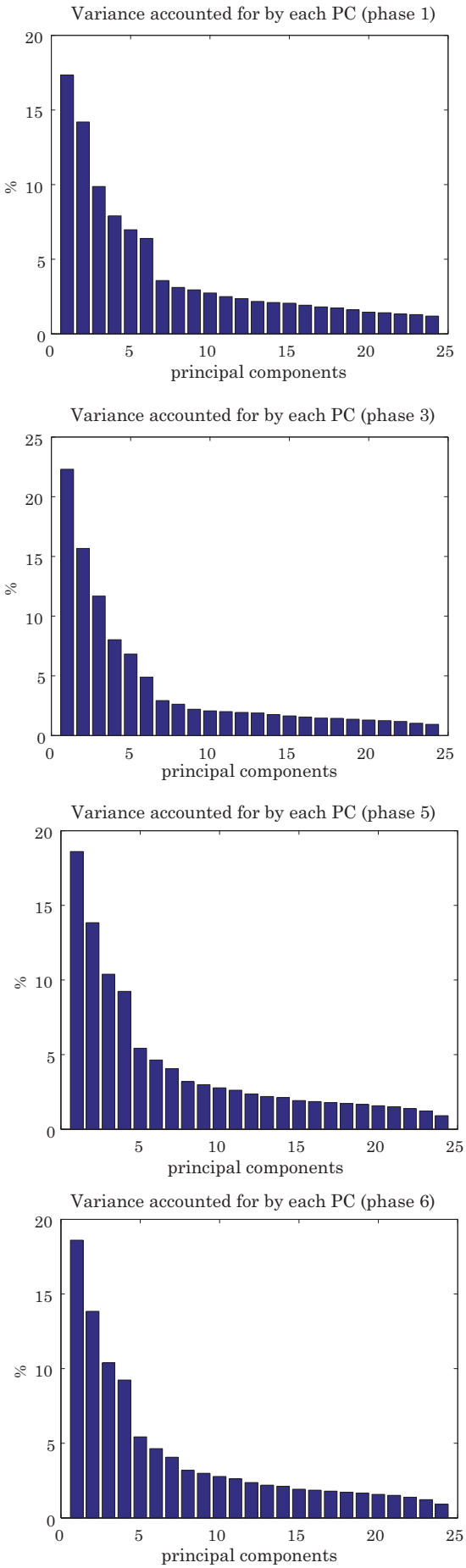


Figure 3.8: Amount of variance accounted for by each principal component, for phases 1, 3, 5 and 6.

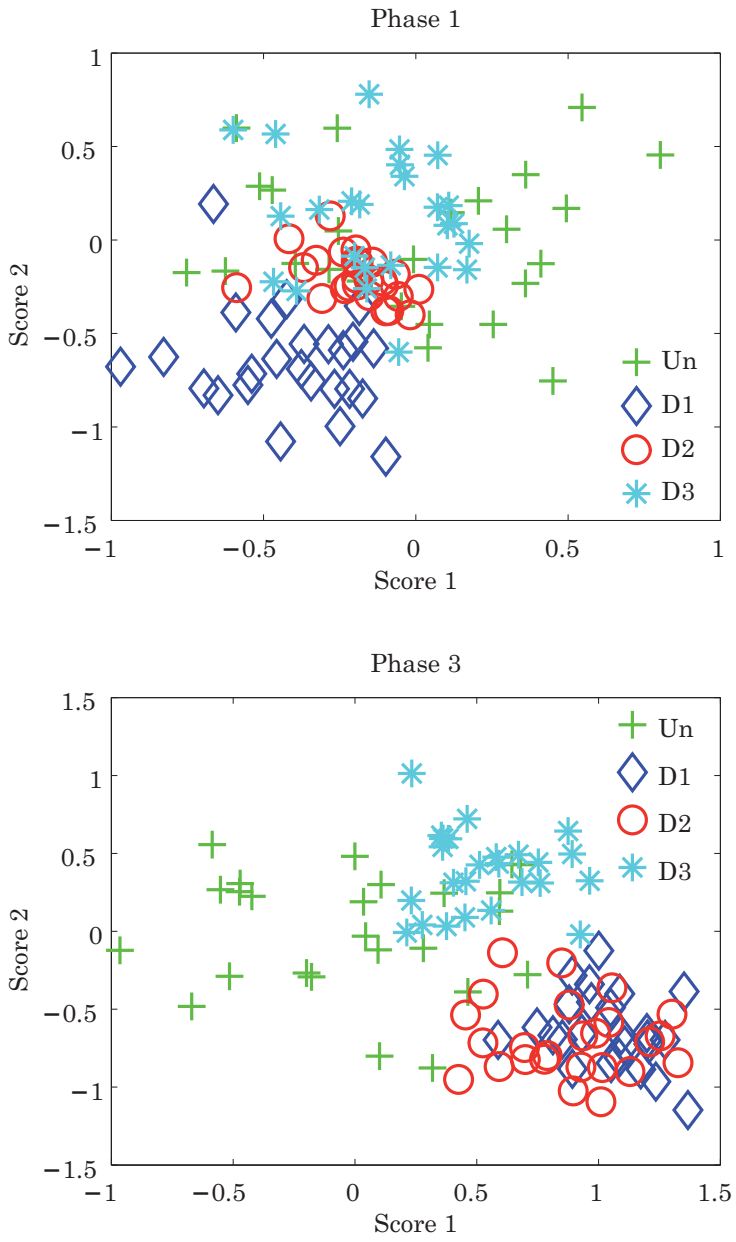


Figure 3.9: Projections onto the two first principal components of several experiments in actuator phases 1 (up) and 3 (down).

76 to 100 to damage 3 ($D3$). The purple solid horizontal line delimits the detection threshold (DT_h). It can be clearly observed that experiments with an affinity value less than DT_h –from the damaged structure– are correctly defined as 'damaged'. Similarly, experiments with an affinity value greater than or equal to DT_h –from the healthy structure– are correctly defined as 'healthy'.

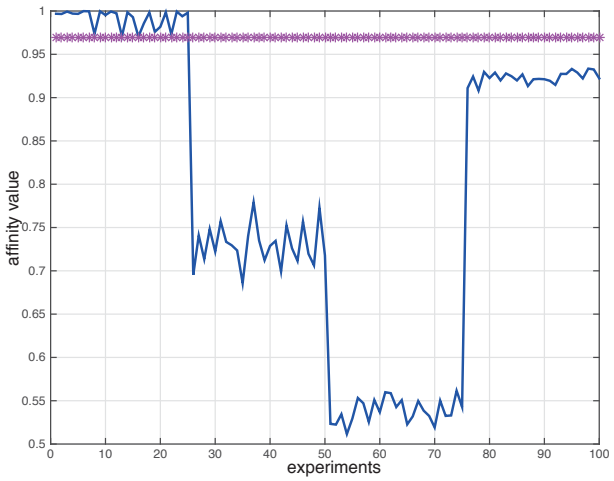


Figure 3.10: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 1).

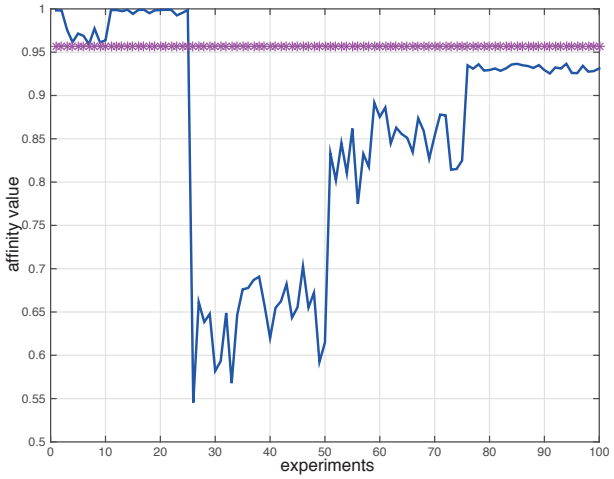


Figure 3.11: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 2).

Actuation phases 1 to 5 show that it is possible to distinguish the healthy and unhealthy states, in addition it can be observed that different affinity values represent the differences between the group of data which indicate that it is possible to determine the presence of three damages. In contrast to the affinity in the rest of the actuation phases, Figure 3.15 is showing that it is possible to detect abnormal situations, however it is not possible to determine the different structural states.

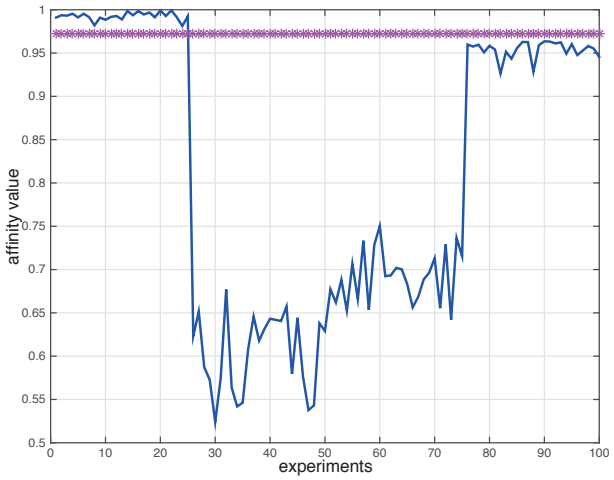


Figure 3.12: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 3).

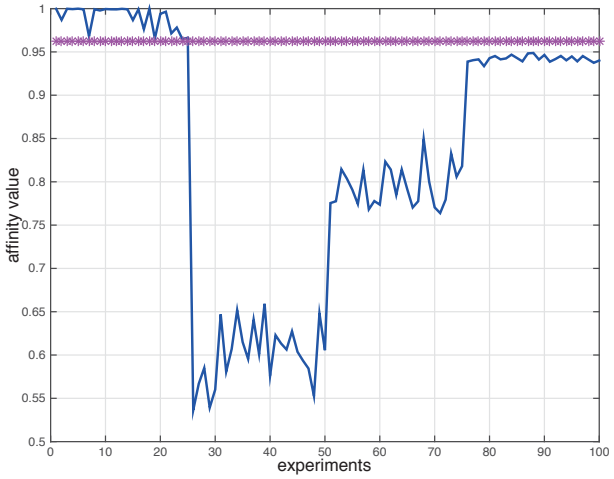


Figure 3.13: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 4).

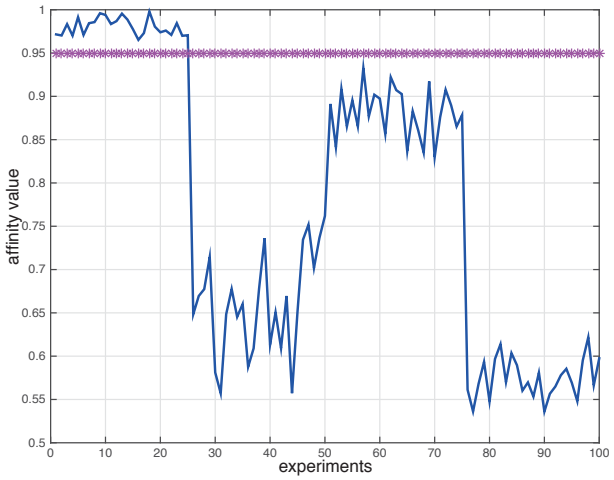


Figure 3.14: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 5).

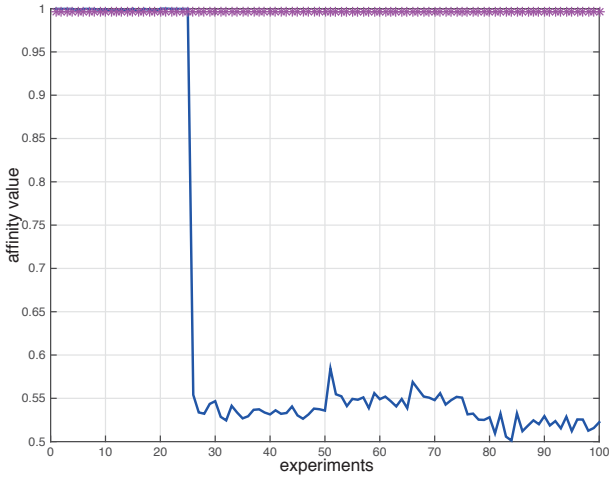


Figure 3.15: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 6).

As it was shown, results from each actuation phase showed different affinity values to the different structural states, this is because the actuators are distributed by the structure in different positions and to different distance to the damages.

3.5 Concluding remarks

In this paper, a new methodology to detect structural changes has been introduced. The methodology includes the use of an artificial immune system (AIS) and the notion of affinity for the damage detection.

One of the advantages of the methodology is the fact that to develop and validate a model is not needed. Additionally, and in contrast to standard Lamb waves-based methods, there is no need to directly analyze the complex time-domain traces containing overlapping, multimodal and frequency dispersive wave propagation that distorts the signals and difficult the analysis. Results shown that different actuation phases present different results

The proposed methodology has been applied to data coming from two sections of an aircraft skin panel. The results indicate that the proposed methodology is able to accurately detect damages by means of the analysis of the affinity values. However, within the proposed methodology, it is not possible to provide a multidamage classification able to identify several simultaneous damages. To ensure the proper performance of the methodology, a study of the effect of changing environmental and operational conditions need to be considered, which is considered as a future work. The methodology can be improved by applying data fusion in order to obtain a only plot with the information from the actuation phases. In this sense, the use of SOM or fuzzy clustering will allow the data fusion and to estimate more information from the data.

Acknowledgements

This work is supported by CICYT (Spanish Ministry of Economy and Competitiveness) through grants DPI2011-28033-C03-01 and DPI2014-58427-C2-1-R and by Universidad Santo Tomás through grants FODEIN 2014, project code 047 and FODEIN 2015, project code 8410083003-036. The authors would also like to express their gratitude to the professor Alfredo Güemes for providing the structure used on this paper and for his suggestions in the experimental setup and data acquisition process.

Bibliography

[1] P. J. Costa Branco, J. A. Dente, and R. Vilela Mendes. Using immunology principles for fault detection. *Industrial Electronics, IEEE Transactions on*, 50:362–373, 2003.

[2] S. da Silva, M. Dias Júnior, and V. Lopes Junior. Damage detection in a benchmark structure using ar-arx models and statistical pattern recognition. *Journal of the Brazilian society of mechanical science and engineering*, Vol 29. No 2, 2007.

[3] S.da Silva, M. Dias Júnior, V. Lopes Junior, and M. J. Brennan. Structural damage detection by fuzzy clustering. *Mechanical systems and signal processing*, 22:1636–1649, 2008.

[4] J. Zhang, K. Worden, and W. J. Staszewski. Sensor optimisation using an immune system metaphor. In *Proceedings of 26th International Modal Analysis Conference (IMAC)*, 2008.

[5] J. R. Vieira de Moura Jr, S. Park, V. Steffen Jr, and D. J. Inman. Fuzzy logic applied to damage characterization through shm techniques. In *Conference Proceedings of the Society for Experimental Mechanics Series.*, 2008.

[6] B. Chen. Agent-based artificial immune system approach for adaptative damage detection in monitoring network. *Journal of network and computer applications*, 33(6):633–645, 2010.

[7] D. Tan, W. Qu, and J. Tu. The damage detection based on the fuzzy clustering and support vector machine. In *Intelligent System Design and Engineering Application (ISDEA) 2010 International Conference on , vol.2, no., pp.598,601, 13-14 Oct. 2010 doi: 10.1109/ISDEA.2010.404.*, 2010.

[8] B. Chen and C. Zang. Emergent damage pattern recognition using immune network theory. *Smart Structures and Systems*, 8 (1):69–92, 2011.

[9] Z. Chilengue, J.A. Dente, and P.J. Costa Branco. An artificial immune system approach for fault detection in the stator and rotor circuits of induction machines. *Electric power systems research*, 81:158–169, 2011.

[10] Y. Zhou, S.Tang, C. Zang, and R. Zhou. *Advances in information technology and industry applications lecture notes in electrical engineering*, volume 136, chapter An artificial immune pattern recognition approach for damage classification in structures, pages 11 – 17. Springer Link, 2012.

[11] W. Xiao. *Structural Health Monitoring and Fault Diagnosis based on Artificial Immune System*. PhD thesis, Worcester Polytechnic Institute, 2012.

[12] Z. Liu, Q. Zhou, Q. Chi, Y. Zhang, Y. Chen, and S. Qi. Structural damage detection based on semi-supervised fuzzy c-means clustering. In *Computer science & education (ICCSE), 2014 9th international conference on*, pages 551–556. IEEE, 2014.

[13] Y.Huang, L. Gong, S. Wang, and L. Li. A fuzzy based semi-supervised method for fault diagnosis and performance evaluation. In *IEEE/ASME Inernational conference on advanced intelligent mechatronic (AIM). Besanon, France*, 2014.

[14] M. Anaya, D. Tibaduiza, and F. Pozo. Data-driven methodology based on artificial immune system for damage detection. In V. Le Cam, L. Melvel, and F. Schoefs, editors, *7th European Workshop on Structural Health Monitoring*, pages 1341–1348, 2014.

[15] L. N. de Castro and J. Timmis. *Artificial Neural Networks in Pattern Recognition*, chapter Artificial Immune Systems: A Novel Approach to Pattern Recognition, pages 67–84. University of Paisley, 2002.

[16] A.K. Eroglu, Z. Erden, and A. Erden. Bioinspired conceptual design (bicd) approach for hybrid bioinspired robot design process. In *Mechatronics (ICM), 2011 IEEE International Conference on. 2011 , Page(s): 905- 910. Digital Object Identifier: 10.1109/ICMECH.2011.5971243.*, 2011.

[17] N. Cruz Cortés. *Sistema inmune artificial para solucionar problemas de optimización*. PhD thesis, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, 2004.

[18] P. J. Delves, S. J. Martin, D. R. Burton, and I. M. Roitt. *Roitt's essential immunology*. Wiley-Blackwell, 2011.

[19] L. N. de Castro and F. J. Von Zuben. Artificial immune systems: Part i basic theory and applications. Technical report, School of Computing and Electrical Engineering, State University of Campinas, Brazil, No. DCA-RT 01/99, December 1999.

[20] D. Trejo Pérez. Optimización global en espacios restringidos mediante un sistema inmune artificial. Master's thesis, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional-Mexixo D.F., 2005.

[21] U. Aickelin and D. Dasgupta. *Search Metodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, chapter Artificial Immune System, pages 375–399. Springer US, 2005.

[22] A. A. Freitas and J. Timmis. Revisiting the foundations of artificial immune systems for data mining. *IEEE Transactions on evolutionary computation*, 11(4):521–540, 2007.

[23] R. Xiao and T. Chen. Relationships of swarm intelligence and artificial immune system. *Internatonal journal of bio-inspired computation*, 5(1):35 – 51, 2013.

[24] X. Wang, X. Z. Gao, and S. J. Ovaska. Fusion of clonal selection algorithm and harmony search method in optimisation of fuzzy classification system. *International journal of bio-inspired computation.*, 1(1/2):80 – 88, 2009.

[25] V. T. Nguyen, T. T. Nguyen, K. T. Mai, and T. D. Le. A combination of negative selection algorithm and artificial immune network for virus detection. In *Future data and security engineering*, volume 8860, pages 97–106. First international conference, FDSE 2014, Ho Chi Minh city, Vietnam, November 2014.

[26] N. Lay and I. Bate. Applying artificial immune system to real-time embedded systems. In *IEEE Congress on evolutionary computation. CEC 2007*, pages 3743–3750. IEEE, September 2007.

[27] B. Mnassri, E. El Adel, and M. Ouladsine. Fault localization using principal component analysis based on a new contribution to the squared prediction error. In *16th Mediterranean Conference on Control and Automation*, 2008.

[28] A. Hyvarinen, J. Kahunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, INC, 2001.

[29] D. A. Tibaduiza. *Design and validation of a structural health monitoring system for aeronautical structures*. PhD thesis, Department of Applied Mathematics III, Universitat Politècnica de Catalunya, 2013.

[30] D. A. Tibaduiza, L. E. Mujica, and J. Rodellar. Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Structural control and health monitoring*, 20:1303 – 1316, 2013.

[31] D. A. Tibaduiza, L. E. Mujica, A. Güemes, and J. Rodellar. Active piezoelectric system using pca. In *Fifth European Workshop on Structural Health Monitoring*, 2010.

[32] D. A. Tibaduiza, M. A. Torres, L. E. Mujica, J. Rodellar, and C. P. Fritzen. A study of two unsupervised data driven statistical methodologies for detecting and classifying damages in structural health monitoring. *Mechanical Systems and Signal Processing*, 41:467 – 484, 2012.

[33] J. Westerhuis, T. Kourti, and J. MacGregor. Comparing alternative approaches for multivariate statistical analysis of batch process data. *Journal of Chemometrics*, 13:397–413, 1999.

[34] L.E. Mujica, M. Ruiz, F. Pozo, J. Rodellar, and A. Güemes. A structural damage detection indicator based on principal component analysis and statistical hypothesis testing. *Smart Materials and Structures*, 23(2), 2014.

[35] D.A. Tibaduiza, L.E. Mujica, J. Rodellar, and A. Güemes. Structural damage detection using principal component analysis and damage indices. *Journal of Intelligent Material Systems and Structures*, 2015.

Chapter 4

Detection and Classification of Structural Changes using Artificial Immune Systems and Fuzzy Clustering

Authors: Maribel Anaya, Diego A. Tibaduiza, Francesc Pozo
Published in:
Journal: International Journal of Bio-Inspired Computation, *to appear*
Publisher: Inder Science
Online ISSN: 1758-0374
Print ISSN: 1758-0366
2014 Impact Factor: 2.77 (Q1, 2/102 in the area of Computer Science, Theory and Methods)

This chapter is a true copy of the original paper published in the journal where the only changes are performed to fit the page setup.

Detection and classification of structural changes using artificial immune systems and fuzzy clustering

Among all the elements that are integrated into a structural health monitoring (SHM) system, methods or strategies for damage detection and classification are nowadays playing a key role in enhancing the operational reliability of critical structures in several industrial sectors. The main contribution of this paper is the application of a new methodology to detect and classify structural changes. The methodology is based on: (i) an artificial immune system (AIS) and the notion of affinity is used for the sake of damage detection and (ii) a fuzzy c-means algorithm is used for damage classification.

One of the advantages of the proposed methodology is the fact that to develop and validate the strategy, a model is not needed. Additionally, and in contrast to standard Lamb waves-based methods, there is no need to directly analyze the complex time-domain traces containing overlapping, multimodal and frequency dispersive wave propagation that distorts the signals and difficult the analysis.

The proposed methodology is applied to data coming from two sections of an aircraft skin panel. The results indicate that the proposed methodology is able to accurately detect damage as well as classify those damages.

Keywords:Artificial Immune Systems; Principal Component Analysis; Damage Indices; Fuzzy Clustering, Affinity Value.

4.1 Introduction

The use of non-destructive inspection methods (NDT) has proved to be a very useful tool for damage detection tasks. However, in some situations where it is impossible to manually inspect a structure –as in the inspection of large-scale structures–, the use of automated methods present significant advantages. Some of these advantages can be summarized as follows: (i) continuous monitoring, since the sensors are permanently attached to the structure; (ii) early damage detection; (iii) damage identification, among others. In this sense, structural health monitoring (SHM) extends the limits of the NDT methods by including the use of data processing algorithms, pattern recognition and continuous monitoring because the sensors are permanently attached to the structure. This is one the reasons why the development of improvements in data processing algorithms is a current demand. The contribution of the present work is the development of a methodology for data-driven damage classification using a bio-inspired algorithm, which is applied to data that comes from a piezoelectric system. More precisely, this work uses an artificial immune system that allows the use of this methodology as a pattern recognition approach. The use of artificial immune systems (AIS) is relatively new in the literature and, compared with the application of other approaches in SHM, there are still a reduced number of works. In the next lines we briefly compile in chronological order the most representative works in the use of AIS.

In 2003 Costa *et al.* [1] developed three module algorithms called T-module, B-module and D-module. These algorithms are based on immunologic principles to detect anomalous situations in a squirrel-cage motor induction. The T-module distinguishes between self and non-self conditions, the B-module analyzes the occurrence of both cells (self and non-self) and finally the D-module is similar to a T-module but with a reduced space. In this work, the normal operation condition of the machine (self) is represented by the frequency spectrum, that can include or not include harmonics.

In 2007, Da Silva *et al.* [2] presented a damage detection algorithm applying an auto-regressive model and auto-regressive model with exogenous input (AR-ARX). This algorithm is based on the structural vibration response measurements and the residual error as damage sensitive index. Data compression is used by means of principal component analysis (PCA) and the fuzzy c-means clustering method is used to quantify the damage sensitive index. In this paper, the authors used a benchmark problem with several damage patterns to test the algorithm. As the main result, a structural diagnosis was obtained with high correlation with the actual state of the structure. Later on 2008, Da Silva *et al.* [3] developed a strategy to perform structural health monitoring. This strategy included three different phases: (i) the use of principal component analysis to reduce the dimensionality of the time series data; (ii) the design of an autoregressive-moving-average (ARMA) model using data from the healthy structure under several environmental and operational conditions; and finally (iii) the identification of the state of the structure through a fuzzy clustering approach. In this paper, the authors compared the performance of two fuzzy algorithms, fuzzy c-means (FCM) and Gustafson-Kessel (GK) algorithms. The proposed strategy was applied to data from a benchmark structure at Los Alamos National Laboratory. The work showed that the GK algorithm outperforms the FCM algorithm, because the first algorithm considers an adaptive distance norm and allows clusters with several geometrical distributions.

Also in 2008, Zhang *et al.* [4] used a clonal selection algorithm to solve a combinatorial optimization problem called *sensor optimization*. This problem consists in choosing an appropriate distribution of a set of sensors in a structure to detect impacts. To test the algorithm, the authors used a composite plate instrumented with 17 lead zirconium titanate (PZT) transducers.

De Moura *et al.* [5] presented a fuzzy based meta-model to detect damages in a flat structure under corrosion conditions. This work considers data obtained from a SHM approach based on electro-mechanic impedance. Chen [6], in 2010, applied an agent-based artificial immune system for adaptative damage detection. In the approach, a group of agents is used as immune cells (B-cells) patrolling over a distributed sensor network installed in the structure. The damage diagnosis is based on the analysis of structural dynamic response data. Each mobile agent inspects the structure using agent based cooperation protocols. In 2010, Tan *et al.* [7] presented a damage detection algorithm based on fuzzy clustering and support vector machines (SVM). In this work, as a first step, the wavelet packet transform is used to decompose the accelerator data from the structure and extract the energy of each wavelet component. Consequently, this energy is used as a damage index. In further steps, damages are classified by means of fuzzy clustering. As a final phase, damages are identified using a vector machine. The numerical example illustrated in this work shows that the proposed method is able to identify the damage from the spatial truss structure. Also in 2010, Casciati [8, 9] presented two statistical approaches for the problem of damage detection and localization. In [8], the detection is based on comparing the SSE (sum of squares errors) histograms of the undamaged case and the *new* set of data which could either represent an undamaged or damaged case. A similar approach is introduced in

[9], where the comparison is performed using linear response surface (RS) models to represent the different structural states of the structure. In both works, the localization of the damages is based on the identification of the largest discrepancies resulting from the comparison. In 2011, Chen and Zang [10] presented an algorithm based on immune network theory and hierarchical clustering algorithms. Chilengue *et al.* [11] presented and artificial immune system (AIS) approach to detect and diagnose failures in the stator and rotor circuits of an induction machine. In the approach, the dynamic of the machine is compared before and after the fault condition. Similarly, the alpha-beta ($\alpha\beta\gamma$) transformation (also known as the Clarke transformation) was applied to the stator current to obtain a characteristic pattern of the machine that is finally applied to the pattern recognition algorithm.

In 2012, Zhou *et al.* [12], inspired by Chen's work, developed a damage classifier in structures based on the immune principle of clonal selection. Using evolution algorithms and the immune learning, a high quality memory cell is generated in the classifier generated able to recognize several damage patterns. In 2012, Xiao [13] developed a structural health monitoring and fault diagnosis system based on artificial immune system. In this approach, the antigen represents the structural state (health or damage), whereas the antibody represents database information to identify a damage pattern. In this work, the feature space is formed by natural frequencies and modal shapes collected by simulation of the structure in free vibration and seismic response. Quite recently, Liu *et al.* [14], in 2014, proposed a structural damage detection method using semi-supervised fuzzy c-means clustering method, wavelet packet decomposition and data fusion. This method is applied to detect damage in a four-level benchmark model. The data that was used include 11 damage pattern and 9 samples per damage. The method uses a Daubechies wavelet filter and 6 decompositions levels. According to the results, the method can achieve a reasonable detection performance. Huang *et al.* [15], in 2014, proposed an automatic methodology to know the status of a machine. The introduced method includes a semi-supervised fuzzy-based method to detect the faults or anomalies in the machine and to classify the unknown faults. The authors described two steps for the learning procedure: (i) a fuzzy c-means clustering to get candidates of labels (fuzzy centers); and (ii) a label matching by filtering out of the unreasonable labels candidates. The proposed method is validated in a roller bearing test top diagnose the state of the machine.

The previous methods can be classified differently according to the algorithms that operate on the extracted features to detect, quantify or localize the damage state of the structure [16]. This statistical discrimination of features for damage detection can be classified in two main groups: (i) supervised learning and (ii) unsupervised learning. The unsupervised learning is applied to data that does not contain examples from the damaged structure. If the data from both the undamaged and damaged structure is available, the supervised approach can be used to classify and quantify damage.

Compared with the works previously reviewed, the methodology described on the current work presents a new point of view, since we use an artificial immune system (AIS) as a pattern recognition approach for structural damage classification. We also use some other techniques such as PCA –both for data reduction and to compute damage indices to create the feature vectors from the healthy structure and under different states– or fuzzy clustering –for damage classification–.

This paper is organized as follows. Section 2 describes a theoretical background that includes basic concepts about the methods and elements used in the methodology. Section 3 includes the damage detection methodology followed by the description about the experimental

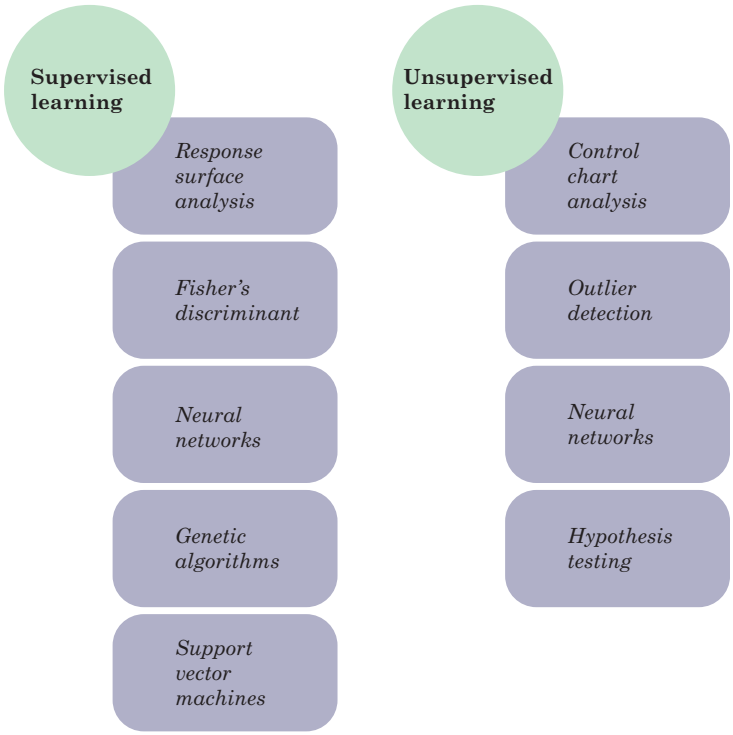


Figure 4.1: A state-of-the-art classification of the statistical models developed to enhance the structural health monitoring process [16].

setup in Section 4. The experimental results are included in Section 5. Finally, some conclusions are drawn.

4.2 General framework

The current work is based on data-driven analysis. This means that the damage detection will be developed by analyzing and interpreting the data collected in several experiments from the structures under to diagnose. To perform this analysis, a bio-inspired methodology based on features extraction for pattern recognition is developed. For the sake of clarity, basic concepts and fundamentals about the methods that will be used are presented in the following subsections.

4.2.1 Bio-inspired systems

The adaptation of the different living beings of the planet in harsh environments and the development of skills to solve the inherent problems in the interaction with the world of nature has resulted in the evolution of the species in order to survive and avoid their extinction. Some examples are the communication abilities, the reasoning, the physical structures design or the response of the body to external agents, among others [17].

Taking advantage of the fact that nature provides robust and efficient solutions to many different problems, more and more researchers on different areas work in the development of biologically inspired hardware and algorithms. The inspiration process is called *biomimetic* or *bioinspired* and aims to apply the developments in the field of biology to the engineering

developments [18].

4.2.2 Natural immune systems

The human immune system (HIS) is a complex and robust defense mechanism composed by a large network of specialized cells, tissues and organs. The system further includes an elevated number of sensors and a high processing capability. The human immune system has proved its effectiveness in the detection of foreign elements by protecting the organism against disease. The principal skills of the human immune system are:

- To discriminate between its own cells (self) and foreign cells (non-self).
- To recognize different invaders (called antigens) in order to ensure the protection of the body.
- To learn from specific antigens and adapt to them in order to improve the immune response to this kind of invader.

In general, when a foreign particle wants to gain access to the organism, it has to break several defense levels provided by the immune system that protects the organism. The idea of several defense levels is illustrated in Figure 4.2. These levels can be summarized as follows [19]:

- *External barriers*. These are the first and the major line of defense into the human body. This level can include elements such as the skin, the mucus secreted by the membranes, the tears, the saliva and the urine. All of these elements present different physiological conditions that are harmful to the antigens, as the temperature or the pH level, among others. The response of these barriers is equal for any foreign invader [20].
- *Innate immune system*. This barrier refers to the defense mechanisms that are activated immediately or within a short lapse of time of an antigen's arrival in the body. The innate immune system operates when the first barrier has been broken This system, in opposition to the adaptive immune system, is not adaptive[19].
- *Adaptive immune system*. This is the last defense level and reacts to the stimulus of foreign cells or antigens that evade both the external barriers and the innate immune defense [19]. Adaptive immunity creates some sort of memory that leads to an improved response to future encounters with this antigen.

With respect to different type of cells, the immune system includes cells born in the bone marrow that are usually called *white blood cells*, *leukocytes* or *leucocytes* [21]. Among the white blood cells, it is possible to highlight the T-cells and the B-cells. On one hand, the T-cells are so called since their maturation takes place in the thymus.Besides, this kind of cells have high mobility and can also be found in the blood and the lymph [22]. One can distinguish three types of T-cells:

- the T-helper cells, involved in the activation of B-cells;
- the T-killer cells that destroy the invaders; and finally

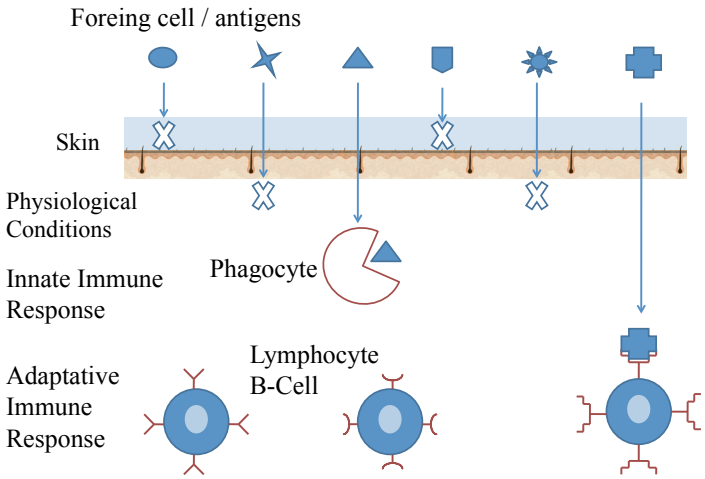


Figure 4.2: Schematic representation of a system's defense barriers.

- the T-suppressor cells that avoid the allergic reactions [23].

On the other hand, the B-cells produce and secrete a special protein called antibody, which recognizes and binds the antigen. The responsibility of each B-cell is the production of a specific antibody. This protein is then used for signaling other cells what elements have to be removed from the body [22]. When the antigen passes over the first barrier of the immune system, the HIS performs the following steps to eliminate the invader [22]:

1. The specialized cells of the immune system, called *antigen presenting cells* (APCs) (macrophages, for instance). These cells active the immune response by ingesting the antigen and dividing it into simple substances known as antigenic peptides.

2. These peptides are joined to the molecules called *major histocompatibility complex* (MHC), inside of the macrophage, and the result passes to the immune cell surface.

3. The T-cells have receptor molecules able to identify and recognize different combinations of MHC-peptide. When the receptor molecule recognizes the combination, the T-cell is activated and sends a chemical signal to other immune cells.

4. The B-cells are activated by chemical signals and they initiate the recognition of the antigen in the bloodstream. This process is performed by the receptor molecules in the B-cells.

5. The mission of the B-cells –when they are activated– is to secrete antibodies to bind the antigens they find, and to neutralize and eliminate them from the body.

The T- and B-cells that have recognized the antigen proliferate and, some of them, become memory cells. These memory cells remain in the immune system to eliminate the same antigen –in the future– in a more effective manner [17, 22].

Three immunological principles are used in artificial immune systems [13, 17, 22]:

Table 4.1: Analogy between the biological immune system and artificial immune system [13].

Biological Immune System	Artificial Immune System in SHM
antibodies	a detector of a specific pattern
antigens	structural health or damage condition
matured antibodies	database or information system for damage detection
recognition of antigens	identification of health and damage condition
process of mutation	training procedure
immune memory	memory cells

- *Immune network theory*. This theory was first introduced by Niels Jerne in 1974 and describes how the immune memory is built by means of the dynamic behaviour of the immune system cells. These cells can recognize by themselves, detect invaders, as well as interconnect between them to stabilize the network [19].

- *The negative selection*. The negative selection is a process that allows the identification and eradication of the cells that react to the own body cells. This ensures a convenient operation of the immune system since it is able to distinguish between foreign molecules and self-molecules thus avoiding autoimmune diseases. This process is similar to the maturation of T-cells carried out in the thymus [17].

- *The clonal selection*. This is a mechanism of the adaptive immune responses in which the cells of the system are adapted to identify an invader element [22]. Antibodies that are able to recognize or identify an antigen can proliferate. Those antibodies unable to recognize the antigens are eliminated. The new cells are clones of their parents and they are subjected to an adaptation process by mutation. From the new antibody set, the cells with the greatest affinity with respect to the primary antigen are selected as memory cells therefore excluding the rest.

4.2.3 Artificial immune systems

Artificial immune systems (AIS) are an adaptive and bio-inspired computational systems based on the processes and performance of the human immune system (HIS) and its properties –diversity, error tolerance, dynamic learning, adaptation, distributed computation and self-monitoring– [24, 25]. Nowadays, these computational systems are used in several research areas such as pattern recognition [18], optimization [22, 26], computer security [27], among others [28]. Table 4.1 presents the analogy between the natural and artificial immune systems applied to the field of structural health monitoring.

In the implementation of an artificial immune system, it is fundamental to bear in mind two important aspects:

- To define the role of the antigen (ag) and the antibody (ab) in the context of the application. Both are represented or coded in the same way. This representation is generally given by a vector of binary or real numbers [23].

- To define the mechanism that measures the degree of correspondence between an antigen and an antibody. This measure is usually related to the distance between them [17]. If both an antigen and an antibody are represented by $L-$dimensional arrays,

$$ab \in \mathbb{R}^L,$$
$$ag \in \mathbb{R}^L,$$

the distance d between them can be computed using, for instance, the Euclidean distance (related to the $2-$norm) or the so-called Manhattan distance (related to the $1-$norm) as in equations (4.1) and (4.2), respectively [21]:

$$d(ab, ag) = \|ab - ag\|_2 = \sqrt{\sum_{i=1}^{L}(ab_i - ag_i)^2} \tag{4.1}$$

$$d(ab, ag) = \|ab - ag\|_1 = \sum_{i=1}^{L}|ab_i - ag_i| \tag{4.2}$$

Finally, there exists the adaptation process of the molecules in the artificial immune system. This adaptation allows to include the dynamic of the system, for instance, the antibodies excitation, cloning all the excited antibodies and the interconnection between them. All these elements are adapted from the three immunologic principles previously introduced.

4.2.4 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a classical method used in applied multivariate statistical analysis with the goal of dimensionality reduction and, more precisely, feature extraction and data reduction. It was developed by Karl Pearson in 1901 and integrated to the mathematical statistics in 1933 by Harold Hotelling [29]. The general idea in the use of PCA is to find a smaller set of variables with less redundancy [30]. To find these variables, the analysis includes the transformation of the current coordinate space to a new space to re-express the original data trying to filtering the noise and redundancies. These redundancies are measured by means of the correlation between the variables.

Matrix unfolding

The application of PCA starts –for each actuation phase– with the collected data arranged in a three dimensional matrix $n \times L \times N$. The matrix is subsequently unfolded –as illustrated in Figure 4.3–in a two dimensional $n \times (N \cdot L)$ matrix as follows:

$$\mathbf{X} = \begin{pmatrix} x_{11}^1 & x_{12}^1 & \cdots & x_{1L}^1 & x_{11}^2 & \cdots & x_{1L}^2 & \cdots & x_{11}^N & \cdots & x_{1L}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1}^1 & x_{i2}^1 & \cdots & x_{iL}^1 & x_{i1}^2 & \cdots & x_{iL}^2 & \cdots & x_{i1}^N & \cdots & x_{iL}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1}^1 & x_{n2}^1 & \cdots & x_{nL}^1 & x_{n1}^2 & \cdots & x_{nL}^2 & \cdots & x_{n1}^N & \cdots & x_{nL}^N \end{pmatrix} \tag{4.3}$$

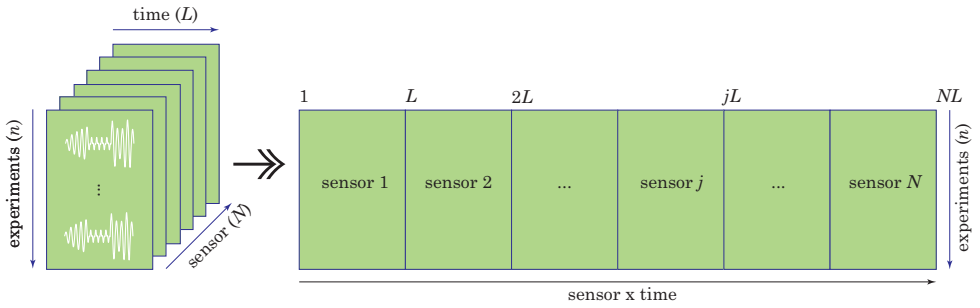


Figure 4.3: The collected data arranged in a three dimensional matrix is unfolded in a two dimensional matrix [31].

Matrix $\mathbf{X} \in \mathcal{M}_{n\times(N\cdot L)}(\mathbb{R})$ –where $\mathcal{M}_{n\times(N\cdot L)}(\mathbb{R})$ is the vector space of $n \times (N \cdot L)$ matrices over $\mathbb{R}$– contains data from N sensors at L discretization instants and n experimental trials [32]. Consequently, each row vector $x_i^T = \mathbf{X}(i,:) \in \mathbb{R}^{N\cdot L},\ i = 1,\ldots,n$ represents, for a specific experimental trial, the measurements from all the sensors. Equivalently, each column vector $\mathbf{X}(:,j) \in \mathbb{R}^n,\ j = 1,\ldots,N\cdot L$ represents measurements from one sensor in the whole set of experimental trials.

In other words, the objective is to find a linear transformation orthogonal matrix $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ that will be used to transform the original data matrix $\mathbf{X}$ according to the following matrix multiplication

$$\mathbf{T} = \mathbf{XP} \in \mathcal{M}_{n\times(N\cdot L)}(\mathbb{R}). \tag{4.4}$$

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 $\mathbf{PP}^{\mathrm{T}} = \mathbf{I}$, 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

$$\mathbf{X} = \mathbf{TP}^T.$$

Group scaling

Since the data in matrix $\mathbf{X}$ come from experimental trials and could have different magnitudes and scales, it is necessary to apply a preprocessing step to scale the data using the mean of all measurements of the sensor at the same time and the standard deviation of all measurements of the sensor [32].

More precisely, for $k = 1, 2, \ldots, N$ we define

$$\mu_j^k = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^k,\ j = 1, \ldots, L, \tag{4.5}$$

$$\mu^k = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} x_{ij}^k, \tag{4.6}$$

$$\sigma^k = \sqrt{\frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}(x_{ij}^k - \mu^k)^2}, \tag{4.7}$$

where μ_j^k is the mean of the n measures of sensor k at the time instant j; μ^k is the mean of all the measures of sensor k; and σ^k is the standard deviation of all the measures of sensor k. Therefore, the elements x_{ij}^k of matrix $\mathbf{X}$ are scaled to define a new matrix $\check{\mathbf{X}}$ as

$$\check{x}_{ij}^k := \frac{x_{ij}^k - \mu_j^k}{\sigma^k}, \quad i = 1, \ldots, n,\ j = 1, \ldots, L,\ k = 1, \ldots, N. \tag{4.8}$$

When the data are normalized using equation (4.8), the scaling procedure is called *variable scaling* or *group scaling* [32]. According to former studies of the authors [31, 33, 34], group scaling presents a better performance than other kind of normalizations. The reason is that group scaling considers changes between sensors and does not process them independently. Further discussion on this issue can be found in [32, 33].

For simplicity, and throughout the rest of the paper, the scaled matrix $\check{\mathbf{X}}$ is renamed as simply $\mathbf{X}$. The mean of each column vector in the scaled matrix $\mathbf{X}$ can be computed as

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\check{x}_{ij}^k &= \frac{1}{n}\sum_{i=1}^{n}\frac{x_{ij}^k - \mu_j^k}{\sigma^k} = \frac{1}{n\sigma^k}\sum_{i=1}^{n}\left(x_{ij}^k - \mu_j^k\right)\\ &= \frac{1}{n\sigma^k}\left(\sum_{i=1}^{n} x_{ij}^k - n\mu_j^k\right)\\ &= \frac{1}{n\sigma^k}\left(n\mu_j^k - n\mu_j^k\right) = 0.\end{aligned}$$

Since the scaled matrix $\mathbf{X}$ is a *mean-centered* matrix, it is possible to calculate the covariance matrix as follows:

$$\mathbf{C_X} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R}). \tag{4.9}$$

The covariance matrix $\mathbf{C_X}$ is a $(N \cdot L) \times (N \cdot L)$ symmetric 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 follows:

$$\mathbf{C_X P} = \mathbf{P}\Lambda \tag{4.10}$$

where the columns of $\mathbf{P} \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvectors of $\mathbf{C_X}$. The diagonal terms of matrix $\Lambda \in \mathcal{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ are the eigenvalues $\lambda_i,\ i = 1, \ldots, N \cdot L$, of $\mathbf{C_X}$ whereas the

off-diagonal terms are zero, that is,

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

The eigenvectors $p_j,\ j = 1, \ldots, N \cdot L$, 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* or the *loading vectors* 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

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

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

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

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 $m-$dimensional space to get a data matrix $\hat{\mathbf{X}}$ as follows:

$$\hat{\mathbf{X}} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T \in \mathcal{M}_{n \times m}(\mathbb{R}). \tag{4.11}$$

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

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}, \tag{4.12}$$

or, equivalenty,

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{E} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T + \mathbf{E}. \tag{4.13}$$

The residual error matrix $\mathbf{E}$ describes the variability not represented by the data matrix $\hat{\mathbf{X}}$, and can be also expressed as

$$\mathbf{E} = \mathbf{X}(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T). \tag{4.14}$$

Even though 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 [35].

4.2.5 Damage detection indices based on PCA

Several damage detection indices based on PCA have been proposed and applied with excellent results in pattern recognition applications. In particular, two damage indices are commonly

used: (i) the Q index (also known as SPE, *square prediction error*) and (ii) the Hotelling's T^2 index.

The Q index of the ith experimental trial x_i^T measures the magnitude of the vector $\tilde{x}_i^T := \tilde{\mathbf{X}}(i,:)$, that is, the events that are not explained by the model of principal components [36], and it is defined as follows:

$$Q_i = \tilde{\mathbf{X}}(i,:)\tilde{\mathbf{X}}(i,:)^T = x_i^T(\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T)x_i. \tag{4.15}$$

The T^2 index of the ith experimental trial x_i^T is the weighted norm of the projected vector $\hat{t}_i^T := \hat{\mathbf{T}}(i,:) = x_i^T\hat{\mathbf{P}}$, that is, a measure of the variation of each sample within the PCA model and it is defined as follows:

$$T_i^2 = \sum_{j=1}^{\ell} \frac{\hat{t}_{i,j}^2}{\lambda_j} = \hat{t}_i^T\Lambda^{-1}\hat{t}_i = x_i^T(\hat{\mathbf{P}}\Lambda^{-1}\hat{\mathbf{P}}^T)x_i \tag{4.16}$$

4.3 Damage detection methodology

The damaged detection methodology that we present in this paper involves an active piezoelectric system to inspect the structure. This active system consists of several piezoelectric transducers (lead zirconium titanate, PZT) distributed on different positions of the structure and working both as actuators or sensors in different actuation phases. Each PZT is able to produce a mechanical vibration if some electrical excitation is applied (actuator mode). Besides, the PZT are able to detect time varying mechanical response data (sensor mode). In each phase of the experimental stage, just one PZT is used a the actuator (exciting the structure). Then, the propagated signal through the structure is collected by using the rest of PZT, which are used as sensors. This procedure is repeated in as many actuation phases as the number of PZT on the structure.

To determine the presence of damage in the structure, the data from each actuation phase will be used in the proposed artificial immune system. The proposed methodology is performed in three steps: (i) data pre-processing and feature extraction; (ii) training process and; (iii) testing. More precisely, in the first step the collected data is organized, pre-processed and dimensionally reduced –by means principal component analysis– to obtain relevant information. The damage indices in equations (4.15)-(4.16) are used to define the feature vectors. The training step includes the evolution of the data to *generate* good representatives for each pattern, damage or structural condition. A good accuracy in the damage detection using AIS depends on a good training. Finally, the testing step includes new data to evaluate the training step and the knowledge of the current state of the structure.

4.3.1 Data pre-processing, PCA modeling and feature extraction

For each different phase (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 $\mathbf{X}$ is defined and arranged as in equation (4.3) in Section 4.2.4 and scaled as stated in Section 4.2.4. PCA modeling basically consists of computing the projection matrix $\mathbf{P}$ for each phase as in equation (4.5). Matrix $\mathbf{P}$, renamed $\mathbf{P}_{\text{model}}$, provides an improved and dimensionally limited

representation of the original data $\mathbf{X}$. The number of principal components retained at each different phase account for at least 90% of the cumulative variance.

Subsequently, the data from different structural states are projected into each PCA model in order to obtain the scores and calculate the damage detection indices T^2 and Q as in Sections 4.2.4 and 4.2.5. This way, for each experiment, a two dimensional *feature vector*

$$f^i = (T_i^2, Q_i)^T \in \mathbb{R}^2,\ i = 1, \ldots, \nu \tag{4.17}$$

is defined, where ν is the total number of experiments. The feature vector could include more components, as the scores, for instance. Several test were then performed in this sense with the combination of scores and damage indices. However, the results indicated that the single use of T^2 and Q lead to the best results. One of the reasons about the use of the damages indices can be found in [36]. In this paper, Rodellar *et al.* showed that the use of scores is not sufficient for damage detection when two scores do not account for a high cumulative variance. This result implies that it is necessary to use another type of measurement or statistic to obtain an accurate discrimination of the presence of damage in a structure.

To keep the affinity values within the range $[0, 1]$, the norm of the feature vectors $f_i,\ i = 1, \ldots, \nu$ is normalized to the unit circle. The normalization process uses the maximum norm of the feature vectors, that is,

$$M := \max_{i=1,\ldots,\nu} \|f^i\|,$$

where

$$\|f^i\| = \sqrt{\left(f_1^i\right)^2 + \left(f_2^i\right)^2},$$

and therefore the normalized feature vector f^i_{norm} of $f^i = (T_i^2, Q_i)^T$ is as follows

$$f^i_{\text{norm}} = \left(\frac{T_i^2}{M}, \frac{Q_i}{M}\right).$$

Since all the feature vectors are located within a unit circle, the Euclidean distance between any feature vectors is less than or equal to 2. The *healthy data set* (HDS) is defined as

$$HDS = \bigcup_{i=1}^{\nu} \{f^i\}. \tag{4.18}$$

4.3.2 Training step

This step can be modified according to different goals. For instance, in the most basic case in damage identification –the detection–, the training only needs to consider the feature vectors that come from data of the healthy structure. However, in a more complex analysis –the classification, for instance–, the training process must include the feature vectors of data coming from the structure in different and known structural states. The steps to perform the training in the basic case are summarized as follows:

- Randomly select $2k \in \mathbb{N},\ 2k < \nu$, feature vectors. The remaining $\nu - 2k$ feature vectors will be used in the testing process. This set of $2k$ feature vectors is divided in two subsets of the same size k, the *antibody training set* (AB_{training}) and the *antigen training set* (AG_{training}). This step is represented in Figure 4.4.

Figure 4.4: Random selection of the antibody (AB_{training}) and antigen (AG_{training}) training sets.

- Compute the affinity between the antibodies and antigens of the AB_{training} and AG_{training} sets, respectively. The affinity between an antibody and an antigen is defined as

$$\mathit{aff}(ab, ag) := 1 - \frac{1}{2} d(ab, ag),$$

where $d(ab, ag)$ is the distance –defined in equation (4.1)– between the feature vectors of ab and ag, respectively. Since the Euclidean distance between any feature vectors is less than or equal to 2, their affinity lies within the range $[0, 1]$.

- Evolve the antibodies. The evolution of the antibodies is performed when these are stimulated by an invading antigen invader and it consists on the mutation of the antibody. The mutation is performed by mutating the feature vectors of the cloned antibodies as shown in equation (4.19)

$$ab_{\text{evolved}} = ab + MV \cdot \phi, \tag{4.19}$$

where ab_{evolved} is the mutated antibody and MV represents the mutation value –a value used to indicate the mutation degree of the feature vector of an antibody–. In the present implementation, the mutation value is defined as in equation (4.20)

$$MV = 1 - CV, \tag{4.20}$$

where CV is the *clonal value* –a value that measures the response of an artificial B-cell to an antigen– and is equal to the affinity between the antibody and the stimulating antigen. The vector

$$\phi = (\phi_1, \phi_2)^T \in \mathbb{R}^2$$

in equation (4.19) is a randomly generated vector. Each element $\phi_i,\ i = 1, 2$ of the random vector is a normally distributed random variable with mean zero and standard deviation $\sigma = 0.5$.

The mutated antibody feature vectors must lie within the unit circle. Therefore, the norm of the feature vector for each mutated antibody is immediately checked after the mutation according to the following procedure:

- if $\|ab_{\text{evolved}}\| \leq 1$, then no normalization is performed.
- if $\|ab_{\text{evolved}}\| > 1$, then

$$ab_{\text{evolved}} = (\|ab\| + \mathcal{U} \cdot (1 - \|ab\|)) \cdot \frac{ab_{\text{evolved}}}{\|ab_{\text{evolved}}\|},$$

where $\mathcal{U}$ is a uniform random function with a value within the range of $[0, 1]$.

The norm of the mutated antibody is greater than the norm of the original antibody and less than 1.

The *clonal rate* (CR) is an integer value used to control the number of antibody clones allowed. The number of clones (NC) is defined in equation (4.21)

$$NC = \lfloor CR \cdot CV \rfloor, \tag{4.21}$$

where $\lfloor \cdot \rfloor$ is the floor function. In this paper the value of CR is 8.

The highest affinity antibody is chosen as the candidate memory cell for possible updating of memory cell set.

- Define the threshold. A threshold T_h is defined in order to update the memory cell set to improve the representation quality of memory cells for the healthy state of the structure. This threshold is defined as a weighted affinity of the two elements in the healthy data set (HDS) in equation (4.18) with the maximum Euclidean distance. That is,

$$\Delta = \max_{i,j=1,\ldots,n} \|f^i - f^j\| \tag{4.22}$$

$$\delta = \frac{7}{25}\Delta \tag{4.23}$$

$$T_h = 1 - \frac{1}{2}\delta \tag{4.24}$$

Then a comparison between the candidate memory cell and all the elements in the healthy data set (HDS) is performed through the affinity. If the affinities are greater than or equal to the threshold, the candidate memory cell becomes memory cell of the healthy state of the structure. Otherwise, the candidate memory cell is eliminated. The main outcome of this step is the *memory cell set* of the healthy state ($MCSH$) of the structure. This algorithmic training process is represented in Figure 4.5.

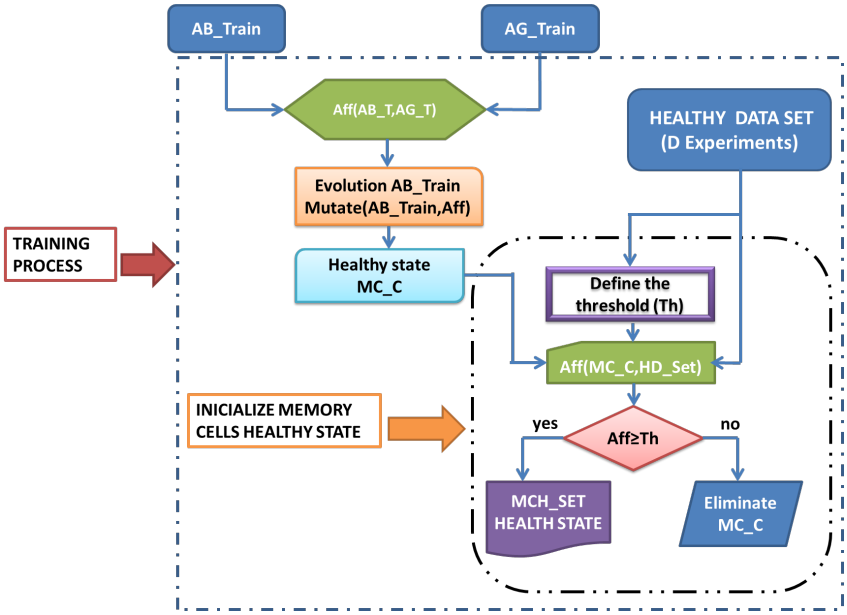


Figure 4.5: Training process in an artificial immune system applied (IAS) to structural health monitoring (SHM).

4.3.3 Testing step

The damage detection algorithm is finally illustrated in Figure 5. The damage detection is based on the affinity values between the elements in the memory cell set of the healthy state ($MCSH$) –acting as antibodies– and the data coming from the structure to test (TD, test data) –acting as antigens–. A *detection threshold* (DT_h) is defined in equation (4.25) for this purpose,

$$DT_h = \min_{\substack{ab \in MCSH \\ i \in \{1,\dots,\nu\}}} aff(ab, f^i), \tag{4.25}$$

that is, the minimum affinity between the elements in the memory cell set of healthy state (MCSH) and the elements in the healthy data set (HDS).

When the affinity is less than the threshold DT_h, we say that the data has been collected from a damaged state of the structure. Otherwise, the data comes from an undamaged structure.

4.3.4 Damage classification

After the detection of damage, the data that comes from the *damaged* structure will be classified. To this end, this data is used as input to a fuzzy clustering algorithm.

Fuzzy clustering is an unsupervised technique normally used to organize or classify data into groups mainly to obtain an accurate representation of a system behavior [37]. The algorithm builds these groups by grouping data with similar information or features. In this technique, a datum is simultaneously associated to many clusters or groups by using a membership function. A huge value to the membership indicates a high confidence in the assignment of the data to the cluster [38].

Data that can be applied to the clustering technique can be numerical (quantitative), categorical (qualitative) or a mixture of both. These data are usually observations of a physical process

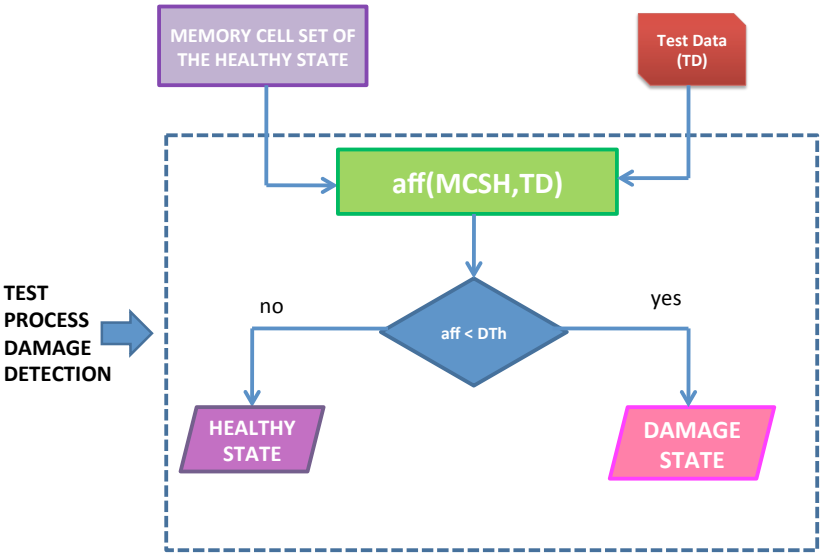


Figure 4.6: Damage detection process

wherein each observation is an $p-$dimensional row vector, $p \in \mathbb{N}$, that group p measured variables as in equation (4.26):

$$\boldsymbol{\psi} = (\psi_1, \psi_2, \ldots, \psi_p)^T \in \mathbb{R}^p \tag{4.26}$$

A set of $q \in \mathbb{N}$ observations is denoted by means of a $p \times q$ matrix where the columns are named *features* or *attributes* and the rows are named *patterns*. Matrix $\boldsymbol{\Psi}$ in equation (4.27)

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1p} \\ \psi_{21} & \psi_{22} & \cdots & \psi_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{q1} & \psi_{q2} & \cdots & \psi_{qp} \end{pmatrix} \tag{4.27}$$

is referred as data matrix [39].

The definition of a cluster depends on the context of the application. Generally the cluster is a group of data that are most similar to one another that the data of other group. The similarity is a mathematical measure that is often defined by a distance norm. The geometrical shapes, sizes and densities of the clusters can be spherical, hollow or elongated. Furthermore, the cluster can be a linear or nonlinear subspace; besides, this cluster can be overlapping each other, well separated or continuously connected [40, 41].

The clustering methods can be classified according to the characteristics of the subsets – fuzzy or crisp (hard)–. In the fuzzy clustering method it is possible that an object or data with different degrees of membership belong to several clusters. However, the hard clustering method just allows a data to belong to a cluster [37].

There are several algorithms in fuzzy clustering. In this work, the fuzzy c-means algorithm (FCM) was used. This technique was developed and introduced by Joe Dunn in 1973 and improved in 1981 by Jim Bezdek. This technique is frequently used in the field of pattern recognition [42, 43]. In this algorithm, a $p-$dimensional data set is grouped into a specific

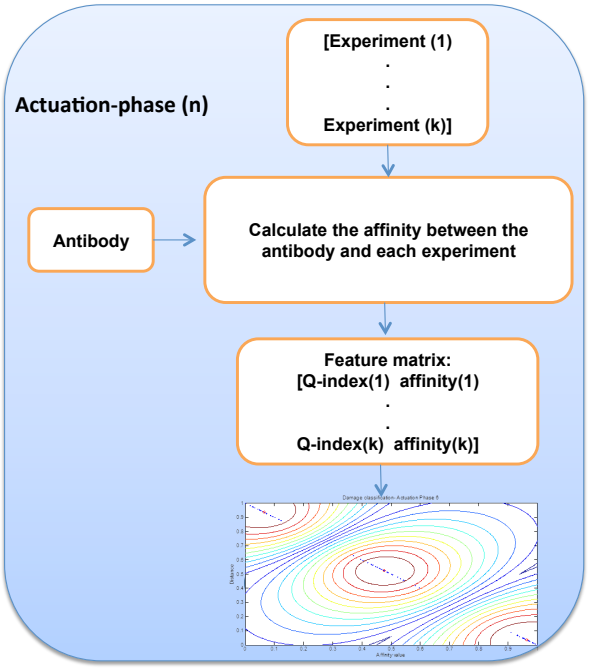


Figure 4.7: Damage classification process

number c of clusters depending on a membership degree. Each cluster is represented by its own center c_i. The measure of membership between a data point and a cluster is given by the Euclidean distance [39, 40].

In this paper, for each actuator phase, an antibody ab of the healthy data set is selected. Assuming that we have three different damages and ν experiments per damage, the cardinality of the damage data set (DDS) is

$$\#DDS = 3\nu.$$

Then, for each experiment of the damage data set (DDS), a two dimensional feature vector $f^i,\ i = 1, \ldots, 3\nu$ is computed as in equation (4.17). Then, the ith row in the feature matrix $\mathbf{\Psi}$ is the two dimensional vector

$$(Q_i, \mathit{aff}(ab, f^i)) \in \mathbb{R}^2,\ i = 1, \ldots, 3\nu, \tag{4.28}$$

where Q_i is the damage index defined in equation (4.15). The scheme of the damage classification step is illustrated in Figure 4.7.

A feature matrix $\mathbf{\Psi}$ is built for each actuator phase. In the particular case of this work, the number of clusters is 3, and the weighting exponent is 4. The weighting exponent determines the fuzziness of the clusters. Finally, the tolerance of the clustering method is 10^{-6}.

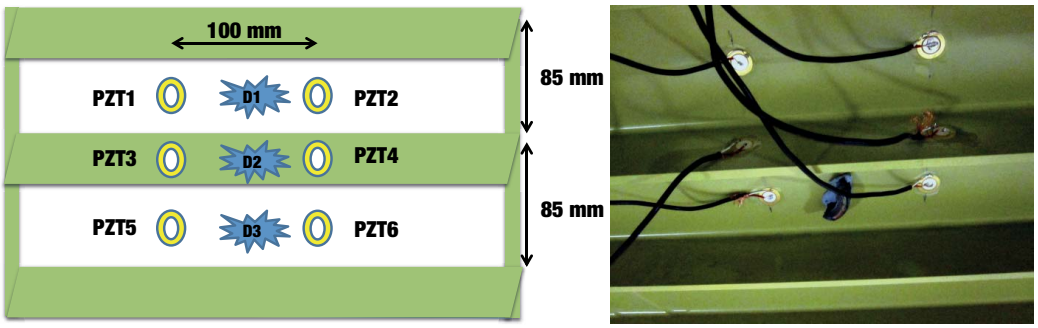
4.4 Experimental setup and experimental results

4.4.1 Experimental setup

To test the proposed methodology, data from an aircraft skin panel is used. The structure is divided in small sections by means of stringers and ribs as shown in Figure 4.8. To validate the

proposed methodology, two sections of this structure were used. The dimensions of each section and the damage description are depicted in Figure 4.9. The damaged areas were simulated by adding different masses at different locations. These sections were instrumented with 6 PZT transducers: two in the upper section; two in the lower section; and two in the rib. The transducers dimensions are: 26 mm diameter and 0.4 mm thickness.

Figure 4.8: Aircraft skin panel



(a) a. (b) b.

Figure 4.9: Damage description.

4.4.2 Experimental results

As said in Section 4.3.1, the experiments are performed in 6 independent phases: (i) piezoelectric transducer 1 (PZT1) is configured as actuator and the rest of PZTs as sensors; (ii) PZT2 as actuator; (iii) PZT3 as actuator; (iv) PZT4 as actuator; (v) PZT5 as actuator; and (vi) PZT6 as actuator.

To apply the proposed methodology, and for each phase, the collected data is arranged in a matrix as in equation (4.3) in Section 4.2.4. With this *unfolded* data, the PCA model $\mathbf{P}$ is built as explained in Sections 4.2.4 and 4.3.1 using data from the healthy structure. In Figure 4.10 the amount of variance accounted for by each principal component is illustrated, for phases 1, 3, 5 and 6.

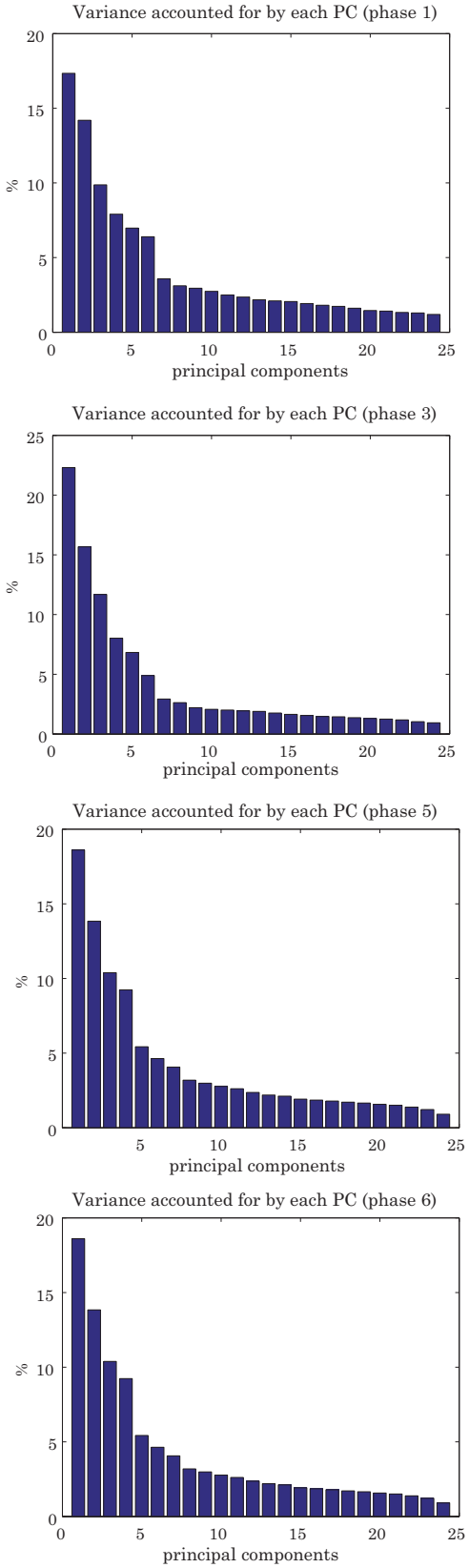


Figure 4.10: Amount of variance accounted for by each principal component, for phases 1, 3, 5 and 6.

For each actuator phase, the number of principal components adopted varies since the principal components retained must account for at least 90% of the cumulative variance. Although there is not an accurate criterion to state a percentage of cumulative variance to be retained for a good representation, a high percentage can ensure that most of the variability is incorporated into the statistical model.

Figure 4.11 show the projections onto the two first principal components of several experiments that come from the undamaged and damaged structure under consideration. It can be clearly observed that no separation of damaged/undamaged can be determined using the scatter plot. These are then two motivating depictions in the sense that with the proposed methodology we will be able to both detect damage in the structure as well as classify it.

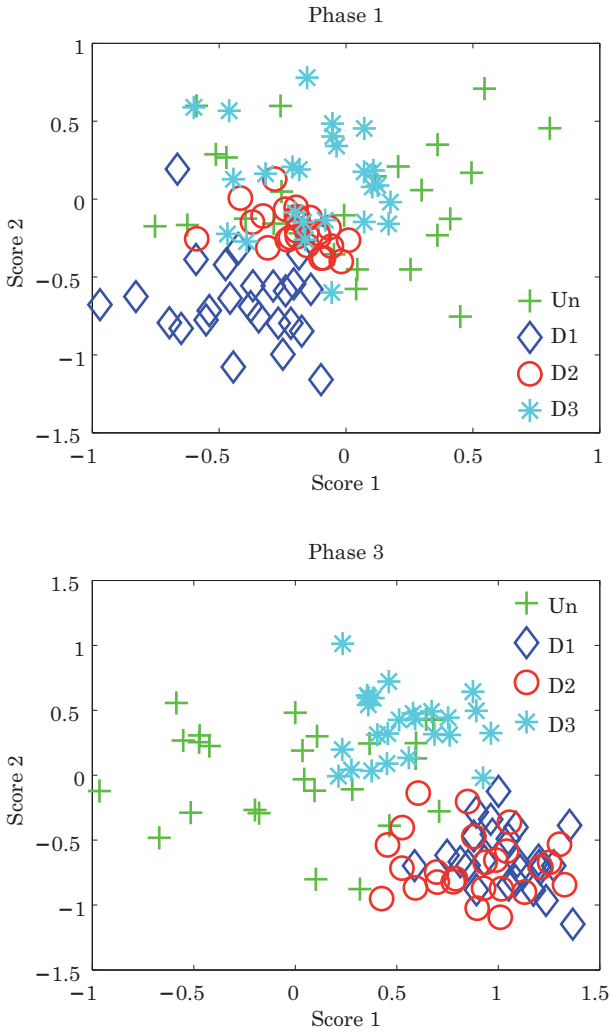


Figure 4.11: Projections onto the two first principal components of several experiments in actuator phases 1 (up) and 3 (down).

Damage detection

After the baseline modeling, the data coming from the structure to be diagnosed is projected onto the PCA model. Then, for each experiment, the feature vector in equation (4.17) formed

by the two damage indices T^2 and Q is defined.

The ability of the proposed method to detect damages in the structure is illustrated in Figures 4.12 to 4.14. In these figures, the affinity of a memory cell from the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose is depicted. The 25 first experiments correspond to data that come from the undamaged structure, while the remainder 75 experiments come from the damaged structure. More precisely, experiments 25 to 50 correspond to damage 1 ($D1$), experiments 51 to 75 to damage 2 ($D2$) and experiments 76 to 100 to damage 3 ($D3$). The pink solid horizontal line delimits the detection threshold (DT_h). It can be clearly observed that experiments with an affinity value less than DT_h –from the damaged structure– are correctly classified as 'damaged'. Similarly, experiments with an affinity value greater than or equal to DT_h –from the healthy structure– are correctly classified as 'healthy'.

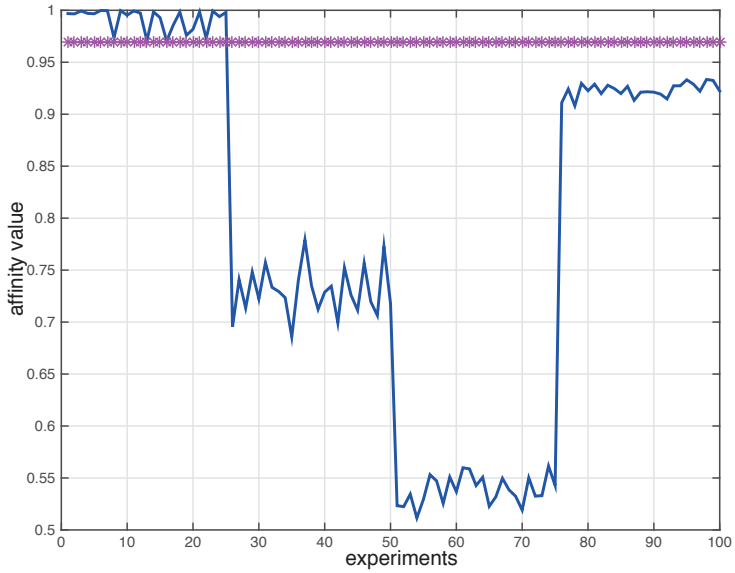


Figure 4.12: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 1).

Damage classification

After the damage detection based on the affinity, the classification is performed by applying the fuzzy c-means clustering algorithm as summarized in Figure 4.7. Figures 4.15 to 4.18 show the result of the damage classification for actuator phases 1, 2, 3 and 6. As it can be seen in these Figures, data is grouped according to its similarity in three clusters, where each data point has a degree of belonging to each cluster inversely proportional to its distance to the center of each cluster.

From Figures 4.12-4.13 (corresponding to phases 1 and 3) the three types of damage can be visually discerned since the related affinities are quite different. This segregation is preserved and even improved in the damage classification in Figures 4.15 and 4.17. However, in Figure 4.14, the affinities for the three sets of experiments are very similar and cannot be precisely separated. In spite of that, and due to the fact that we include the Q index as the first column in the feature matrix in equation (4.27), the separation and classification is quite remarkable.

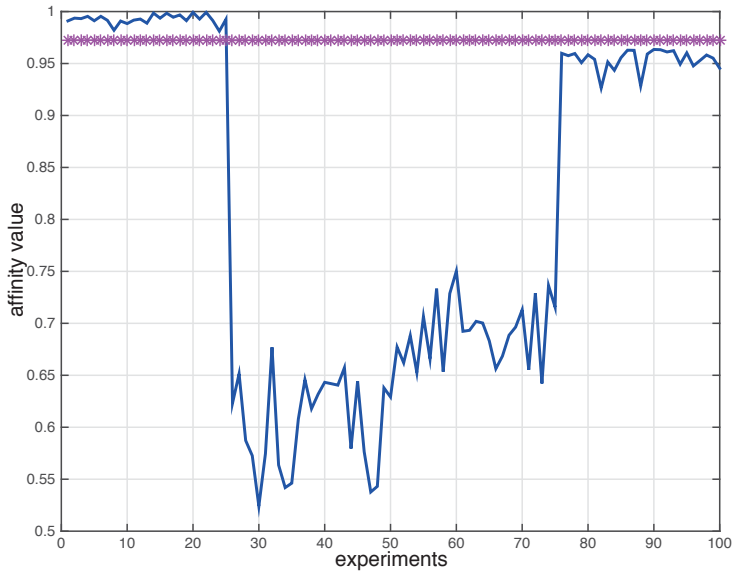


Figure 4.13: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 3).

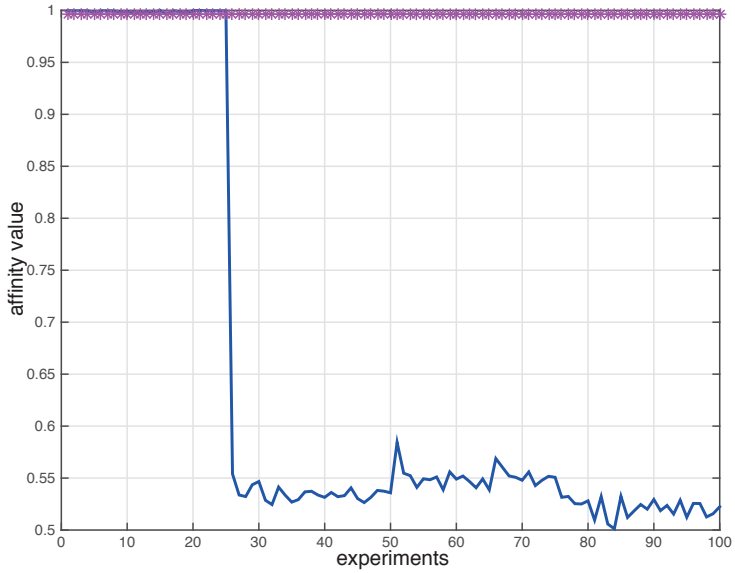


Figure 4.14: Affinity values between a memory cell of the memory cell set of the healthy state (MCSH) and the data coming from the structure to diagnose (phase 6).

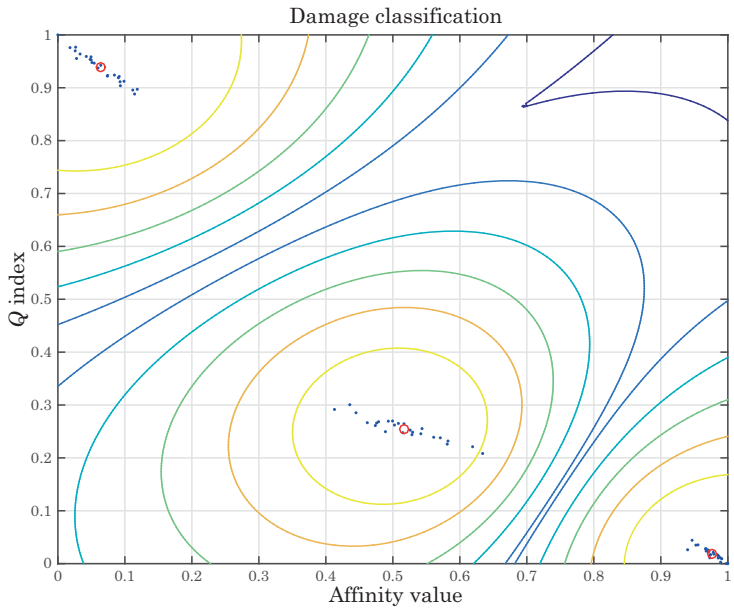


Figure 4.15: Damage classification through the c-means algorithm (phase 1).

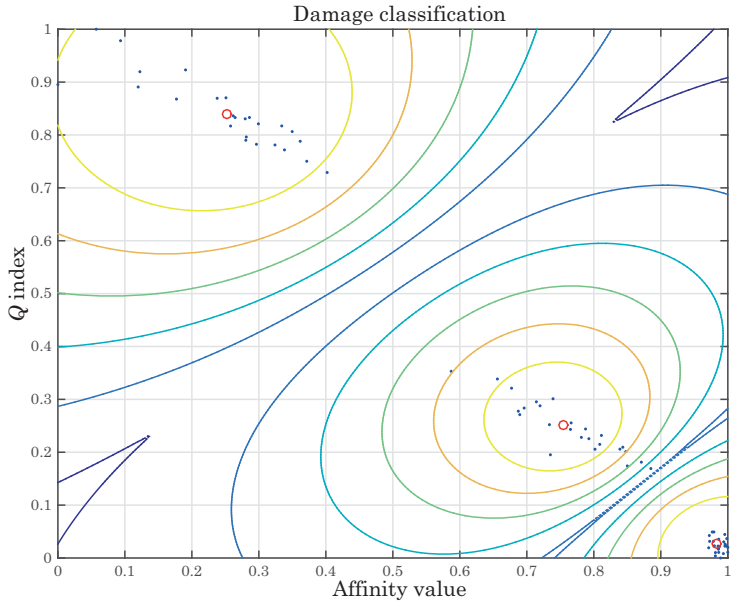


Figure 4.16: Damage classification through the c-means algorithm (phase 2).

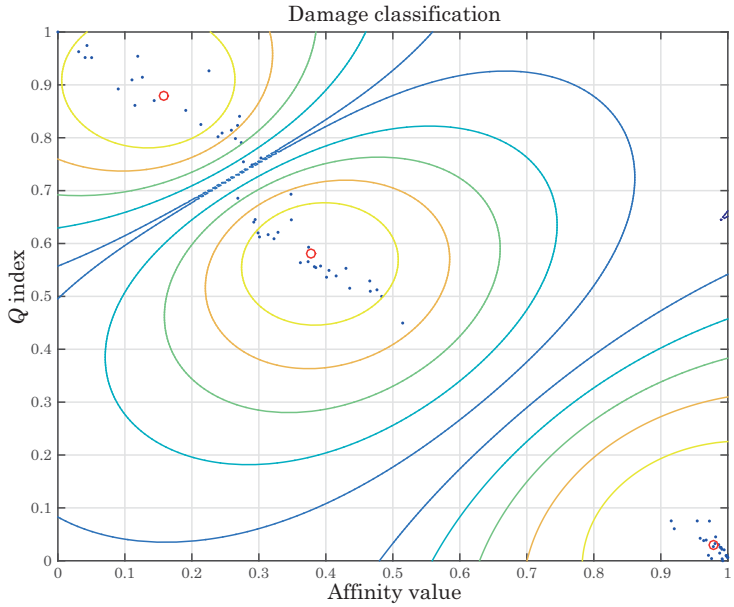


Figure 4.17: Damage classification through the c-means algorithm (phase 3).

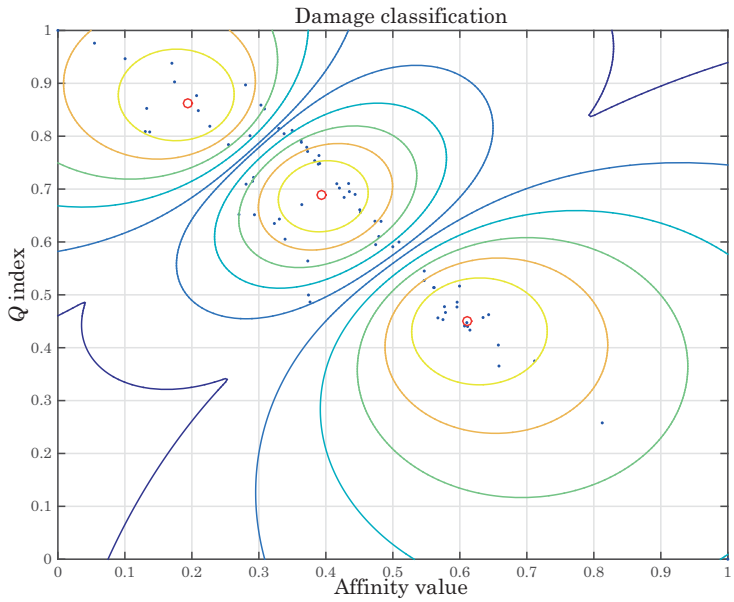


Figure 4.18: Damage classification through the c-means algorithm (phase 6).

4.5 Concluding remarks

In this paper, a new methodology to detect and classify structural changes has been introduced. The methodology is twofold: (i) an artificial immune system (AIS) and the notion of affinity is used for the sake of damage detection and (ii) a fuzzy c-means algorithm is used for damage classification.

One of the advantages of the methodology is the fact that to develop and validate a model is not needed. Additionally, and in contrast to standard Lamb waves-based methods, there is no need to directly analyze the complex time-domain traces containing overlapping, multimodal

and frequency dispersive wave propagation that distorts the signals and difficult the analysis.

The proposed methodology has been applied to data coming from two sections of an aircraft skin panel. The results indicate that the proposed methodology is able to accurately detect damage as well as classify those damages. However, within the proposed methodology, it is not possible to provide a multidamage classification able to identify several simultaneous damages. To ensure the proper performance of the methodology, a study of the effect of changing environmental and operational conditions need to be considered, which is considered as a future work.

Acknowledgements

This work is supported by CICYT (Spanish Ministry of Economy and Competitiveness) through grants DPI2011-28033-C03-01 and DPI2014-58427-C2-1-R. The authors would also like to express their gratitude to the professor Alfredo Güemes for providing the structure used on this paper and for his suggestions in the experimental setup and data acquisition process.

Bibliography

[1] P. J. Costa Branco, J. A. Dente, and R. Vilela Mendes. Using Immunology Principles for Fault Detection. *Industrial Electronics, IEEE Transactions on*, Vol 50, Pages 362–373, 2003.

[2] S. da Silva, M. Dias Júnior, and V. Lopes Junior. Damage Detection in a Benchmark Structure Using AR-ARX Models and Statistical Pattern Recognition. *Journal of The Brazilian Society of Mechanical Science and Engineering*, Vol 29, No 2. Pages 174–184. 2007.

[3] S. da Silva, M. Dias Júnior, V. Lopes Junior, and M. J. Brennan. Structural Damage Detection by Fuzzy Clustering. *Mechanical systems and signal processing*, Vol 22, Pages 1636–1649. 2008.

[4] J. Zhang, K. Worden, and W. J. Staszewski. Sensor Optimisation Using an Immune System Metaphor. In *Proceedings of 26th International Modal Analysis Conference (IMAC)*, 2008.

[5] J. R. V. de Moura Jr, S. Park, V. Steffen Jr, and D. J. Inman. Fuzzy Logic Applied to Damage Characterization Through SHM Techniques. In *Conference Proceedings of the Society for Experimental Mechanics Series*, 2008.

[6] B. Chen. Agent-Based Artificial Immune System Approach for Adaptative Damage Detection in Monitoring Network. *Journal of Network and Computer Applications*, Vol 33, No 6, Pages 633–645. 2010.

[7] D. Tan, W. Qu, and J. Tu. The Damage Detection Based on The Fuzzy Clustering and Support Vector Machine. In *Intelligent System Design and Engineering Application, International Conference on*, (ISDEA) 2010, Vol 2, Pages 598–601, 2010.

[8] S. Casciati. Statistical Approach to a SHM Benchmark Problem. *Journal of Smart Structures and Systems*, Vol 6, No 1, Pages 17–27, 2010.

[9] S. Casciati. Response Surface Models to Detect and Localize Distributed Cracks in a Complex Continuum. *Journal of Engineering Mechanics*, Vol 136, No 9, Pages 1131–1142, 2010.

[10] B. Chen, and C. Zang. Emergent Damage Pattern Recognition Using Immune Network Theory. *Journal of Smart Structures and Systems*, Vol 8, No 1, Pages 69–92. 2011.

[11] Z. Chilengue, J.A. Dente, and P.J. Costa Branco. An artificial immune system approach for fault detection in the stator and rotor circuits of induction machines. *Electric power systems research*, 81:158–169, 2011.

[12] Y. Zhou, S.Tang, C. Zang, and R. Zhou. *Advances in information technology and industry applications lecture notes in electrical engineering*, volume 136, chapter An artificial immune pattern recognition approach for damage classification in structures, pages 11 – 17. Springer Link, 2012.

[13] W. Xiao. *Structural Health Monitoring and Fault Diagnosis based on Artificial Immune System*. PhD thesis, Worcester Polytechnic Institute, 2012.

[14] Z. Liu, Q. Zhou, Q. Chi, Y. Zhang, Y. Chen, and S. Qi. Structural damage detection based on semi-supervised fuzzy c-means clustering. In *Computer science & education (ICCSE), 2014 9th international conference on*, pages 551–556. IEEE, 2014.

[15] Y.Huang, L. Gong, S. Wang, and L. Li. A fuzzy based semi-supervised method for fault diagnosis and performance evaluation. In *IEEE/ASME Inernational conference on advanced intelligent mechatronic (AIM). Besanon, France*, 2014.

[16] H. Sohn, C. R. Farrar, F. M. Hemez, D. Shunk, D. W. Stinemates, and B. R. Nadler. *A Review of Structural Health Monitoring Literature: 19962001*. Los Alamos National Laboratory Report, LA-13976-MS. 2001.

[17] L. N. de Castro and J. Timmis. *Artificial Neural Networks in Pattern Recognition*, chapter Artificial Immune Systems: A Novel Approach to Pattern Recognition, Pages 67–84. University of Paisley, 2002.

[18] A. K. Eroglu, Z. Erden, and A. Erden Bioinspired Conceptual Design (bicd) Approach for Hybrid Bioinspired Robot Design Process. In *Mechatronics (ICM), 2011 IEEE International Conference on*. 2011 , Page(s): 905- 910. Digital Object Identifier: 10.1109/ICMECH.2011.5971243, 2011.

[19] N. Cruz Cortés. *Sistema Inmune Artificial para Solucionar Problemas de Optimización*. Phd thesis, Centro de Investigacin y de Estudios Avanzados del Instituto Politécnico Nacional, 2004.

[20] P. J. Delves, S. J. Martin, D. R. Burton and I. M. Roitt. *Roitt's essential immunology*. Wiley-Blackwell. 2011.

[21] L. N. de Castro and F. J. Von Zuben. *Artificial Immune Systems: Part I Basic Theory and Applications*. Technical Report, School of Computing and Electrical Engineering, State University of Campinas, Brazil, 1999.

[22] D. Trejo Pérez. *Optimización Global en Espacios Restringidos Mediante un Sistema Inmune Artificial*. Master's thesis, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional-Mexixo D.F., 2005.

[23] U. Aickelin, and D. Dasgupta. *Search Metodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, chapter Artificial Immune System, pages 375–399. Springer US, 2005.

[24] A. A. Freitas and J. Timmis. Revisiting the foundations of artificial immune systems for data mining. *IEEE Transactions on evolutionary computation*, 11(4):521–540, 2007.

[25] R. Xiao and T. Chen. Relationships of swarm intelligence and artificial immune system. *Internatonal journal of bio-inspired computation*, 5(1):35 – 51, 2013.

[26] X. Wang, X. Z. Gao, and S. J. Ovaska. Fusion of clonal selection algorithm and harmony search method in optimisation of fuzzy classification system. *International journal of bio-inspired computation.*, 1(1/2):80 – 88, 2009.

[27] V. T. Nguyen, T. T. Nguyen, K. T. Mai, and T. D. Le. A combination of negative selection algorithm and artificial immune network for virus detection. In *Future data and security engineering*, volume 8860, pages 97–106. First international conference, FDSE 2014, Ho Chi Minh city, Vietnam, November 2014.

[28] N. Lay and I. Bate. Applying artificial immune system to real-time embedded systems. In *IEEE Congress on evolutionary computation. CEC 2007*, pages 3743–3750. IEEE, September 2007.

[29] B. Mnassri, E. El Adel, and M. Ouladsine. Fault localization using principal component analysis based on a new contribution to the squared prediction error. In *16th Mediterranean Conference on Control and Automation*, 2008.

[30] A. Hyvarinen, J. Kahunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, INC, 2001.

[31] D. A. Tibaduiza. *Design and validation of a structural health monitoring system for aeronautical structures*. PhD thesis, Department of Applied Mathematics III, Universitat Politècnica de Catalunya, 2013.

[32] D. A. Tibaduiza, L. E. Mujica, and J. Rodellar. Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Structural control and health monitoring*, 20:1303 – 1316, 2013.

[33] D. A. Tibaduiza, L. E. Mujica, A. Güemes, and J. Rodellar. Active piezoelectric system using pca. In *Fifth European Workshop on Structural Health Monitoring*, 2010.

[34] D. A. Tibaduiza, M. A. Torres, L. E. Mujica, J. Rodellar, and C. P. Fritzen. A study of two unsupervised data driven statistical methodologies for detecting and classifying damages in structural health monitoring. *Mechanical Systems and Signal Processing*, 41:467 – 484, 2012.

[35] L.E. Mujica, M. Ruiz, F. Pozo, J. Rodellar, and A. Güemes. A structural damage detection indicator based on principal component analysis and statistical hypothesis testing. *Smart Materials and Structures*, 23(2), 2014.

[36] D.A. Tibaduiza, L.E. Mujica, J. Rodellar, and A. Güemes. Structural damage detection using principal component analysis and damage indices. *Journal of Intelligent Material Systems and Structures*, 2015.

[37] R. Babushka. *Fuzzy modeling for control*. Kluwer Academic Publishers Norwell. 1998.

[38] L. Rokach, and O. Maimon. *Data Mining and Knowledge Discovery Handbook*, chapter Clustering Methods, Pages 321–352. Springer.2005

[39] B. Balasko, J. Abonyi, and B. Feil. *Fuzzy Clustering and Data Analysis Toolbox(for Use with Matlab)*.

[40] MathWorks. *Matlab Fuzzy Logic Toolbox User's Guide*.

[41] Md. E. Karim, F. Yun, and S. P. V. Siva Krishna Madani. *Fuzzy Clustering Analysis* Blekinge Institute of Technology. School of Engineering. Department of mathematics and Science. 2010.

[42] M. S. Yang, P. Y. Hwang, and D. H. Chen. Clustering Algorithms for Mixed Feature Variables. *Journal of Fuzzy Sets and Systems*, Vol 141, Pages 301–317. 2004.

[43] C. Döring, M. J. Lesot, and R. Kruse. Data Analysis With Fuzzy Clustering Methods. *Journal of Computational Statistics and Data Analysis*, Vol 51, Pages 192–214. 2006.

Chapter 5

Conclusions and future research

5.1 Comment and concluding remarks

This thesis shows different data-driven methodologies for damage detection and classification using bioinspired algorithms. Among this algorithms, artificial immune systems, fuzzy clustering and self-organizing maps were considered. These pattern recognition approaches have the advantage of the baseline comparison, which can be related not only to the structure damage but also with the performance of the active sensor through its nominal behaviour. Some conclusions and comments about the results obtained with the development of this thesis are listed below:

5.1.1 Instrumentation and data acquisition

The use of piezoelectric transducers resulted in a viable method to inspect composites and aluminium structures. Some advantages about the inspection with piezoelectric transducers includes high sensitivity to the damages, easy installation and operation since relatively long-distance inspection which can be covered with low attenuation and reduced price compared with other sensors. Additionally, these kind of sensors can be used as passive or active sensors since they can work both as a sensor or as actuators. Some limitations in the use of piezoelectrics for inspection processes are: low output, this means that it is necessary to use an additional circuit to amplify the excitation/collected signals and high impedance output.

In this thesis, several PZTs were used to conform a piezoelectric active system which corresponds to an inspection system where each PZT was used as an actuator and as a sensor in different actuation phases. It is necessary to remark that the use of the active piezoelectric system has allowed to prove the PCA based methodology and validate the results from different actuation phases to identify a healthy structural state or the presence of damages at the surface by changing the PZT used as actuator.

This methodology has been used by the previous works in the CoDAlab group and still working with excellent results. Some disadvantages in the use of piezoelectric include debonding, cristal cuts, among others. In this sense the use of another sensors need to be explored, however, with the development of this thesis, a methodology to detect and classify typical damages in piezoelectric sensors was developed.

The use of different actuation phases allowed to obtain different concepts about the presence of a damage that were analyzed by plotting each result and analyzing the information in each

plot. In the same way, with the aim of classification, the data fusion allowed to obtain a unified result.

Finally, it is worth remarking that the use of the active piezoelectric system allowed to prove the PCA based methodologies and validate the results from different actuation phases to identify a healthy structural state or the presence of damages at the sensors into the sensor network.

5.1.2 Data pre-processing

This step is one of the most important in the developed methodologies, since it adds some desirable features to the data used by the several approaches. Some of the advantages in the use of data processing are the following:

The use of unfolding data has allowed to process data in two dimensions by organizing the information from three dimensions (experiment $\times$ time $\times$ sensor) to two dimensions (experiments $\times$ (time)·(sensor)). This allows to simplify the analysis of the information by each actuation phase.

The normalization is the procedure to standardize data set, so that signal changes –from different sources and collection points– as well as changes in the operational and environmental conditions can be scaled using the mean and standard deviation of all measurements. In this thesis, the unfolding and the normalization (also known as group scaling) is used as in previous works by the advisors.

Although DWT it was not used in all the approaches, its use presented some advantages in the development of the algorithm for temperature analysis, since gives some benefits such as noise reduction, dimension reduction and feature extraction. It can be an advantage if real-time applications are needed because the analysis is performed to some coefficients that represent a reduced number of data.

5.1.3 Structural damage detection and classification

The damage identification problem was divided in two: (i) damage detection and (ii) damage classification. In addition, the methodology was used to detect and classify faults in sensors. Specifically, the approaches were tested in structures instrumented with piezoelectric sensors.

Although the damage detection and classification processes have been addressed by several authors, this thesis contributes to these problems by proposing the use of robust baselines or by the use of artificial immune systems. One of the advantages of the developed methodologies is the fact that to develop and validate a model is not needed. Additionally, and in contrast to standard Lamb waves-based methods, there is no need to directly analyze the complex time-domain traces containing overlapping, multimodal and frequency dispersive wave propagation that distorts the signals and difficult the analysis. Results shown that different actuation phases present different results. In this sense, the following conclusions can be summarized:

It was found that the number of principal components to carry out the damage detection is related to the dispersion of the variance. Overlaps were found when just two principal components were used, this means a good representation requires more than two scores to discern between the different structural states. In this sense, the use of the damage indices allowed to improve the damage detection process discerning features in each state of the structure.

The methodologies cannot discern multiple damages at the same time. That is, multiple damages are detected as a new damage, this is because the algorithms use data from the PZTs

collecting vibration signals.

Although the damage identification process was performed under laboratory conditions, in this thesis, the effects of changes in the temperature were explored using different kind of structures with excellent results.

With respect to the use of some realistic delamination and debonding problems, it was shown that the proposed approach helps in generating a robust damage detection and classification system to handle different kind of damages under different temperatures.

Some conditions such as temperature variations were evaluated to determine its influence in the damage detection process. It was found that temperature change the results in the detection because the collected signal changes. To solve this problem, a robust baseline was proposed with excellent results.

With respect to the use of the SOM, it is possible to define the following conclusions: Results with the cluster map and the U-matrix showed some differences between them. In spite of these differences, both results allowed to detect and classify the damages. The U-matrix presented an advantage compared with the cluster map because it gives a easy-understand representation due to the inclusion of the boundaries between the different states of the structure. In the classification process the inclusion of data from a different structural state provide a new cluster in the training process. This is because the map creates different zones according to its similarity. The outliers are better isolated in the U-matrix representation because it allows the identification of each structural state under evaluation.

With respect to the use of artificial immune systems, some advantages and conclusions can be summarized: The encoding of the information is very easy and intuitive since healthy state represents the antibodies and the different states (damaged or not) are represented by the antigens.

The use of affinity values is a good way to determine the differences between the antigens and the antibodies. This difference was not equal at each actuation phase. However, the use of all the plots allowed to determine the presence of damages.

The evolution and the use of immune learning algorithms make it possible for the classifier to generate a high quality memory cell to recognize various structure damage patterns.

The use of AIS and fuzzy clustering allowed the data fusion and the organization of the information from all the actuation phases to classify all the structural states.

5.1.4 Sensor fault detection and classification

The same methodologies developed for damage detection can be used to detect and classify all the sensor faults in spite of the differences between cristal cuts, cristal removal and debonding. Better results were obtained with the use of the damage indices and SOM. One probable reason is because only two components are not sufficient to discriminate differences between the different scenarios. One of the disadvantage in the methodology is the big quantity of plots to analyze, in this case two plots by each actuation phase. However this work shows that damage indices plots can be evaluated without the use of the score plots and fused into a cluster map or a U-Matrix.

5.2 Future research

Some issues still remain open and were not covered by this thesis. The following subjects are outlined as future works:

- To validate the algorithms by using variations of different environmental and operational conditions. In this sense, humidity and a wide range of temperature variations need to be used to test the algorithms.
- To use different kind of sensors. Since the algorithms are based on pattern recognition approaches, it is expected that the same algorithms can work with minor changes.
- To evaluate the best sensor distribution to detect damages. The knowledge of an optimized distribution of the sensors to install in a structure can help to increase the damage detection process.
- To evaluate the algorithms in more complex structures and in-service structures. Although the developed methodologies presented good results in the detection and classification of damages and in the sensor fault detection, the methodology need to be evaluated in-service structures under real operational and environmental conditions.
- Some of the of damages used in this thesis are not very realistic and for a full validation of the method, it is necessary to test structures with additional real damages to the explored in this thesis.

Appendix A

Publications

During the period of the development of this thesis, as result of this research, as well as result of collaborations with other groups and the research stay performed by the author, were obtained the next contributions to books, journals and relevant works in conferences:

A.1 Book chapters

1. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Structural damage assessment using an artificial immune system. In: Emerging Design Solutions in Structural Health Monitoring Systems, IGI-Global, 2015, doi: 10.4018/978-1-4666-8490-4.ch005.

2. M.A. Torres Arredondo, D.A. Tibaduiza, I. Buethe, L.E. Mujica, **M. Anaya**, J. Rodellar, C.P. Fritzen. Methodologies of damage identification using non-linear data-driven modelling. In: Encyclopedia of Information Science and Technology, IGI-Global, 2015, doi: 10.4018/978-1-4666-5888-2.ch093.

3. M.A. Torres Arredondo, D.A. Tibaduiza, I. Buethe, L.E. Mujica, **M. Anaya**, J. Rodellar, C.P. Fritzen. Validation of damage identification using non-linear data-driven modelling. In: Encyclopedia of Information Science and Technology, IGI-Global, 2015, doi: 10.4018/978-1-4666-5888-2.ch094.

A.2 Journals

1. **M. Anaya**, D.A. Tibaduiza, M.A. Torres-Arredondo, F. Pozo, M. Ruiz, L.E. Mujica, J. Rodellar and C.P. Fritzen. Data-driven methodology to detect and classify structural changes under temperature variations. Smart Materials and Structures, 23(4):045006, 2014, doi:10.1088/0964-1726/23/4/045006.

2. **M. Anaya**, D.A. Tibaduiza, F.Pozo. Detection and classification of structural changes using artificial immune systems and fuzzy clustering. International Journal on Bio-Inspired Computation, to appear.

3. **M. Anaya**, D.A. Tibaduiza, F.Pozo. A bioinspired methodology based on an artificial immune system for damage detection in structural health monitoring. Shock and Vibration

(Special issue on Structural Dynamical Monitoring and Fault Diagnosis), 2015, article ID 648097, 15 pages, 2015, doi:10.1155/2015/648097.

4. D.A. Tibaduiza, **M. Anaya**, R. Castro, E. Forero, F. Pozo. A sensor fault detection methodology in piezoelectric active systems used in structural health monitoring applications. Accepted for publication in the Journal conference Series: Materials Science and Engineering (MSE). IOP-publishing.

5. E. Forero, D.A. Tibaduiza, **M. Anaya**. Detection and characterization of defects in moving parts of wind turbines. Accepted for publication in the Journal conference Series: Materials Science and Engineering (MSE). IOP-publishing.

A.3 Conferences

1. **M. Anaya**, D.A. Tibaduiza, F. Pozo. Artificial immune system (AIS) for damage detection under variable temperature conditions. 8th European Workshop on Structural Health Monitoring (8EWSHM), Bilbao (Spain), 2016.

2. **M. Anaya**, D.A. Tibaduiza, , E. Forero, R.Castro, F. Pozo. An acousto-ultrasonics pattern recognition approach for damage detection in wind turbine structures. 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA), Bogotá (Colombia), 2015.

3. **M. Anaya**, D.A. Tibaduiza, M.A. Torres, F. Pozo. Principal component analysis and self-organizing maps for damage detection and classification under temperature variations. 10th International Workshop on Structural Health Monitoring (IWSHM), Stanford (CA, United States of America), 2015.

4. D.A. Tibaduiza, M.A. Torres, **M.Anaya**. Structural Health Monitoring of Wind Turbine Blades Using Statistical Pattern Recognition. 10th International Workshop on Structural Health Monitoring (IWSHM), Stanford (CA, United States of America), 2015.

5. D.A. Tibaduiza, **M. Anaya**, R. Castro, E. Forero, F. Pozo. A sensor fault detection methodology in piezoelectric active systems used in structural health monitoring applications. 2nd International Congress of Mechanical Engineering and Agricultural Science, Bucaramanga (Colombia), 2015.

6. E. Forero, D.A. Tibaduiza, **M. Anaya**. Detection and characterization of defects in moving parts of wind turbines. 2nd International Congress of Mechanical Engineering and Agricultural Science, Bucaramanga (Colombia), 2015.

7. **M. Anaya**, D. Tibaduiza, F. Pozo. Data driven methodology based on artificial immune systems for damage detection. 7th European Workshop on Structural Health Monitoring, Nantes (France), 2014.

8. **M. Anaya**, D. Tibaduiza, F. Pozo. Structural Damage Classification using Artificial Immune Systems and Fuzzy Clustering. 6th World Conference on Structural Control and Monitoring, Barcelona (Spain), 2014.

9. **M. Anaya**, D. Tibaduiza, F. Pozo. Artificial Immune Systems for Damage Detection. 6th World Conference on Structural Control and Monitoring, Barcelona (Spain), 2014.

10. D.A. Tibaduiza, L.E. Mujica, **M. Anaya**, J. Rodellar, A. Güemes. Principal component analysis vs independent component analysis for damage detection. 6th European Workshop on Structural Health Monitoring, Dresden (Germany), 2012.

11. D.A. Tibaduiza, L.E. Mujica, **M. Anaya**, J. Rodellar, A. Güemes. Independent component analysis for detecting damage on aircraft wing skeleton. EACS 2012-5th European Conference on Structural Control, Genoa (Italy), 2012.