

# Wireless Intelligent Sensing Devices

DTI Project No: TP/2/SC/6/I/10347

Part of the DTI Technology Programme  
Sensors for Industrial and Environmental Applications Theme

## FEATURE EXTRACTION

<b>Date:</b>	29 <sup>th</sup> November 2006
<b>WISD Document Number:</b>	WISD-DOC-0008
<b>Originating Company:</b>	University of Bristol
<b>Originating Reference:</b>	(Report Number D0008)

<b>Prepared by:</b>	P. J Escamilla-Ambrosio
---------------------	-------------------------

<b>Approved by:</b>	N.A.J. Lieven
---------------------	---------------

**WESTLAND**  
an AgustaWestland Company

**conekt**  
powered by IBM

 University of  
BRISTOL

 SEA



## TABLE OF CONTENTS

1	Introduction	1
1.1	Structural health monitoring	2
1.2	Wireless Intelligent Sensing Devices	3
1.3	Feature extraction algorithms for wireless intelligent sensing devices	4
2	Soft computing feature extraction algorithm	6
2.1	Wavelet transform theory	7
2.1.1	Continuous wavelet transform	8
2.1.2	Discrete wavelet transform	8
2.2	Fuzzy logic theory	12
2.3	Feature extraction algorithm	17
3	Soft computing feature extraction algorithm applied to tie bar data	21
3.1	Algorithm implementation	22
3.2	Feature extraction and comparison analysis results	23
3.2.1	Comparison analysis by test	24
3.2.2	Comparison analysis by tie bar	28
3.3	Analysis of results	29
3.4	Validation of the SCFEA	30
3.5	Tie bar critical degradation detection	32
3.6	Real time simulation	34
4	Soft computing feature extraction algorithm applied to pitch link data	34
4.1	Algorithm implementation	35
4.2	Real time simulation	38
5	Conclusions and future work	39
	References	40

# Feature Extraction

## Abstract

Structural health monitoring is the process of implementing a damage identification strategy for aerospace, civil and mechanical engineering infrastructure. Under this context, feature extraction is the process of identifying damage-sensitive information from measured data. Feature extraction is an essential component of a SHM system needed to convert raw sensor data into useful information about the structural health condition. The need for robust health monitoring and prognosis of components in remote or difficult to access locations is driving the advancement of sensing hardware and processing algorithms. The Wireless Intelligent Sensing Devices (WISD) research project aims to attend to this need. In this document a feature extraction algorithm, referred to as soft computing feature extraction algorithm, is developed as part of the WISD project to extract damage-sensitive information from measured response data of tie bar and bearing system of pitch link components of the main rotor hub of a Lynx Helicopter. The feature extraction algorithm is based on a combining of discrete wavelet transform theory and fuzzy logic theory. The results of applying the proposed feature extraction approach to tie bar and pitch link data are presented. Additionally, methods for pattern recognition and critical degradation detection of tie bar and critically worn detection of pitch link are proposed. Results show that the proposed algorithms are capable of detecting critical degradation of tie bar and have the ability to discriminate between unworn (undamaged) and worn (damaged) pitch link bearings.

## 1. Introduction

This report is generated by the University of Bristol to describe research work carried out in the area of Feature Extraction as the fourth deliverable from the Wireless Intelligent Sensing Devices (WISD) research project. The University of Bristol (UB) is member of the WISD consortium, which other members are Westland Helicopters Ltd (WHL), TRW Conekt (TRW), and Systems Engineering & Assessment Ltd (SEA). The WISD research project is supported by The Department of Trade and Industry (DTI).

The need for robust health monitoring and prognosis of components in remote or difficult to access locations is driving the advancement of sensing hardware and processing algorithms. The Wireless Intelligent Sensing Devices (WISD) research project aims to attend to this need. Therefore, the main goal of the WISD project is to advance the development of autonomous, self-powered, wireless sensors with built-in intelligence, in order to provide an accurate health state and life prediction of engineered structures. It is intended that a WISD will then only transmit information when a feature of the state of health of the monitored structure has altered and requires attention or maintenance. Thus, it is expected that an array of WISDs will provide an efficient alternative to the current approach of streaming raw data back to a central monitoring unit. The target application for the sensing system developed by the WISD project is health monitoring of the main rotor hub of a Lynx Helicopter.

In this report the main findings of ongoing research in the area of feature extraction and statistical model development (pattern classification), under the context of the WISD research project, are presented. Thus, in the remainder of this section some background concepts are presented. First the concept of structural health monitoring is reviewed, then a definition of what is considered a WISD is presented together with a review of data interrogation procedures used in the past in wireless sensor devices. Then, in section 2, after providing a short description of wavelet theory and fuzzy logic theory, a feature extraction algorithm referred to as the soft computing feature extraction

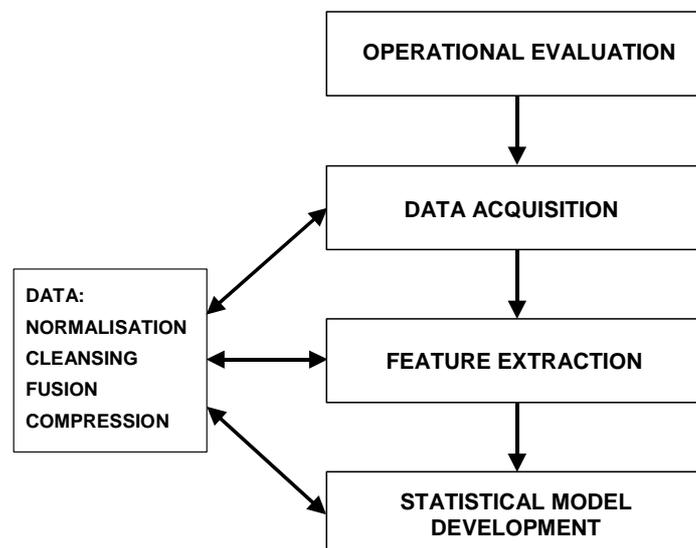
algorithm, which combines these two technologies, is proposed. The results of the algorithm applied to tie bar data and pitch link data components of the main rotor hub of a Lynx Helicopter are presented in sections 3 and 4, respectively. Finally, conclusions to this work are given in section 5.

## 1.1 Structural health monitoring

Structural health monitoring (SHM) can be defined as the process of implementing a damage identification strategy for aerospace, civil and mechanical engineering infrastructure [Farrar et al, 2001]. The goal of implementing a SHM system is to improve the safety and reliability of engineering structures by detecting damage before it reaches a critical state. Therefore, the process of implementing a SHM system involves the observation of a structure or mechanical system over time using periodically spaced measurements, the extraction of damage-sensitive features from these measurements, and the statistical analysis of these features to determine the current state of health of the system. The output of this process is periodically updated information regarding the ability of the structure or mechanical system to continue to perform its desired function in light of the inevitable ageing and degradation resulting from operational environments.

The implementation of a SHM system can be summarised into the four-step flow chart shown in figure 1 [Farrar et al, 2001] [Farrar et al, 2006]. A short description of each one of these processes is given below.

*Operational evaluation:* This process defines and quantifies the damage that is to be detected and describes the benefits to be gained from implementing the SHM system. This process also sets what unique aspects of the system will be monitored and how to perform the monitoring as well as specifying the features of the damage to be detected.



**Figure 1:** Flow chart for implementing a SHM system.

*Data acquisition:* This component of the SHM process involves selecting the excitation methods, the sensor types, number and locations, and the hardware platform for data acquisition, storage, processing, and communication. It is necessary to remark that the data acquisition and sensing systems do not measure damage. Rather, they measure the response of a system to its operational and environmental loading or the response to inputs from actuators embedded with the sensing

system. They deliver raw measurement data. However, depending on the sensing technology selected and the type of damage to be identified, the sensor readings include features that may be more or less directly correlated to the presence and location of damage.

*Feature extraction:* Feature extraction is the process of identifying damage-sensitive information from measured data. A damage-sensitive feature is some quantities extracted from the measured system response data that is correlated with the presence of damage in a structure [Farrar et al, 2006]. The main objective of the feature extraction process is to extract damage-sensitive features that change in some consistent manner with increasing damage level. Ultimately, the goal is to distinguish a damaged structure from an undamaged one based on the extracted features in a robust and accurate manner. Two alternative feature extraction methods have been mainly proposed in the SHM literature; model based and waveform based. The model based feature extraction method consists on fitting some model, either physics based or data based, to the measured system response data. The parameters of these models or the predictive errors associated with these models then become the damage-sensitive features. Alternatively, the waveform based approach extract features by directly comparing the sensor waveforms or spectra of these waveforms.

*Statistical model development:* This process is concerned with the implementation of the algorithms that analyse the distributions of the extracted features in order to determine the damage state of the structure. The algorithm used to perform this task can be categorised into three types: (1) Group Classification, (2) Regression Analysis, and (3) Outlier Detection. The selection of the appropriate algorithm to use depends on the data available. For example, algorithms performing supervised learning can be applied when examples of data are available from damaged and undamaged structures. If data were available only from the undamaged structure, then an algorithm implementing unsupervised learning would be more adequate. The statistical models are typically used to answer a series of questions regarding the presence, location, type and extent of damage.

Inherent in the data acquisition, feature extraction and statistical model development sections of the SHM process are data normalisation, cleansing, fusion and compression [Farrar et al, 2006]. Under the context of SHM, data normalisation is the process of separating changes in sensor reading caused by damage from those caused by varying operational and environmental conditions. Data cleansing is the process of selectively choosing data to pass on to, or reject from, the feature selection process. Data fusion is the process of combining information from multiple sensors in an effort to enhance the fidelity of the damage detection process. Data compression is the process of reducing the dimensionality of the data, or the feature extracted from the data, in order to facilitate an efficient storage of information and to enhance the statistical quantification of these parameters. These four activities can be implemented in either hardware or software and usually a combination of the two approaches is used.

The two processes of feature extraction and statistical model development for feature classification commonly are referred together to as data interrogation procedures. They are the essential components of a SHM system needed to convert the sensor data into useful information about the structural health condition. Furthermore, to successfully implement a SHM strategy, the data acquisition system will have to be developed in conjunction with these data interrogation procedures.

## **1.2 Wireless Intelligent Sensing Devices**

In general terms, an intelligent sensing device includes a memory for data storage; a radio frequency communicator configured to receive signals from and transmit signals to an external device; a processor; one or more sensors; and a power supply. The key component of a WISD is

the onboard processor, which allows the device to perform its own local data interrogation tasks, only transmitting the results. Consequently, a significant reduction in power consumption is achieved (It has been demonstrated that one byte of data transmission consumes the same energy as approximately 11000 cycles of computation using low powered DSPs [Tanner *et al.*, 2003]).

The main technological advancement in a WISD is its ability to make decisions, not simply stream raw data. Accordingly, the output of a WISD is a feature of the structure or a health state indicator rather than raw data itself. Furthermore, by being self-powered and equipped with a wireless transmitter, a WISD is able to act autonomously or be triggered remotely to provide an intelligent assessment of the state of health of the monitored structure.

### 1.3 Feature extraction algorithms for wireless intelligent sensing devices

Recently, research studies have been carried out in order to develop feature extraction algorithms for SHM capable of being implemented in the onboard processor of wireless intelligent sensing devices. Tanner *et al* [2003] implemented an SHM algorithm in an off-the-shelf wireless sensing and data processing hardware known as “Motes”, which were developed at the University of California, Berkeley. The Mote system consist of modular circuit boards integrating sensors, microprocessor, A/D converters, and wireless transmitter all of them powered by two AA batteries [Kurata *et al.*, 2005]. In the implemented SHM algorithm the cross-correlation coefficient between the time responses measured from two accelerometers mounted across a joint in a demonstration structure was used as the feature for damage detection. Thus, damage is detected using a statistical process control approach [Montgomery, 2005] [Sohn *et al.*, 2000]. Firstly, in a training phase, data from a known healthy condition of the structure were used to establish the upper and lower bounds of the statistical process control. The control limits were set at  $\mu \pm 1.5\sigma$ , based on the cross-correlation coefficient sample mean  $\mu$  and standard deviation  $\sigma$ . The values of the cross-correlation coefficients were calculated using a recursive algorithm on the wireless sensor device and broadcasted to a base station connected to a PC. After the bounds were calculated on the PC, they were hard-coded back onto the processor on the wireless sensor. Secondly, in a monitoring phase, the cross-correlation coefficients were calculated from newly measured acceleration time signals and checked against the previously determined control limits to determine if any of the cross-correlation values was an outlier. Finally, damage is declared based on the frequency of occurrence of outliers. A binary result could then either be shown on the motes’ LED or transmitted wirelessly to a base station. The whole process proved to be, however, very limited, allowing only the most rudimentary data interrogation algorithms to be implemented and not in a completely autonomous operational mode.

Lynch *et al* [2004] developed a hardware system to implement SHM algorithms using off-the-shelf components. The hardware platform includes sensing circuits and a wireless transmission unit coupled with a computational core incorporating two microcontrollers working in a master-slave configuration for power efficiency. This hardware system allows a decentralised collection, analysis and broadcast of a structure’s health. The feature extraction and damage detection algorithm implemented in this hardware platform is the statistical time-series approach proposed by Sohn and Farrar [2001]. In this approach the time histories of vibration signals of the analysed structure in its undamaged state are measured under a variety of environmental and operational conditions. Then, after normalisation, an autoregressive (AR) model of dimension  $p$  (denoted as AR( $p$ )) is fitted to the measured data:

$$y_k = \sum_{i=1}^p b_i^y y_{k-i} + r_k^y \quad (1)$$

where  $y_k$  denotes the response of the structure at sample index  $k$ ,  $b_i^y$  are coefficients on the previous observations  $y_{k-i}$ , and  $r_k^y$  is the AR model residual error term. As it is assumed that the residual error  $r_k^y$  of the AR model is influenced by the unknown input to the system, a second time-series model, an autoregressive with exogenous inputs (ARX) model of dimension  $a$  and  $b$  (denoted as ARX( $a, b$ )) is adopted to model the relationship between the residual error and the measured response of the system:

$$y_k = \sum_{i=1}^a \alpha_i y_{k-i} + \sum_{j=0}^b \beta_j r_{k-j}^y + \varepsilon_k^y \quad (2)$$

where  $\alpha_i$  and  $\beta_j$  are coefficients on past measurements and the residual error of the AR model, respectively. The coefficients of the AR-ARX time series models and the standard deviation of the residual error of the fitted ARX models form a database (denoted by the superscript DB) of baseline models describing the structure in its undamaged state. The residual of the ARX model,  $\varepsilon_k^y$ , is the damage sensitive feature used to detect the existence of damage in the structure.

To detect damage, after measuring the response of the structure  $y_k$  in an unknown state (damage or undamaged), an AR model is fitted. The coefficients of this AR model are then compared to the library of baseline AR-ARX coefficients. The closest AR-ARX model pair is selected from the library based on the Euclidian distance,  $D$ , of the newly derived AR model and the database AR model coefficients,  $b_i^y$  and  $b_i^{DB}$ , respectively. This is:

$$D = \sum_{i=1}^p (b_i^{DB} - b_i^y)^2 \quad (3).$$

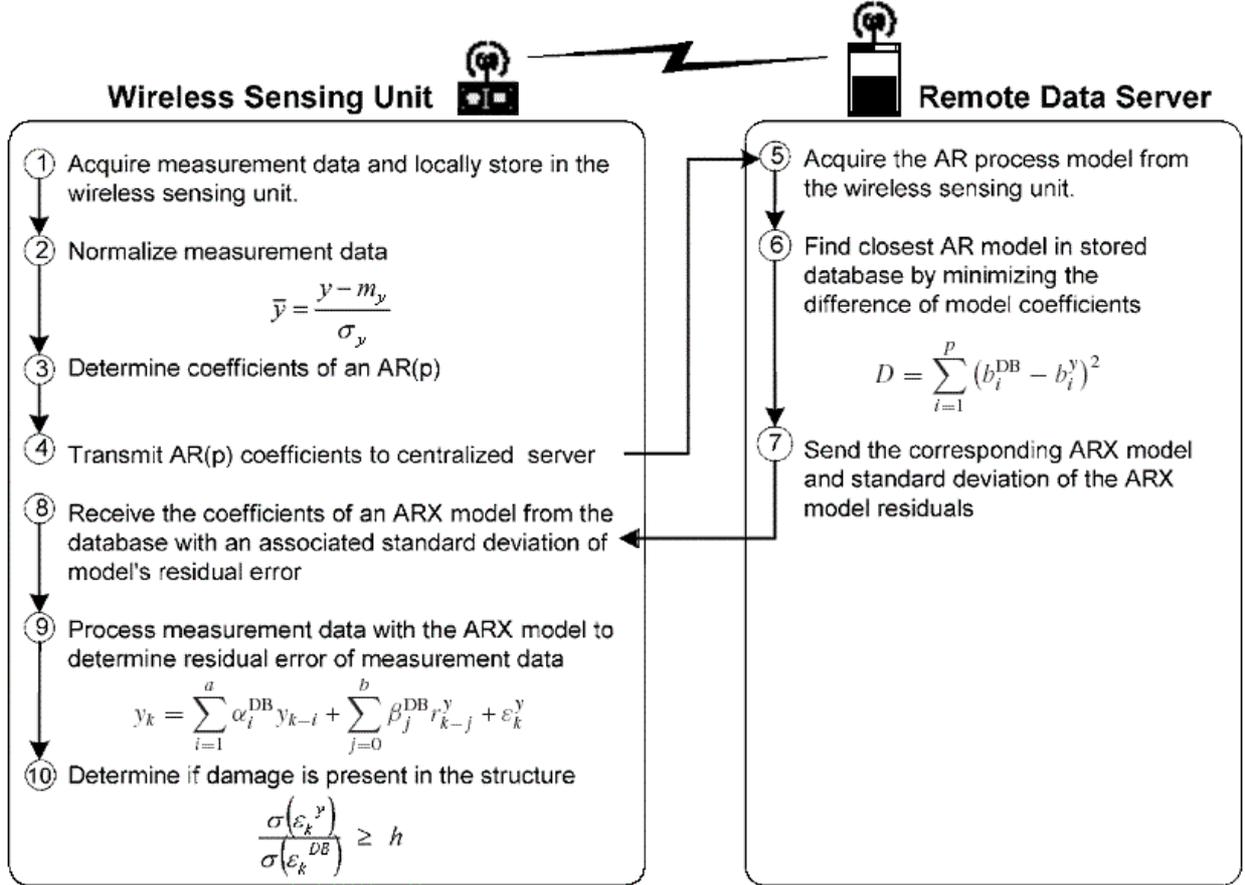
Thus, if no structural damage is experienced and the operational conditions of the two models are close to one another, the selected AR model from the database will closely approximate the measured response. On the contrary, if damage has been sustained by the structure, even the closest AR model of the database will not approximate the measured structural response well. Therefore, the measured response of the structure in the unknown state,  $y_k$ , and the residual error of the fitted AR model,  $b_k^y$ , are substituted into the database ARX model to determine the residual error,  $\varepsilon_k^y$ , of the ARX model:

$$y_k = \sum_{i=1}^a \alpha_i^{DB} y_{k-i} + \sum_{j=0}^b \beta_j^{DB} r_{k-j}^y + \varepsilon_k^y \quad (4)$$

If the structure is in a state of damage, the standard deviation  $\sigma(\varepsilon_k^y)$  of the ARX model residual,  $\varepsilon_k^y$ , will vary from the standard deviation  $\sigma(\varepsilon_k^{DB})$  of the ARX model residual corresponding to the undamaged structure,  $\varepsilon_k^{DB}$ . In particular, damage can be identified when the ratio of the standard deviation of the model residual error exceeds a threshold value,  $h$ , established from good engineering judgment [Sohn and Farrar 2001]:

$$\frac{\sigma(\varepsilon_k^y)}{\sigma(\varepsilon_k^{DB})} \geq h \quad (5)$$

The statistical time-series damage detection method was implemented in the proposed wireless sensing unit as is shown in figure 2. Given the memory limitations of the wireless sensing unit, storage of a database of AR and ARX coefficients was done using a remote data server. The wireless sensing unit is primarily responsible for the determination of AR model coefficients as well as processing the data through the ARX model that is obtained from the remote server. The wireless sensing unit, after calculating the ARX residual error, makes the ultimate decision if damage is potentially present in the system within the vicinity of its respective node.



**Figure 2:** Implementation of the statistical time-series feature extraction and damage detection approach in a wireless sensor unit [Lynch *et al* 2004].

## 2. Soft computing feature extraction algorithm

It is known that current helicopter rotors spin at near constant revolutions per minute (RPM) throughout a flight mission. Consequently, it is expected that signals coming from monitored components of the main rotor hub will be periodic signals. In addition, data signals provided by Westland Helicopters Ltd from results of fatigue testing of tie bars, components of the main rotor hub of a Lynx Helicopter, show a periodic (or cyclic) behaviour. Hence, in this work it is assumed that the measured signals from which damage sensitive features are going to be extracted are periodic signals with a known period (or frequency).

A feature extraction approach, referred to as soft computing feature extraction algorithm (SCFEA), has been proposed as part of the WISD research project in order to identify damage sensitive information from periodic signals. This approach has been developed inspired on the work of Li *et al* [2004]. The SCFEA combines wavelet transform theory and fuzzy logic theory. The general idea explored in the proposed approach is as follows, if a feature vector can be extracted which

represents the characteristics of a cycle or a series of cycles of a signal, then this feature vector can be used to perform comparisons with a cycle or a series of cycles of the signal obtained over different periods of time. This in turn will allow the possibility to assess how the signal is evolving until failure is reached and this outcome can be used for pattern recognition and damage detection. In other words, the main postulate is that the change of the dynamic behaviour of the system being monitored can be expressed in terms of changes in the feature vectors extracted from every cycle (or series of cycles) of the measured signals and compared over time.

The effective content of information in a signal is usually given in its entirety (low frequency) or at slight positions (high frequency). Thus, if two signals (or cycles of a signal) are apparently different, then their traits can be extracted in very different features. But, if the two signals (or cycles of a signal) are approximately the same, then their features should be very similar. Hence, the feature of a signal to be extracted should be sensitive enough to have effective information, and should be robust enough to tolerate noise and distortion. However, sensitivity and robustness are mutually exclusive in nature. In that sense the discrete wavelet transform is used as a tool to decompose a signal into approximation and detail signals, associated with low and high frequencies. While, due to its ability to tolerate imprecision and uncertainty, fuzzy sets are used to provide a robust representation for the signal information. The concept of fuzzy sets is used then to serve as a bridge between sensitivity and robustness in feature extraction for a signal. The sensitivity in feature extraction of the SCFEA can be adapted by tuning or increasing the number of fuzzy sets defined for the detail and approximation signals. The output of the SCFEA will be a feature vector that can be used for pattern classification.

In the next sections the theoretical background used to develop the SCFEA is reviewed. First, an overview of wavelet transform theory is given. Then, the main concepts of fuzzy logic are presented. After that, the proposed feature extraction algorithm is presented.

## **2.1 Wavelet transform theory**

Wavelet transform theory was originally developed in the 1980s at about the same time by mathematicians and seismologists as a new tool for the frequency analysis of geophysical signals [Goupillaud *et al.*, 1984] [Grossmann and Morlet, 1984] [Daubechies, 1988]. Over the last two decades, due to its many useful features, the wavelet transform has been used in a broad range of applications, such as seismic records analysis, image coding, electrocardiogram analysis, feature extraction, pattern recognition, voice processing, image processing, signal denoising, and data compression.

The wavelet transform (WT) can be viewed as an alternative to the traditional Fourier transform (FT) for the analysis of signals. Unlike the FT analysis, which employs complex exponential or global sine and cosine functions as the basis functions, the WT analysis uses single localised “small waves” or wavelets as the basis functions. Each wavelet function, commonly referred to as basis wavelet or mother wavelet, is defined by two parameters: its scale (relating to frequency) and its position (relating to time).

Although the FT has proven to be extremely valuable to analyse periodic, time-invariant, or stationary phenomena, the frequency spectrum of a signal as a result of the FT is not localised in time due to the infinite sinusoid basis functions. This characteristic implies that the Fourier coefficients of a signal are determined by the entire signal support. Consequently, any local behaviour of a signal cannot be easily traced from its FT. In contrast, in WT analysis long waves, corresponding to larger scale values, are used for more precise low-frequency information and shorter waves, corresponding to smaller scale values, are used for the time locality of high-

frequency information. Therefore, the WT is a more suitable and powerful tool for the analysis of transient, time-varying, or non-stationary phenomena, as both frequency (scales) and time information can be obtained simultaneously from the WT of a signal. Hence, the two theories, rather than competing, are complementary since there are applications where the WT analysis is better suited than the FT analysis and vice versa.

### 2.1.1 Continuous wavelet transform

The Continuous Wavelet Transform (CWT) of a function  $f(t) \in L^2(\mathbb{R})$  (the space of square integrable functions) is defined by

$$(W_\psi f)(u, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-u}{s}\right) dt \quad (6)$$

where  $u, s \in \mathbb{R}$  are real continuous variables,  $s \neq 0$ . The continuous wavelet transform comprises the continuous translation and dilation of a basis function defined by:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (7)$$

where the factor  $1/\sqrt{s}$  is used to normalise the energy so that it stays at the same level for different values of  $s$  and  $u$ . The wavelet function  $\psi_{u,s}(t)$  is expanded in time (or space) when  $s$  is increased, and is displaced in time (or space) when  $u$  is varied. For this reason,  $s$  is called the scaling (or dilation) parameter, which captures the local frequency content, and  $u$  is called the translation (or shifting) parameter, which localises the wavelet basis function at time  $t = u$  and its vicinity.

### 2.1.2 Discrete wavelet transform

In many practical applications the CWT is discretised in the scaling and dilation parameters for computational efficiency. Thus, instead of calculating the wavelet transform over the continuous range of  $s$  and  $u$ , the wavelet transform is calculated only at the discrete values defined by  $s = s_0^{-j}$  and  $u = nu_0 s_0^{-j}$ . Then, by substituting these discrete scaling and translation values in (7) it becomes:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{s_0^{-j}}} \psi\left(\frac{t - nu_0 s_0^{-j}}{s_0^{-j}}\right) = s_0^{j/2} \psi(s_0^j t - nu_0) \quad (8).$$

where  $j, n \in \mathbb{Z}$  (the set of all integers),  $s_0 > 1$  and  $u_0 > 0$  are fixed dilation and translation steps, respectively. Note in (8) that the translation factor  $u_0$  has been made dependent on the dilation step  $s_0$ . Most commonly,  $s_0$  and  $u_0$  are selected in order to have a dyadic grid along the frequency and time axes. This is  $s_0 = 2$ , and  $u_0 = 1$ . Therefore, by substituting these values in (8) it gives:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^{-j}}} \psi\left(\frac{t - n2^{-j}}{2^{-j}}\right) = 2^{j/2} \psi(2^j t - n) \quad j, n \in \mathbb{Z} \quad (9).$$

Thus, any signal in  $L^2(\mathbb{R})$  can be represented as a superposition of dyadic dilations and translations of a single wavelet function  $\psi(t)$  (also known as wavelet series representation), this is:

$$f(t) = \sum_{j,n} d_{j,n} \psi_{j,n}(t) = \sum_{j,n} d_{j,n} 2^{j/2} \psi(2^j t - n) \quad (10)$$

where the two-dimensional set of coefficients  $d_{j,n}$  is called the Discrete Wavelet Transform (DWT) of  $f(t)$  and (10) is the inverse DWT. If the functions  $\psi_{j,n}(t)$  form an orthogonal basis for the space of signals of interest, then a more specific form of the DWT indicating how the  $d_{j,n}$ 's are calculated can be written using the inner product operator as:

$$f(t) = \sum_{j,n} \langle \psi_{j,n}(t), f(t) \rangle \psi_{j,n}(t) \quad (11).$$

Recall that the inner product of two functions  $x(t)$  and  $y(t)$  is defined as:

$$\langle x(t), y(t) \rangle = \int x(t)y(t)dt \quad (12),$$

and two wavelets  $\psi_{j,n}(t)$  and  $\psi_{j,k}(t)$  are orthogonal if their inner product is equal to zero, this is:

$$\langle \psi_{j,n}(t), \psi_{j,k}(t) \rangle = \int \psi_{j,n}(t)\psi_{j,k}(t)d(t) = 0 \quad (13).$$

The construction of these orthogonal bases can be directly related to the theory of multiresolution signal approximations [Mallat, 1989]. Specifically, this leads to an equivalence between wavelet bases and quadrature mirror filters used in discrete multirate filter banks. These filter banks can then be used to implement a fast DWT algorithm that requires only  $O(N)$  operations for signals of size  $N$ . This implementation of the DWT is reviewed next.

Lets consider the wavelet series representation of a square integrable function  $f(t)$  introduced in (10):

$$f(t) = \sum_{j,n} d_{j,n} \psi_{j,n}(t) = \sum_{j,n} d_{j,n} 2^{j/2} \psi(2^j t - n) \quad (14)$$

for every  $j \in \mathbb{Z}$  the following functions can be defined [Coca and Billings, 2001]:

$$w_j(t) = \sum_n d_{j,n} \psi_{j,n}(t) = 2^{j/2} \sum_n d_{j,n} \psi(2^j t - n) \quad (15).$$

It follows that the function  $f(t)$  can then be expressed as:

$$f(t) = \sum_{j \in \mathbb{Z}} w_j(t) = \dots + w_{-1}(t) + w_0(t) + w_1(t) + \dots \quad (16).$$

But, since the scale is roughly speaking a substitute for frequency in the time domain, equation (16) can be interpreted as a decomposition of the function  $f(t)$  in frequency bands, where higher values of  $j$  correspond to higher frequency bands.

If all the functions  $w_j(t)$  are added together up to a scale  $j \subset \mathbb{Z}$ , then a new set of functions  $\{v_j(t)\}_{j \in \mathbb{Z}}$  can be defined as:

$$v_j(t) = \sum_{l=-\infty}^{j-1} w_l(t) \quad (17).$$

Then, for every  $j \in \mathbb{Z}$ , the functions  $v_j(t)$  can be substituted in (16) to obtain:

$$f(t) = v_j(t) + w_j(t) + w_{j+1}(t) + \dots \quad (18).$$

Continuing with the frequency band interpretation of equation (16), equation (18) can be interpreted in the frequency domain as a decomposition of the function  $\hat{f}(\omega)$ , the FT of  $f(t)$ , in a low-frequency band  $\hat{v}_j(\omega)$ , plus additional, high-frequency bands represented by  $\hat{w}_j(\omega) + \hat{w}_{j+1}(\omega) + \dots$ . From an approximation point of view  $v_j(t)$  is a coarse approximation of the function  $f(t)$  with the additional, finer detail provided by the functions  $w_j(t) + w_{j+1}(t) + \dots$ . The value of  $j$  which corresponds to the initial resolution level controls the amount of detail contained in the function  $v_j(t)$  relative to the original function  $f(t)$ .

The alternative representation of  $f(t)$  given in equation (18) can be realised through the utilisation of a scaling basis function  $\phi(t) \in L^2(\mathbb{R})$  which dyadic dilations and translations are used to expand the functions  $v_j(t)$ , for every integer  $j$ , in the same way  $\psi(t)$  is used to represent  $w_j(t)$  in (15). This yields the following representation of  $v_j(t)$  in terms of dilating and translating the scale basis function:

$$v_j(t) = \sum_{n \in \mathbb{Z}} c_{j,n} \phi_{j,n}(t) = 2^{j/2} \sum_{n \in \mathbb{Z}} c_{j,n} \phi(2^j t - n), \quad j \in \mathbb{Z} \quad (19)$$

where  $\{c_{j,n}\}$  are the coefficients of the expansion. Consequently, substituting (15) and (19) in (18) the function  $f(t)$  has a new series representation in terms of scaling and wavelet functions [Coca and Billings, 2001] [Burrus *et al.*, 1998]:

$$f(t) = \sum_{n \in \mathbb{Z}} c_{j,n} \phi_{j,n}(t) + \sum_j \sum_{n \in \mathbb{Z}} d_{j,n} \psi_{j,n}(t) \quad (20).$$

A scaling function which can be used to expand any square integrable function as in equation (20) for any integer  $j$  is said to generate a multi-resolution approximation over the space of square summable functions [Mallat, 1989] [Burrus *et al.*, 1998].

The coefficients in the wavelet expansion (20) are called the DWT of the signal  $f(t)$ . If the functions  $\phi_{j,n}(t)$  and  $\psi_{j,n}(t)$  are orthogonal, then the  $j$  level scaling coefficients are found by taking the inner products:

$$c_{j,n} = \langle f(t), \phi_{j,n}(t) \rangle = \int f(t) 2^{j/2} \phi(2^j t - n) dt \quad (21)$$

$$d_{j,n} = \langle f(t), \psi_{j,n}(t) \rangle = \int f(t) 2^{j/2} \psi(2^j t - n) dt \quad (22).$$

In practice, instead of dealing directly with the scaling and wavelet functions, a method using the theory of filter banks has been developed to calculate the coefficients  $c_{j,n}$  and  $d_{j,n}$  in a recursive way. First, a relationship between the expansion coefficients at a lower scale level in terms of those at a

higher scale is derived using the basic recursion equation, also known as the multiresolution analysis equation [Burrus *et al.*, 1998]:

$$\phi(t) = \sum_k h(k)\sqrt{2}\phi(2t-k) \quad (23)$$

this equation states that the scaling function  $\phi(t)$  can be expressed in terms of a weighted sum of shifted  $\phi(2t)$ . Thus, by scaling and translating the time variable it gives:

$$\phi(2^j t - n) = \sum_k h(k)\sqrt{2}\phi(2(2^j t - n) - k) = \sum_k h(k)\sqrt{2}\phi(2^{j+1} t - 2n - k) \quad (24)$$

which, after defining a change of variable  $m = 2n + k$ , and substituting it in (24), results in:

$$\phi(2^j t - n) = \sum_m h(m-2n)\sqrt{2}\phi(2^{j+1} t - m) \quad (25).$$

Then, using (25) and interchanging the sum and integral, (21) can be written as:

$$c_{j,n} = \sum_m h(m-2n) \int f(t) 2^{(j+1)/2} \phi(2^{j+1} t - m) dt \quad (26)$$

but the integral in (26) is the inner product with the scaling function at a scale of  $j+1$  giving:

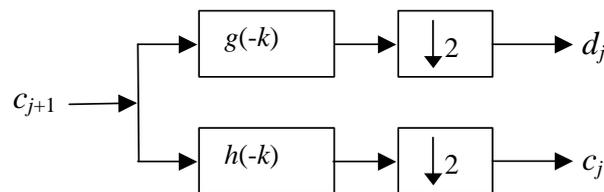
$$c_{j,n} = \sum_m h(m-2n) c_{j+1,m} \quad (27).$$

Following the same procedure the corresponding relationship for the wavelet coefficients is obtained:

$$d_{j,n} = \sum_m g(m-2n) c_{j+1,m} \quad (28).$$

Equations (27) and (28) show that the scaling and wavelet coefficients at different levels of scale can be obtained by convolving the expansion coefficients at scale  $j$  by the time-reversed recursion coefficients  $h(-k)$  and  $g(-k)$  then down-sampling or decimating (taking every other term, the even terms) to give the expansion coefficients at the next level of  $j-1$ . In other words, the scale- $j$  coefficients are “filtered” by two FIR digital filters with coefficients  $h(-k)$  and  $g(-k)$  after which down-sampling gives the next coarser scaling and wavelet coefficients.

The block diagram of figure 3 illustrates the implementation of equations (27) and (28). In this figure the down-pointing arrows denote a down-sampling by two and the other blocks denote FIR filters or equivalently a convolution by  $h(-k)$  or  $g(-k)$ .



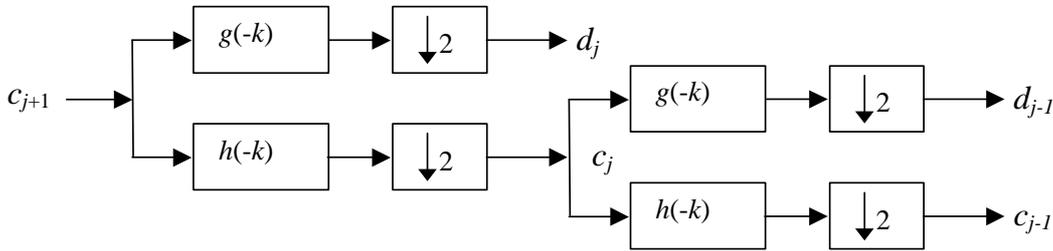
**Figure 3:** Representation of the implementation of a two-band analysis filter bank.

The filter implemented by  $h(-k)$  is a lowpass filter, and the one implemented by  $g(-k)$  is a highpass filter. The filters are designed so that the impulse response of the filter  $g$  is related to the impulse response of the filter  $h$  by:

$$g(k) = (-1)^{1-k} h(1-k) \quad (29).$$

Equation (29) specifies that the filter  $g$  is the mirror filter of  $h$ . In signal processing,  $g$  and  $h$  are called *quadrature mirror filters*.

Note that the total number of data points at the output of the filter bank is the same as the number of data points coming in. The number of data points is doubled by having two filters, but because the down-sampling, this number is halved back to the original number. If more decomposition levels are required, then the splitting, filtering, and decimation structure can be repeated on the scaling coefficients as is illustrated in figure 4.



**Figure 4:** Implementation of a two-stage two-band analysis filter bank tree.

For a more detailed description of filter banks and its utilisation to implement the DWT the reader is referred to [Burrus *et al.*, 1998], [Mallat, 1989], and [Strang and Nguyen, 1997].

## 2.2 Fuzzy logic theory

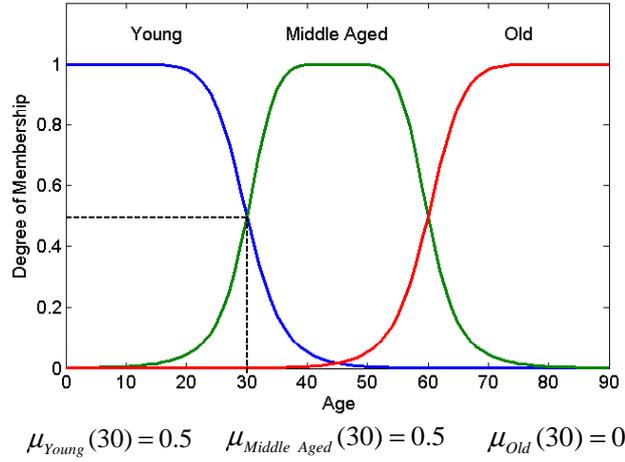
Imprecision and uncertainty are inherent concepts to the inexact nature of human reasoning. As a result, our way of interpreting the world is generally seen as a function of vague propositions, uncertain data and appreciative judgements. However, this way of thinking is not taken into account in traditional logic, where only two fundamental premises exist: true and false, 0 and 1. Lotfi A. Zadeh noticed this and created a new logic, called fuzzy logic [Zadeh, 1973], in order to attempt to capture the uncertainty present in our reasoning when interpreting the world. This logic is based on the theory of *fuzzy sets* proposed by Zadeh [1965] in his seminal paper of 1965. The main concept in fuzzy sets theory is that an object is no longer restricted to be completely a member or not a member of a set. Instead, any element is allowed to have a grade of membership intermediate between full membership and non-membership, this is a membership value in the whole range [0,1]. In other words, whereas in traditional sets theory a set has sharp borders, in fuzzy sets theory a fuzzy set has soft borders allowing an object a smooth transition between being a member or not a member of a particular set. This smooth transition is characterised by a membership function, which gives fuzzy sets flexibility in modelling commonly used linguistic expressions, such as “the temperature is high”, or “the speed is fast”.

Formally, a fuzzy set is defined as follows. Let  $X$  be a universe of discourse (e.g. a space of points or objects) and  $x$  denote a generic element of  $X$ . Then, a fuzzy set  $A$  in  $X$  is characterised by a membership function  $\mu_A : X \rightarrow [0,1]$  which associates with each element  $x$  of  $X$  a real number

$\mu_A(x)$  in the interval  $[0,1]$ , with the value  $\mu_A(x)$  representing the degree (or grade) of membership of  $x$  in  $A$  [Zadeh, 1977b]. Thus, a fuzzy set  $A$  in  $X$  can be represented as the set of ordered pairs:

$$A = \{(x, \mu_A(x)) \mid x \in X\} \quad (30)$$

where  $0 \leq \mu_A(x) \leq 1$ , with  $\mu_A(x) = 0$  represents no membership and  $\mu_A(x) = 1$  represents full membership. As an example, let  $X$  consist of the ages of all people. Three fuzzy subsets of  $X$ , Young, Middle aged, and Old, are shown in figure 5 representing those people that are young middle aged and old, respectively. Hence, these membership functions determine that a 30-years-old person belongs equally to the fuzzy sets Young and Middle Aged, but she or he does not belong to the fuzzy set Old.



**Figure 5:** Fuzzy sets defined to represent the age of people.

Similarly to traditional set theory, the corresponding basic operations of intersection, union, and complement, are defined in fuzzy logic. The operations of intersection and union are specified in general by a T-norm and a T-conorm operator, respectively [Jang *et al*, 1997]. Below, the two more frequently used T-norm and T-conorm operators are presented.

1. *Intersection:* The intersection of two fuzzy sets  $A$  and  $B$  is a fuzzy set  $C$ , denoted as  $C = A \cap B$ , whose membership function is related to those of  $A$  and  $B$  by:

$$\text{Minimum:} \quad \mu_C(x) = \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \quad (31)$$

$$\text{Algebraic product:} \quad \mu_C(x) = \mu_{A \cap B}(x) = \mu_A(x) \cdot \mu_B(x) \quad (32)$$

2. *Union:* The union of two fuzzy sets  $A$  and  $B$  is a fuzzy set  $C$ , denoted as  $C = A \cup B$ , whose membership function is related to those of  $A$  and  $B$  by:

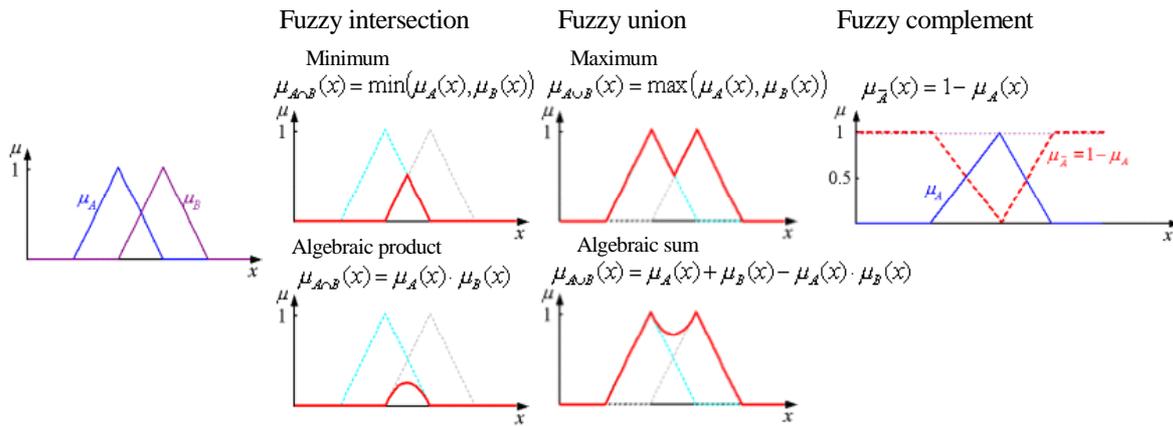
$$\text{Maximum:} \quad \mu_C(x) = \mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) \quad (33)$$

$$\text{Algebraic sum:} \quad \mu_C(x) = \mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x) \quad (34)$$

3. *Complement (negation):* The complement (negation) of fuzzy set  $A$ , denoted by  $\bar{A}$ , is defined as:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (35)$$

Examples of these basic operations are presented in figure 6.



**Figure 6:** Basic operations on fuzzy sets.

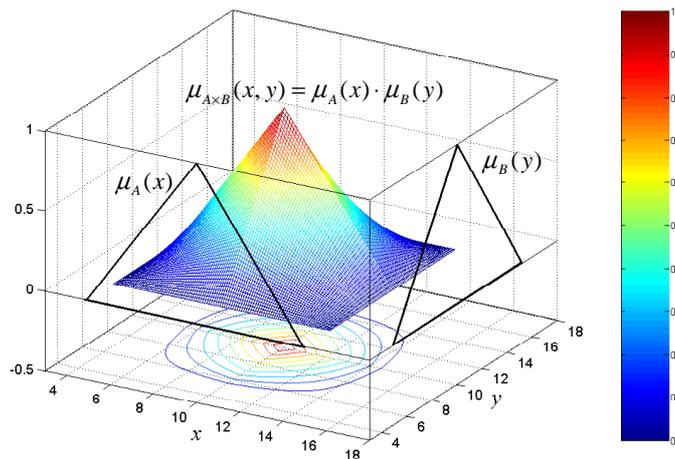
Other important operations in fuzzy sets, which are also direct generalisations of operations on ordinary sets are:

4. *Cartesian product:* Let  $A$  and  $B$  be fuzzy sets in  $X$  and  $Y$ , respectively. The Cartesian product of  $A$  and  $B$ , denoted by  $A \times B$ , is a fuzzy set in the product space  $X \times Y$  with the membership function:

$$\text{Minimum:} \quad \mu_{A \times B}(x, y) = \min(\mu_A(x), \mu_B(y)) \quad (36)$$

$$\text{Algebraic product:} \quad \mu_{A \times B}(x, y) = \mu_A(x) \cdot \mu_B(y) \quad (37)$$

The Cartesian product of two fuzzy sets is characterised by a two-dimensional membership function. Figure 7 shows an example of a two-dimensional membership function generated by performing the Cartesian product of two fuzzy sets using the algebraic product operator. The same figure also shows the contour plot generated by the two-dimensional membership function.



**Figure 7:** Cartesian product using the algebraic product operator.

5. *Binary fuzzy relation:* Let  $A$  and  $B$  be two fuzzy sets in the universes of discourse  $X$  and  $Y$ , respectively. Then, a fuzzy relation of the form  $A \Rightarrow B$ , denoted by  $R$ , from the fuzzy set  $A \subset X$  to the fuzzy set  $B \subset Y$ , is a fuzzy subset of the Cartesian product  $X \times Y$ . This is:

$$R = \{((x, y), \mu_R(x, y)) \mid (x, y) \in X \times Y\} \quad (38)$$

where  $\mu_R(x, y)$  is a two-dimensional membership function, characterised by:

$$\text{Minimum:} \quad \mu_R(x, y) = \mu_{A \times B}(x, y) = \min(\mu_A(x), \mu_B(y)) \quad (39)$$

$$\text{Algebraic product:} \quad \mu_R(x, y) = \mu_{A \times B}(x, y) = \mu_A(x) \cdot \mu_B(y) \quad (40)$$

Common examples of fuzzy relations are expressions of the form “If  $x$  is  $A$ , then  $y$  is  $B$ ”, e.g. “If  $x$  is *large* then  $y$  is *small*”. Fuzzy relations of this kind are repeatedly used in fuzzy inference systems.

Fuzzy relations in different product spaces can be combined through a *composition operation*. Different composition operations have been suggested for fuzzy relations, depending on the operator selected to perform the fuzzy intersection and union operations. The two most popular composition operations are the *max-min* and *max-product*, which are presented below.

6. *Max-min composition:* Let  $R_1$  and  $R_2$  be two fuzzy relations defined on  $X \times Y$  and  $Y \times Z$ , respectively, Then, the max-min composition of  $R_1$  and  $R_2$  is a fuzzy set defined by:

$$R_1 \circ R_2 = \{[(x, z), \max_y \min(\mu_{R_1}(x, y), \mu_{R_2}(y, z))] \mid x \in X, y \in Y, z \in Z\}, \quad (41)$$

which is characterised by the membership function defined by:

$$\mu_{R_1 \circ R_2}(x, z) = \max_y \min(\mu_{R_1}(x, y), \mu_{R_2}(y, z)) \quad (42)$$

7. *Max-product composition:* Assuming the same notation as used in the definition of the max-min composition, the max-product composition is defined as follows:

$$R_1 \circ R_2 = \{[(x, z), \max_y (\mu_{R_1}(x, y) \cdot \mu_{R_2}(y, z))] \mid x \in X, y \in Y, z \in Z\} \quad (43)$$

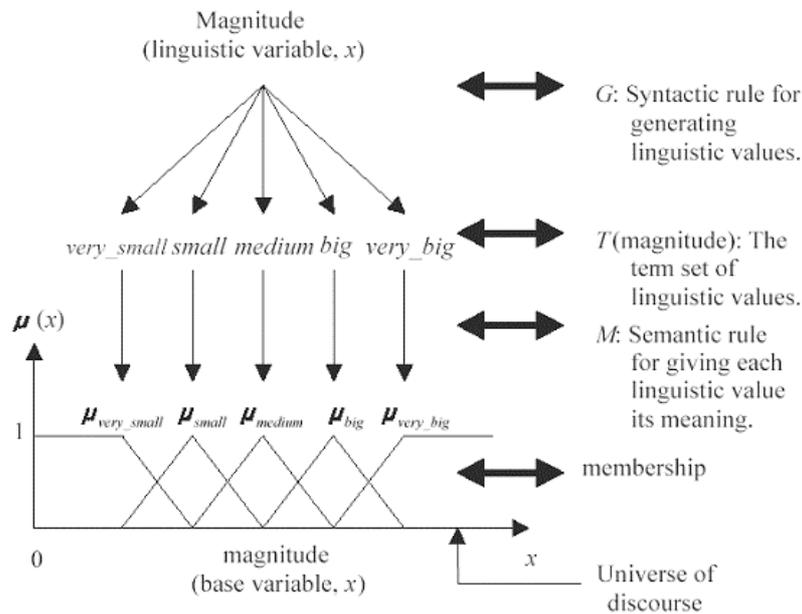
characterised by the membership function:

$$\mu_{R_1 \circ R_2}(x, z) = \max_y (\mu_{R_1}(x, y) \cdot \mu_{R_2}(y, z)) \quad (44).$$

Another important concept in fuzzy logic theory is the concept of *linguistic variable*. A linguistic variable is a device for systematising the use of words or sentences, expressed in a natural or artificial language, for the purpose of characterising the values of variables and describing their interrelations [Zadeh, 1977]. As its name suggest, the values of a linguistic variable are words or sentences, rather than numbers. Each linguistic value is a label of a fuzzy set, defined by a membership function with gradual transition between full membership and non-membership [Zadeh, 1965]. In formal terms:

8. *Linguistic variable:* A linguistic variable is characterised by a quintuple  $(x, T(x), X, G, M)$  in which  $x$  is the name of the variable;  $T(x)$  is the term-set of  $x$ , that is, the collection of its linguistic values (or linguistic terms);  $X$  is a universe of discourse;  $G$  is a syntactic rule which generates the terms in  $T(x)$ ; and  $M$  is a semantic rule which associates with each linguistic value  $A$  its meaning,  $M(A)$ , where  $M(A)$  denotes a fuzzy subset of  $X$ . The meaning of a linguistic value  $A$  is characterised by a membership function,  $\mu_A : X \rightarrow [0, 1]$ , which associates with each  $x$  in  $X$  a degree of membership in  $A$ .

The concept of linguistic variable can be applied to signal analysis. For example, the magnitude of a signal can be viewed as a linguistic variable, whose linguistic values are given in the term set  $T(\text{Magnitude}) = (\text{very small}; \text{small}; \text{medium}; \text{big}; \text{very big})$ . A graphical representation of the linguistic variable “Magnitude” is shown in figure 8.



**Figure 8:** Linguistic variable Magnitude

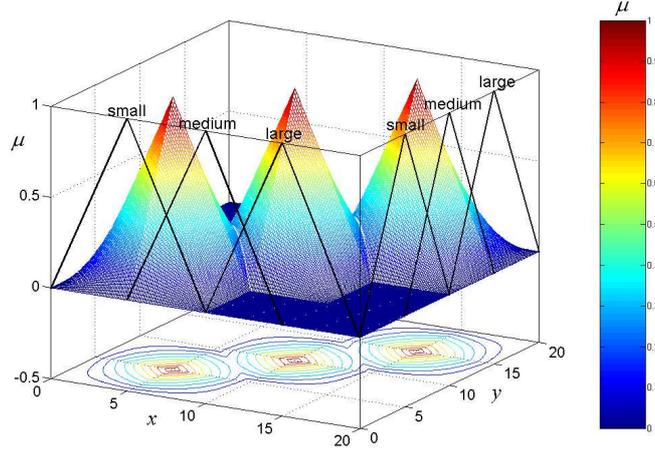
The concepts of fuzzy logic and linguistic variable can be used as a framework to encode structured knowledge in what is known as fuzzy associative memory [Kosko, 1992] [Brown and Harris, 1994] (also referred to as fuzzy inference system [Jang *et al*, 1997]), which can be defined as follows:

9. *Fuzzy associative memory*: A Fuzzy associative memory (FAM) is a rule-based system based on fuzzy sets and fuzzy logic. An FAM system encodes a bank of compound FAM rules that associate multiple inputs or antecedent fuzzy sets with multiple outputs or consequent fuzzy sets. An FAM rule defines an input-output transformation expressed as a logical if-then statement such as, “if this antecedent (group of fuzzy input sets) occurs, then this consequent (fuzzy output set) should be used”.

The knowledge in an FAM is encoded as a set of fuzzy rules and a fuzzy inference algorithm applied to these rules. For example, consider the following single input, single output FAM system, which has three triangular membership functions defined on each variable that represents the linguistic terms *small*, *medium*, and *large*:

FAM fuzzy rules  
 If  $x$  is *small*, then  $y$  is *small*  
 OR If  $x$  is *medium*, then  $y$  is *medium*  
 OR If  $x$  is *large*, then  $y$  is *large*

Each fuzzy rule can be coded as a fuzzy relation, which is characterised by a two-dimensional fuzzy set. The combination of all three fuzzy rules through a composition operator generates a fuzzy relational surface. Figure 9 shows the fuzzy relational surface generated for the FAM system described above using the max-product composition. Note that the each peak in the relational surface corresponds to each one of the three fuzzy rules, and that a contour or fuzzy region in the  $x$ - $y$  plane can be associated to each fuzzy rule.

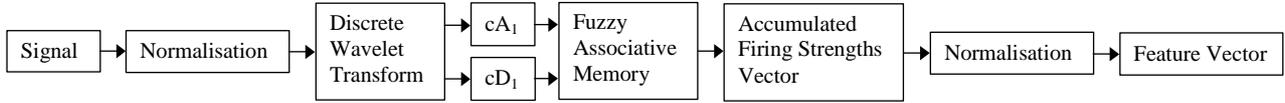


**Figure 9:** Example of a fuzzy relational surface for a single input, single output FAM system.

Fuzzy inference systems (FAMs), and in general fuzzy logic, has found successful applications in a wide variety of fields, such as automatic control, data classification, decision analysis, expert systems, time series prediction, robotics, and pattern recognition.

### 2.3 Feature extraction algorithm

The stages involved in the SCFEA [Li *et al.*, 2004] are summarised in the flow chart shown in Figure 10. A description of each one of these stages is given below.



**Figure 10:** Flow chart of the SCFEA.

*Normalisation input.* The incoming signal  $f(t)$  is normalised before performing discrete wavelet transform as follows:

$$\tilde{f}(t) = \frac{f(t)}{\max(\text{abs}(f(t)))} \quad (45).$$

After normalisation the incoming sampled values of the given signal are between the range  $[-1,1]$ . Note, that the normalisation is performed over a cycle or a given number of cycles of the incoming signal.

*Discrete wavelet transform.* Using the multi-resolution method described in Section 2.1.2 to obtain the DWT [Burrus *et al.*, 1998], the normalised signal is decomposed to its approximation and detail signals, denoted here as  $cA_1(t)$ , and  $cD_1(t)$ , respectively. The approximation signal contains the high-scale, low-frequency components of the original signal, while the detail signal contains the low-scale, high-frequency components of the original signal. Note that only one level DWT decomposition is performed, indicated by the subscript 1.

*Fuzzy associative memory.* The obtained approximation and detail signals can be interpreted as two linguistic variables. The values of these linguistic variables can then be described using fuzzy sets and identified using linguistic labels e.g. *very small*, *small*, *medium*, *big*, etc. Each fuzzy set

(linguistic value) is defined by a membership function with gradual transition between full membership and non-membership. By combining the fuzzy sets among the universes of discourse for the approximation and detail signals, a two dimensions linguistic approximation-detail hyperspace is created. The hyperspace can be interpreted as the fuzzy-wavelet feature plane. Several fuzzy regions cover this feature plane, each of which is associated with a two-dimensional fuzzy set. These fuzzy sets combinations form a set of fuzzy rules or fuzzy associations and all together form a fuzzy associative memory (FAM). Therefore, the data of discrete approximation and detail signals can be scattered on the feature plane. Each data of approximation-detail pair has its membership degrees to the fuzzy sets on the plane. Therefore, by evaluating each FAM rule, each approximation-detail pair activates or “fires” a fuzzy region to a different degree.

In formal terms, lets define the linguistic variables  $x_A$  and  $x_D$  to represent the approximation and detail signals  $cA_1(t)$  and  $cD_1(t)$ , respectively, obtained from a single level DWT decomposition. These two variables are used as linguistic input variables to the SCFEA. The corresponding linguistic value sets for  $x_A$  and  $x_D$  are given by,

$$T_A = \begin{pmatrix} v_{A,1} \\ v_{A,2} \\ \vdots \\ v_{A,nA} \end{pmatrix}, T_D = \begin{pmatrix} v_{D,1} \\ v_{D,2} \\ \vdots \\ v_{D,nD} \end{pmatrix} \quad (46)$$

where  $T_A$  and  $T_D$  are the term sets for the approximation and detail signals, respectively;  $v_{A,i}$  and  $v_{D,j}$ ,  $i = 1,2,\dots,nA$  and  $j = 1,2,\dots,nD$  are linguistic values for  $x_A$  and  $x_D$ , respectively;  $nA$  and  $nD$  are the number of linguistic values for  $TA$  and  $TD$ , respectively. The corresponding membership function sets can then be denoted as follows:

$$\mu_{T_A}(cA(t)) = \begin{pmatrix} \mu_{v_{A,1}}(cA(t)) \\ \mu_{v_{A,2}}(cA(t)) \\ \vdots \\ \mu_{v_{A,nA}}(cA(t)) \end{pmatrix}, \mu_{T_D}(cD(t)) = \begin{pmatrix} \mu_{v_{D,1}}(cD(t)) \\ \mu_{v_{D,2}}(cD(t)) \\ \vdots \\ \mu_{v_{D,nD}}(cD(t)) \end{pmatrix} \quad (47)$$

where  $\mu_{v_{A,i}}$  and  $\mu_{v_{D,j}}$  are the  $i$ th membership function and the  $j$ th membership function for  $x_A$  and  $x_D$ , respectively,  $cA$  and  $cD$  are the base variables for  $x_A$  and  $x_D$ , respectively. The relationship between the linguistic value set and the membership function set is expressed as:

$$S_A(cA(t)) = T_A \otimes \mu_{T_A}(cA(t)) = \begin{pmatrix} v_{A,1}, \mu_{v_{A,1}}(cA(t)) \\ v_{A,2}, \mu_{v_{A,2}}(cA(t)) \\ \vdots \\ v_{A,nA}, \mu_{v_{A,nA}}(cA(t)) \end{pmatrix} = \begin{pmatrix} S_{A,1}(cA(t)) \\ S_{A,2}(cA(t)) \\ \vdots \\ S_{A,nA}(cA(t)) \end{pmatrix}, \quad (48a)$$

$$S_D(cD(t)) = T_D \otimes \mu_{T_D}(cD(t)) = \begin{pmatrix} v_{D,1}, \mu_{v_{D,1}}(cD(t)) \\ v_{D,2}, \mu_{v_{D,2}}(cD(t)) \\ \vdots \\ v_{D,nD}, \mu_{v_{D,nD}}(cD(t)) \end{pmatrix} = \begin{pmatrix} S_{D,1}(cD(t)) \\ S_{D,2}(cD(t)) \\ \vdots \\ S_{D,nD}(cD(t)) \end{pmatrix} \quad (48b)$$

where  $S_A(cA(t))$  and  $S_D(cD(t))$  are the fuzzy set structures for the linguistic variables  $x_A$  and  $x_D$ , respectively. The symbol  $\otimes$  represents the major Cartesian product operator, and it associates in a one-to-one way the elements of the two sets. The fuzzy set structures  $S_A(cA(t))$  and  $S_D(cD(t))$  together form the fuzzy basis set denoted by:

$$S(C(t)) = \{S_A(cA(t)), S_D(cD(t))\} \quad (49)$$

where  $C(t) = [cA(t) \quad cD(t)]^T$  is the crisp (not fuzzy) input vector to the fuzzy feature plane at time  $t$ .

Fuzzy regions on the plane  $x_A - x_D$  are formed by performing fuzzy associations (or fuzzy relations) between the fuzzy sets in  $S_A(cA(t))$  and  $S_D(cD(t))$ . These associations are defined by the Cartesian product  $S_A(cA(t)) \times S_D(cD(t))$ . The number of fuzzy regions  $nR$  on the feature plane is determined by the cardinalities of  $S_A(cA(t))$  and  $S_D(cD(t))$ , this is  $nR = nA \times nD$ . Each fuzzy region is thus defined as a fuzzy relation of two fuzzy sets selected from the basis set  $S(C(t))$ . In general, each fuzzy set structure provides one fuzzy set for each fuzzy region. Thus, each region is generated by associating each fuzzy set in  $S_A(cA(t))$  with each fuzzy set in  $S_D(cD(t))$ , this is:

$$\begin{aligned} R_{1,1}(S(C(t))) &= \{S_{A,1}(cA(t)), S_{D,1}(cD(t))\} = \{(v_{A,1}, \mu_{v_{A,1}}(cA(t))), (v_{D,1}, \mu_{v_{D,1}}(cD(t)))\} = V_{1,1} \otimes \mu_{1,1}(C(t)) \\ &\vdots \\ R_{1,nD}(S(C(t))) &= \{S_{A,1}(cA(t)), S_{D,nD}(cD(t))\} = \{(v_{A,1}, \mu_{v_{A,1}}(cA(t))), (v_{D,nD}, \mu_{v_{D,nD}}(cD(t)))\} = V_{1,nD} \otimes \mu_{1,nD}(C(t)) \\ R_{2,1}(S(C(t))) &= \{S_{A,2}(cA(t)), S_{D,1}(cD(t))\} = \{(v_{A,2}, \mu_{v_{A,2}}(cA(t))), (v_{D,1}, \mu_{v_{D,1}}(cD(t)))\} = V_{2,1} \otimes \mu_{2,1}(C(t)) \\ &\vdots \\ R_{2,nD}(S(C(t))) &= \{S_{A,2}(cA(t)), S_{D,nD}(cD(t))\} = \{(v_{A,2}, \mu_{v_{A,2}}(cA(t))), (v_{D,nD}, \mu_{v_{D,nD}}(cD(t)))\} = V_{2,nD} \otimes \mu_{2,nD}(C(t)) \\ &\vdots \\ R_{nA,nD}(S(C(t))) &= \{S_{A,nA}(cA(t)), S_{D,nD}(cD(t))\} = \{(v_{A,nA}, \mu_{v_{A,nA}}(cA(t))), (v_{D,nD}, \mu_{v_{D,nD}}(cD(t)))\} = V_{nA,nD} \otimes \mu_{nA,nD}(C(t)) \end{aligned} \quad (50)$$

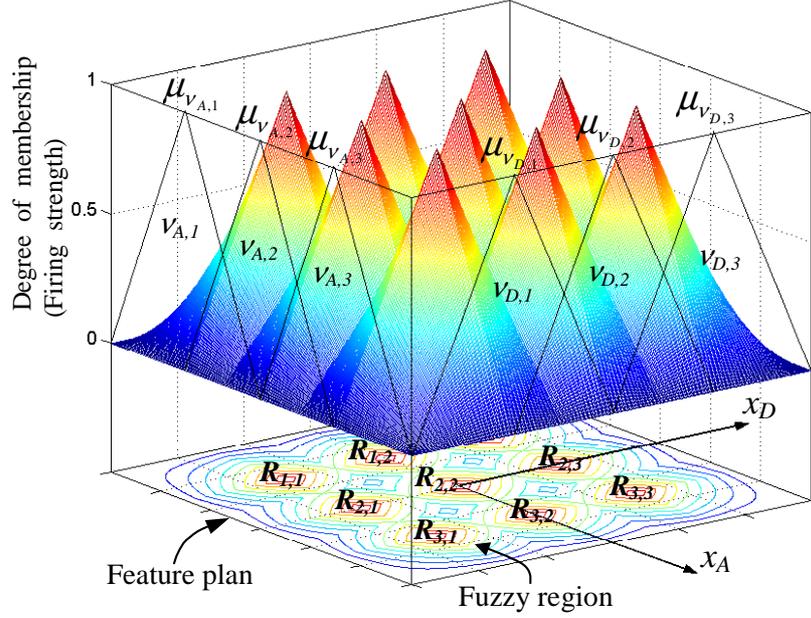
Therefore, the  $i$ th,  $j$ th fuzzy region can be defined as:

$$R_{i,j}(S(C(t))) = \{S_{A,i}(cA(t)), S_{D,j}(cD(t))\} = \{(v_{A,i}, \mu_{v_{A,i}}(cA(t))), (v_{D,j}, \mu_{v_{D,j}}(cD(t)))\} = V_{i,j} \otimes \mu_{i,j}(C(t)) \quad \dots(51)$$

where  $V_{i,j} = (v_{A,i}, v_{D,j})$  is the fuzzy region linguistic value set for the  $i$ th,  $j$ th fuzzy region  $R_{i,j}$  on the feature plane, and  $\mu_{i,j}(C(t)) = (\mu_{v_{A,i}}(cA(t)), \mu_{v_{D,j}}(cD(t)))$  is the corresponding fuzzy region membership function set. Hence, a firing condition for the  $i$ th,  $j$ th region on the feature plane can be expressed in linguistic terms as a fuzzy rule:

$$\text{“Region } R_{i,j} \text{ is activated if } (x_A \text{ is } v_{A,i} \text{ and } x_D \text{ is } v_{D,j})\text{”} \quad (52)$$

Note that the condition  $(x_A \text{ is } v_{A,i} \text{ and } x_D \text{ is } v_{D,j})$  can be represented as a two-dimensional fuzzy set obtained by performing the Cartesian product between the corresponding membership functions. The combination of all the two-dimensional fuzzy sets (fuzzy rules) through a composition operator generates a fuzzy relational surface or FAM. Figure 11 shows an example of a feature plane and relational surface generated for a system with three fuzzy sets defined for the linguistic variables  $x_A$  and three fuzzy sets defined for the linguistic variables  $x_D$ . The surface was generated using the max-product composition operator.



**Figure 11:** Feature plane generated using the max-product composition operator.

The firing strength or degree of activation of the  $i$ th,  $j$ th fuzzy region on the feature plane is given by,

$$(\chi_{i,j}, \beta_{i,j}(t)) = \wedge R_{i,j}(S(C(t))) \quad (53),$$

where  $\wedge$  is the fuzzy intersection operator, which can be any t-norm operator;  $\chi_{i,j} \wedge (V_{i,j})$  is the firing condition stated as “ $x_A = V_{A,i} \wedge x_D = V_{D,j}$ ”; and  $\beta_{i,j}$  is the firing strength calculated as:

$$\beta_{i,j} = \wedge(\mu_{i,j}(C(t))) = \mu_{V_{A,i}}(cA(t)) \wedge \mu_{V_{D,j}}(cD(t)) \quad (54).$$

Adopting the algebraic product as the t-norm operator, then the firing strength will be given by:

$$\beta_{i,j} = \mu_{V_{A,i}}(cA(t)) \cdot \mu_{V_{D,j}}(cD(t)) \quad (55).$$

Equation (55) states that the degree in which the fuzzy region  $R_{i,j}$  is activated (fired) is given by multiplying the degree in which the data point  $cA(t)$  is member of the fuzzy set  $v_{A,i}$  by the degree in which the data point  $cD(t)$  is member of the fuzzy set  $v_{D,j}$ . Therefore, if a data pair  $(cA_1(t), cD_1(t))$  is corresponding to  $(cA(t), cD(t))$ , then this data pair will fire or activate in some degree each one of the fuzzy regions in the feature plane, value obtained applying (55).

*Accumulated firing strength vector.* Assuming that the original signal  $f(t)$  is sampled and there are  $N_0$  data points among a cycle, remember that  $f(t)$  has been assumed to be a periodic (cyclic) signal and it is processed cycle by cycle, then after performing one level DWT decomposition, as defined in section 2.1.2 both the approximation and detail signals will have  $N_0/2$  data points. Each approximation-detail pair  $(cA_1(t), cD_1(t))$ ,  $t = 1, 2, \dots, N_0/2$  will have its firing strength or degree of membership to each one of the fuzzy regions on the feature plane. Thus, the accumulated firing strength for the  $i$ th,  $j$ th fuzzy region can be written as,

$$\beta_{i,j}^{acc} = \sum_{t=1}^{N_0/2} \beta_{i,j}(t) = \sum_{t=1}^{N_0/2} (\mu_{v_{A,i}}(cA(t)) \cdot \mu_{v_{D,j}}(cD(t))) \quad (56)$$

After processing all data pairs  $\{(cA_1(t), cD_1(t)), t= 1,2,\dots, N_0 / 2 \}$ , the accumulated firing strengths for the fuzzy regions on the feature plane are collected together to form the feature vector for a cycle of the signal  $f(t)$ . The feature vector can then be expressed as:

$$\beta = \begin{bmatrix} \beta_{1,1}^{acc} \\ \beta_{1,2}^{acc} \\ \vdots \\ \beta_{nA,nD}^{acc} \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{N_0/2} (\mu_{v_{A,1}}(cA(t)) \cdot \mu_{v_{D,1}}(cD(t))) \\ \sum_{t=1}^{N_0/2} (\mu_{v_{A,1}}(cA(t)) \cdot \mu_{v_{D,2}}(cD(t))) \\ \vdots \\ \sum_{t=1}^{N_0/2} (\mu_{v_{A,nA}}(cA(t)) \cdot \mu_{v_{D,nD}}(cD(t))) \end{bmatrix} \quad (57).$$

*Normalisation output.* Finally, the feature vector (57) is normalised to a unit vector in order to minimise the sampling effect [Li *et al.*, 2004]. In mathematical terms the normalisation to a unit vector is given by,

$$\beta_N = \frac{1}{\sqrt{\sum_{i=1}^{nA} \sum_{j=1}^{nD} [\beta_{i,j}^{acc}]^2}} \begin{bmatrix} \beta_{1,1}^{acc} \\ \beta_{1,2}^{acc} \\ \vdots \\ \beta_{nA,nD}^{acc} \end{bmatrix} \quad (58)$$

The normalised feature vector (58) is the output of the SCFEA.

### 3 Soft computing feature extraction algorithm applied to tie bar data

In this section feature extraction results are presented corresponding to the SCFEA applied to extract damage-sensitive information from measured response data of the tie bar system component of the main rotor hub of a Lynx Helicopter. These results correspond to data gathered in six tests where several tie bars have been subjected to high level ground-air-ground (G-A-G) cyclic load testing until failure. As the signals obtained from the tests have a cyclic behaviour, feature vectors are extracted for every cycle on these signals, among each test. Having available the feature vectors, a comparison analysis is performed by calculating the angle and Model Assurance Criteria (MAC) between a selected reference feature vector and the remaining extracted feature vectors. In this way, the dynamic behaviour of the tie bars before the point of failure is expressed in terms of the variation of the extracted feature vectors over time and when compared to a reference feature vector. Results of feature extraction and comparison analysis are presented by test and by tie bar. From the comparison analysis it is clear that a pattern emerges in the data corresponding to the tie bar that has failed.

Tests of several tie bars were carried out by Westland in a purpose built test rig. Two tie bars were installed back to back in the test rig. Then cyclic twist and axial loads were applied to the tie bars simulating a high level G-A-G cycle load, and tested until failure, with varying results [Gorton, 2006]. Table I summarises the tests carried out and corresponding results. Note that a test is carried out until one of the tie bars fails (with an exception in test 5, where no failure was reached). The axial load (kN), the angle of twist (degrees), and the two tie bar extension displacements (mm)

were the parameters measured, constantly monitored and recorded on a digital chart recorder and stored in computer files. Therefore, soft computing feature extraction and comparison analysis has been carried out for each one of the measured parameters. This analysis has been carried out both by test and by tie bar and results are presented in the next sections.

**Table I:** Summary of the test completed

Test No.	Tie Bar Serial No.	Previous Usage (Flight Hours+cycles)	Tensile Load (KN)	Cycles Completed	Remarks
1	AET7119	671.3 FH	-1.0 to 330.6	216	Failed
	LJA0404	519.0 FH			
2	AET7119	671.3 FH+216 cyc.	-1.0 to 330.6	778	Failed
	BAH4263				
3	AET7119	671.3 FH+994 cyc.	+0.7 to 330.6	7026	Failed
	AEX5714	884.7 FH			
4	LJA0399		+0.7 to 330.6	1296	Failed
	AEX5714	884.7 FH+7026 cyc.			
5	LK0034	0 FH	+0.4 to 285/290	24485	Lord Corp. Man.
	LK0046	0 FH			
6	LJA0401	69.4 FH	+0.7 to 285/290	18553	Failed
	LJA1440	0 FH (New)	+0.7 to 330.6	+3120	
7	LJA0401	69.4 FH+21673 cyc.	+0.7 to 330.6	3845	Failed
	LJA2061	0 FH (New)			

### 3.1 Algorithm implementation

The SCFEA presented in section 2.3 was implemented in the MATLAB/Simulink simulation environment. First the process of normalisation is carried out by implemented equation (45). After normalisation, the DWT given in Section 2.1.2 was implemented using the Haar wavelet (the simplest wavelet) decomposition filter [Daubechies, 1992]. The approximation-detail data pairs ( $cA_1(t), cD_1(t)$ ) obtained from a single level DWT are used then for feature extraction with the fuzzy logic based approach as given in Section 2.3.

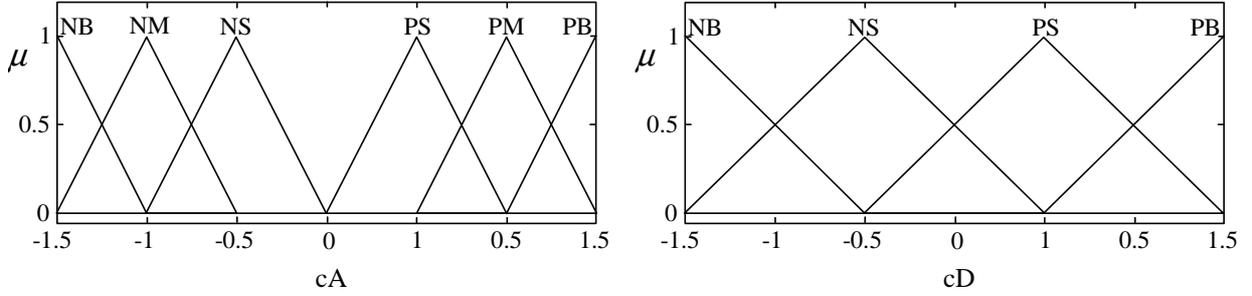
Two linguistic variables are defined  $x_A$  and  $x_D$  to represent the approximation and detail signals ( $cA_1(t), cD_1(t)$ ), respectively. The corresponding base variables are denoted as  $cA$  and  $cD$  for  $x_A$  and  $x_D$ , respectively. The linguistic value sets  $T_A$  and  $T_D$  and the corresponding membership functions for  $x_A$  and  $x_D$  are given as follows:

$$T_A = \begin{bmatrix} NB \\ NM \\ NS \\ PS \\ PM \\ PB \end{bmatrix} \text{ and } T_D = \begin{bmatrix} NB \\ NS \\ PS \\ PB \end{bmatrix}; \mu_A = \begin{bmatrix} \mu_{NB}(cA(t)) \\ \mu_{NM}(cA(t)) \\ \mu_{NS}(cA(t)) \\ \mu_{PS}(cA(t)) \\ \mu_{PM}(cA(t)) \\ \mu_{PB}(cA(t)) \end{bmatrix} \text{ and } \mu_D = \begin{bmatrix} \mu_{NB}(cD(t)) \\ \mu_{NS}(cD(t)) \\ \mu_{PS}(cD(t)) \\ \mu_{PB}(cD(t)) \end{bmatrix}$$

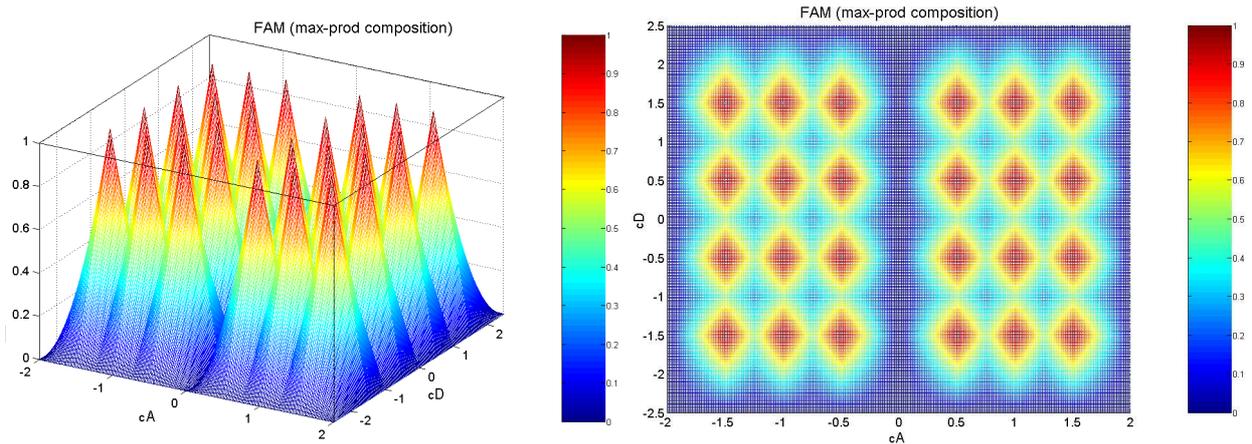
Thus, six fuzzy sets were defined for  $x_A$  (associated with the approximation signal  $cA_1$ ) and four fuzzy sets were defined for  $x_D$  (associated with the detail signal  $cD_1$ ). The fuzzy sets terms mean: NB = Negative Big, NM = Negative Medium, NS = Negative Small, PS = Positive Small, PM = Positive Medium, and PB = Positive Big. The associated membership functions were defined using

triangular functions as is shown in figure 12. The two-dimensional hyperspace or FAM generated by combining the fuzzy sets for the approximation and detail signals is shown in figure 13 together with the generated feature plane.

Each one of the rules generated by combining the fuzzy sets for the detail and approximation signals is listed in figure 14(a). As shown in figure 13, each rule defines a fuzzy region in the feature plane; the identification of these fuzzy regions with the corresponding fuzzy rule number is given in figure 14(b).



**Figure 12:** Universes of discourse and membership functions for the approximation and detail signals. The Greek symbol  $\mu$  is used to represent the degree of membership value.



**Figure 13:** Two-dimensional hyperspace or FAM generated by combining the fuzzy sets for the approximation and detail signals using the max-product composition.

### 3.2 Feature extraction and comparison analysis results

In this section results from feature extraction and comparison analysis are presented both by test and by tie bar for each one of the measured parameters for tests 2 to 7 (the analysis was not carried out for test 1 due that data results for this test were not provided). Note that a feature vector was extracted for every cycle of a given signal.

In order to perform comparison between extracted feature vectors, two measures were defined. The angle between two vectors  $x$  and  $y$  is defined as:

$$\theta(x, y) = \cos^{-1} \left( \frac{(y^T x)}{\|x\| \|y\|} \right) \quad (59)$$

and the Model Assurance Criterion (MAC) value is defined as:

$$MAC(x, y) = \frac{|x^T y|^2}{(x^T x)(y^T y)} \quad (60).$$

Rules

- R1: Region 1 is activated if (cA<sub>1</sub> is NB and cD<sub>1</sub> is NB) R13: Region 13 is activated if (cA<sub>1</sub> is PS and cD<sub>1</sub> is NB)  
R2: Region 2 is activated if (cA<sub>1</sub> is NB and cD<sub>1</sub> is NS) R14: Region 14 is activated if (cA<sub>1</sub> is PS and cD<sub>1</sub> is NS)  
R3: Region 3 is activated if (cA<sub>1</sub> is NB and cD<sub>1</sub> is PS) R15: Region 15 is activated if (cA<sub>1</sub> is PS and cD<sub>1</sub> is PS)  
R4: Region 4 is activated if (cA<sub>1</sub> is NB and cD<sub>1</sub> is PB) R16: Region 16 is activated if (cA<sub>1</sub> is PS and cD<sub>1</sub> is PB)  
R5: Region 5 is activated if (cA<sub>1</sub> is NM and cD<sub>1</sub> is NB) R17: Region 17 is activated if (cA<sub>1</sub> is PM and cD<sub>1</sub> is NB)  
R6: Region 6 is activated if (cA<sub>1</sub> is NM and cD<sub>1</sub> is NS) R18: Region 18 is activated if (cA<sub>1</sub> is PM and cD<sub>1</sub> is NS)  
R7: Region 7 is activated if (cA<sub>1</sub> is NM and cD<sub>1</sub> is PS) R19: Region 19 is activated if (cA<sub>1</sub> is PM and cD<sub>1</sub> is PS)  
R8: Region 8 is activated if (cA<sub>1</sub> is NM and cD<sub>1</sub> is PB) R20: Region 20 is activated if (cA<sub>1</sub> is PM and cD<sub>1</sub> is PB)  
R9: Region 9 is activated if (cA<sub>1</sub> is NS and cD<sub>1</sub> is NB) R21: Region 21 is activated if (cA<sub>1</sub> is PB and cD<sub>1</sub> is NB)  
R10: Region 10 is activated if (cA<sub>1</sub> is NS and cD<sub>1</sub> is NS) R22: Region 22 is activated if (cA<sub>1</sub> is PB and cD<sub>1</sub> is NS)  
R11: Region 11 is activated if (cA<sub>1</sub> is NS and cD<sub>1</sub> is PS) R23: Region 23 is activated if (cA<sub>1</sub> is PB and cD<sub>1</sub> is PS)  
R12: Region 12 is activated if (cA<sub>1</sub> is NS and cD<sub>1</sub> is PB) R24: Region 24 is activated if (cA<sub>1</sub> is PB and cD<sub>1</sub> is PB)

(a)

Fuzzy Regions

<b>c D<sub>1</sub></b>	<b>P B</b>	R 4	R 8	R 1 2	R 1 6	R 2 0	R 2 4
	<b>P S</b>	R 3	R 7	R 1 1	R 1 5	R 1 9	R 2 3
	<b>N S</b>	R 2	R 6	R 1 0	R 1 4	R 1 8	R 2 2
	<b>N B</b>	R 1	R 5	R 9	R 1 3	R 1 7	R 2 1
		<b>N B</b>	<b>N M</b>	<b>N S</b>	<b>P S</b>	<b>P M</b>	<b>P B</b>
	<b>c A<sub>1</sub></b>						

(b)

**Figure 14:** (a) Rules generated by combining the fuzzy sets for the detail and approximation signals. (b) Fuzzy regions and corresponding fuzzy rule number.

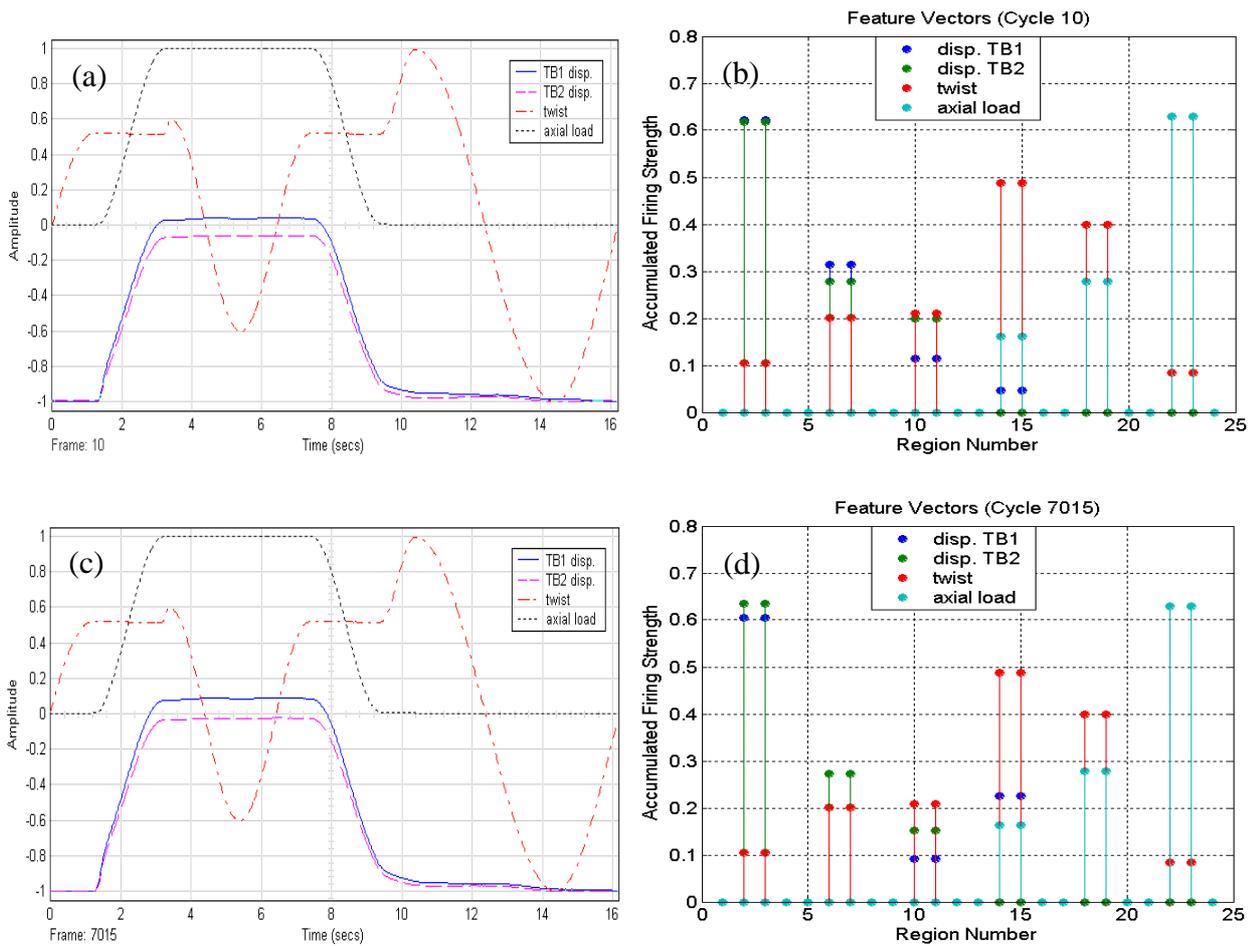
Therefore, if a periodic (cyclic) signal is not changing over time and a feature vector is extracted for every cycle of the signal, then the angle between two feature vectors corresponding to two cycles of the signal taken at different periods of time should be near to zero (and the MAC value should be near to one). Furthermore, if a cycle is taken as a reference signal and its corresponding feature vector is compared with the feature vectors extracted from cycles ahead in time, then in addition to the angles being near to zero the variation of the angle values should be small as well. On the contrary, if the signal is changing over time, e.g. the signal amplitude is bigger and bigger at each consecutive cycle, then this change will be indicated by a growing angle value (and a decreasing MAC value) between the reference feature vector and the feature vectors corresponding to cycles ahead in time.

First, examples of two cycles of the measured signals and their corresponding extracted feature vectors are shown in figure 15. Both cycles are from measurements taken in test 3. The comparison analysis results are presented in the next sections.

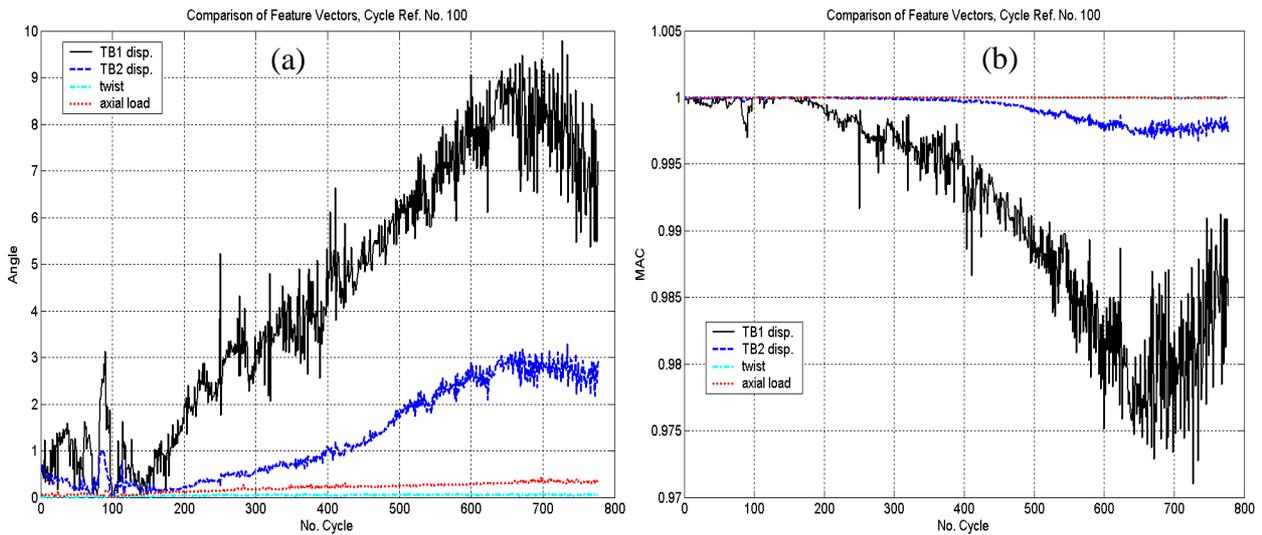
### 3.2.1 Comparison analysis by test

In this section the results from a comparison analysis performed by test are presented. A comparison analysis by test is performed by selecting a reference feature vector and calculating the angle and MAC values when it is compared with the remaining feature vectors among a test. This will produce two curves referred to as angle feature vector comparison analysis curve and MAC

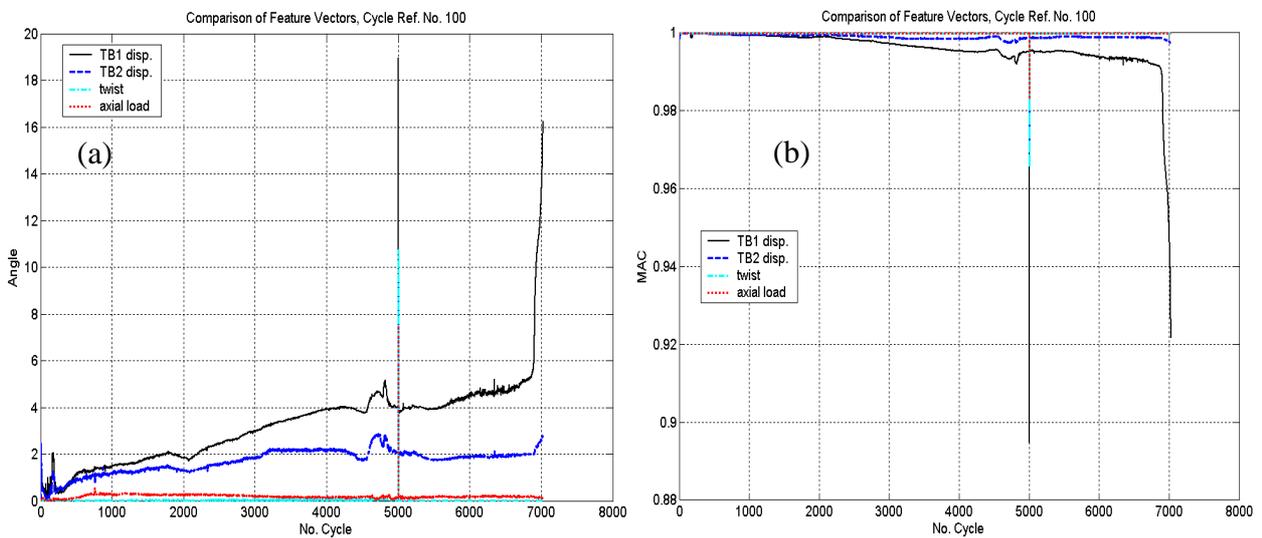
feature vector comparison analysis curve, respectively. Therefore, 12 comparison analyses are presented, corresponding to feature extraction vectors obtained for tests 2 to 7. In all of these comparison analyses the feature vector corresponding to cycle number 100 has been selected as the reference feature vector. This cycle number is selected considering that a warming up process may be present at the start of each test and 100 cycles are enough for the test responses to settle down. In addition, it is assumed that the signals obtained in this cycle are representative of the signals obtained for the undamaged state of the corresponding tie bars. Figures 16 to 21 show the feature vector comparison analysis curves corresponding to tests 2 to 6. Note that, as the tests were performed over a pair of tie bars, the term TB1 refers to tie bar 1 and the term TB2 refers to tie bar 2. The serial numbers of the corresponding tie bars under test and identified as TB1 and TB2 are indicated in each figure.



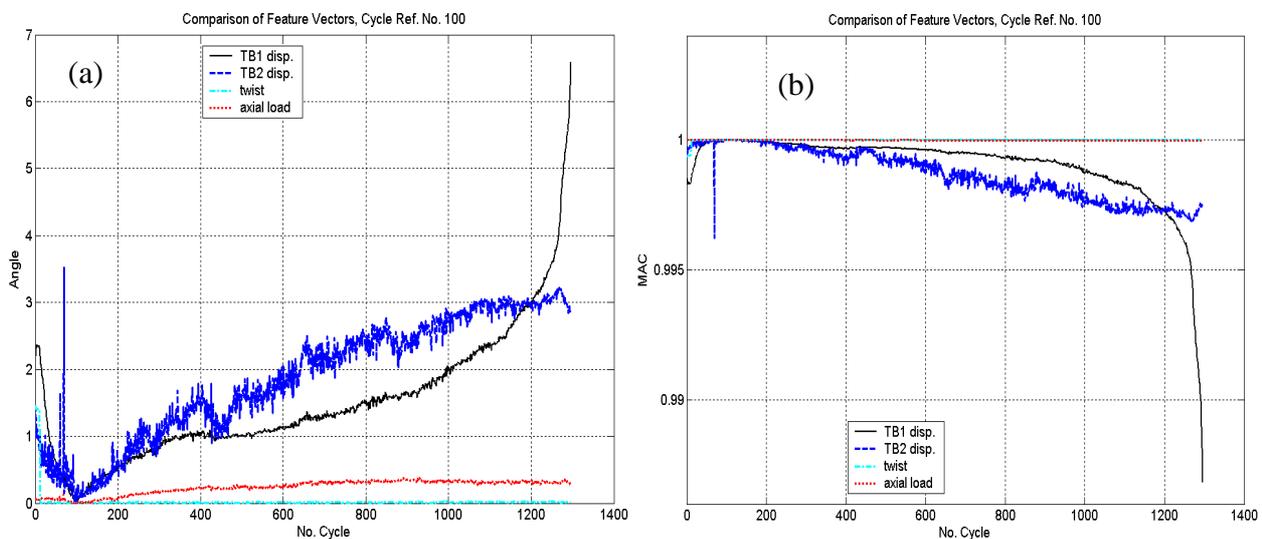
**Figure 15:** (a) Measured signals at cycle 10 and (b) corresponding feature vectors. (c) Measured signals at cycle 7015 and (d) corresponding feature vectors. Both cycles are from measurements taken in test 3.



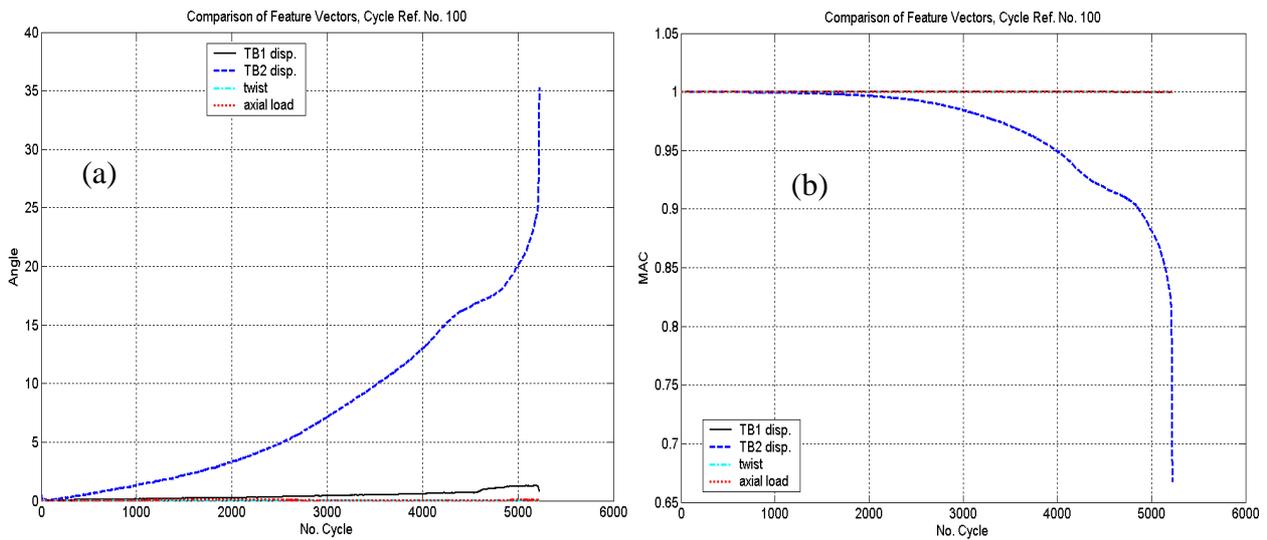
**Figure 16:** Test 2, TB1=AE7119, TB2=BAH4263, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.



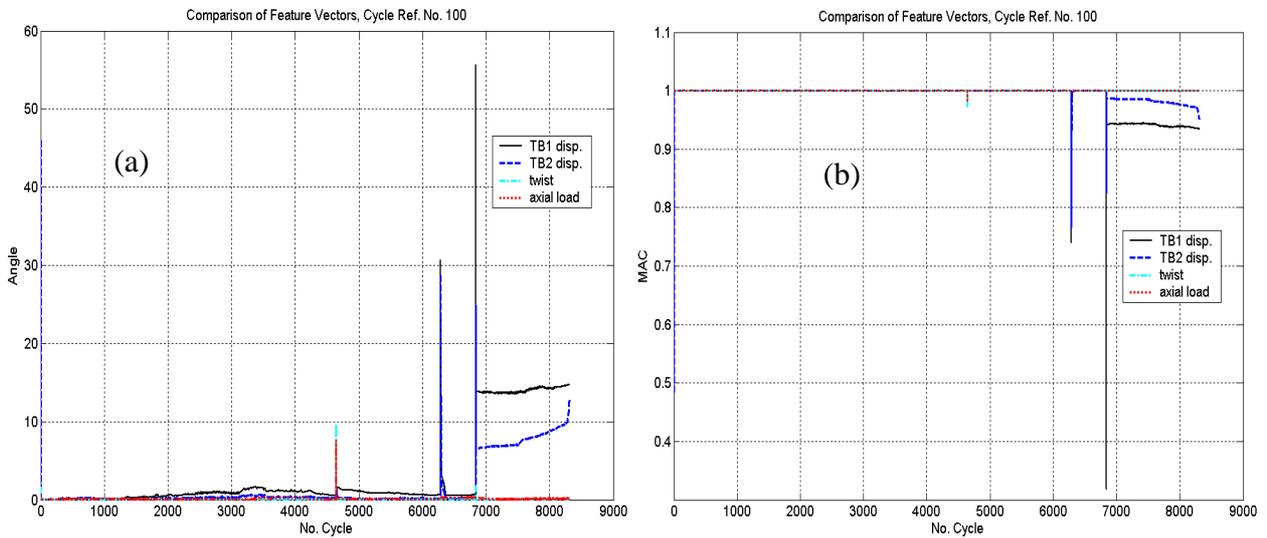
**Figure 17:** Test 3, TB1=AE7119, TB2=AEX5714, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.



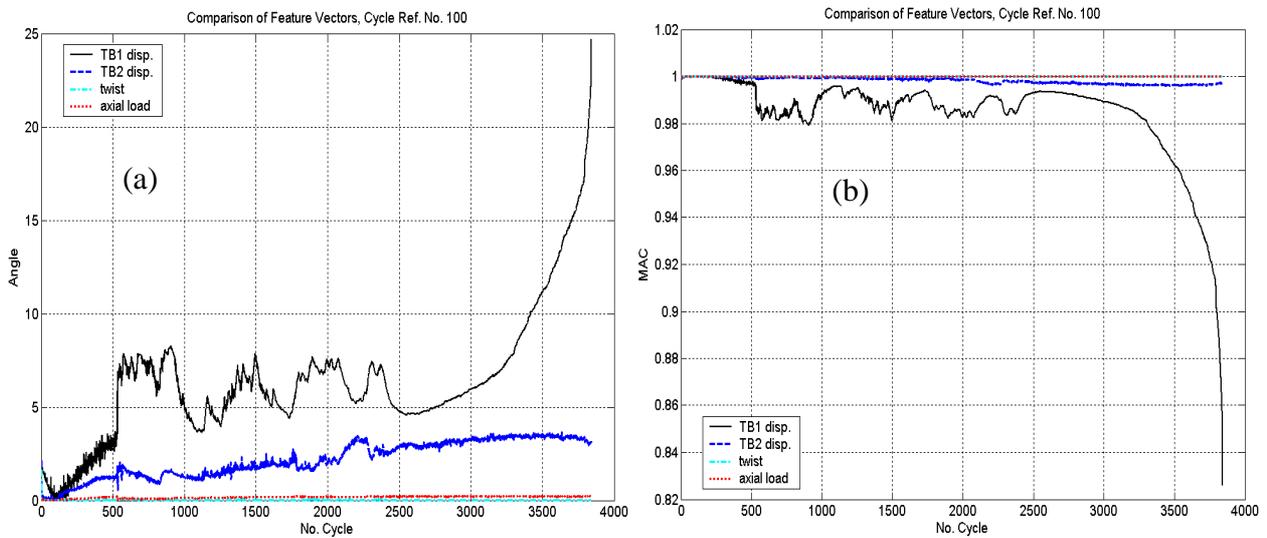
**Figure 18:** Test 4, TB1=LJA0399, TB2=AEX5714, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.



**Figure 19:** Test 5, TB1=LK0034, TB2=LK0046, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.



**Figure 20:** Test 6, TB1=LJA0401, TB2= LJA1440, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.



**Figure 21:** Test 7, TB1=LJA0401, TB2= LJA2061, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve.

### 3.2.2 Comparison analysis by tie bar

In this section the results from the comparison analysis performed by tie bar are presented. A comparison analysis by tie bar of the extracted feature vectors is performed considering individual tie bars among all the tests. Note from Table I that several of the tie bars are used only in one test, thus its comparison analysis by tie bar is the same as the corresponding comparison analysis by test, which means that their results have already been presented in the previous section. The tie bars which comparison analysis by tie bar is the same as the comparison analysis by test are listed below:

Comparison analysis for tie bar BAH4263 is the same as that for test 2, shown in Figure 16.

Comparison analysis for tie bar LJA0399 is the same as that for test 4, shown in Figure 18.

Comparison analysis for tie bar LK0034 is the same as that for test 5, shown in Figure 19.

Comparison analysis for tie bar LK0046 is the same as that for test 5, shown in Figure 19.

Comparison analysis for tie bar LJA1440 is the same as that for test 6, shown in Figure 20.

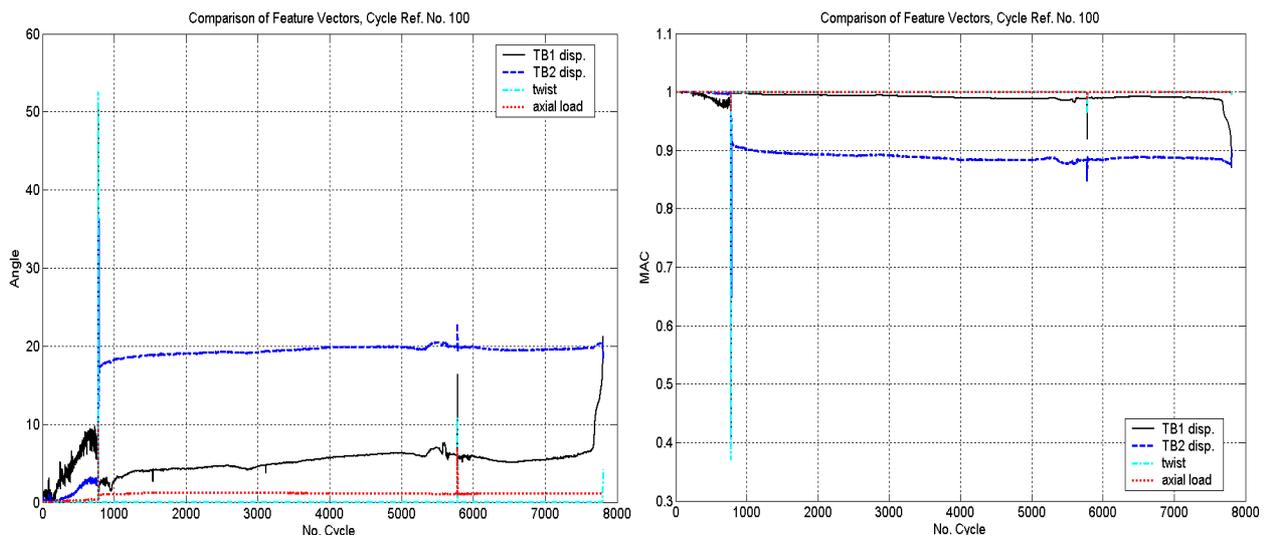
Comparison analysis for tie bar LJA2061 is the same as that for test 7, shown in Figure 21.

Hence, only three tie bars were employed in more than two tests. In order to appreciate the change in the feature vectors over the whole utilisation of these tie bars, a comparative analysis is performed taken as reference the feature vector extracted from cycle 100 and comparing it with the remaining feature vectors for the whole utilisation of the tie bars.

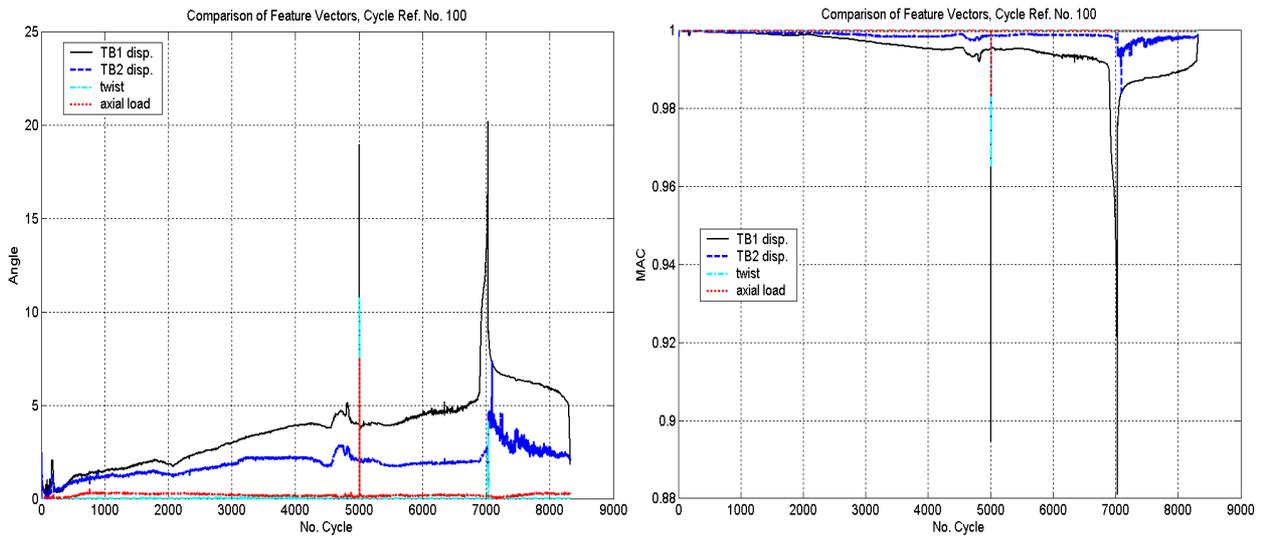
The comparison analysis for tie bar AET7119 is the combination of test 2 and test 3 and it is shown in figure 22.

The comparison analysis for tie bar AEX5714 is the combination of test 3 and test 4 and it is shown in figure 23.

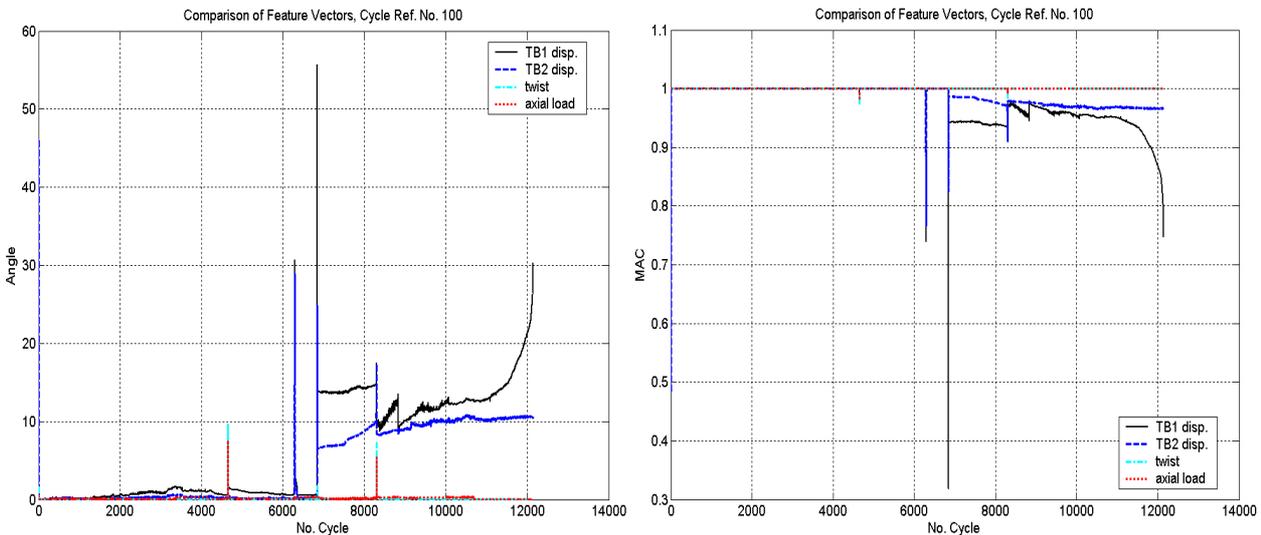
Finally, the comparison analysis for tie bar LJA0401 is the combination of test 6 and test 7 and it is shown in figure 24.



**Figure 22:** TB1=AET7119, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve. The first 778 cycles correspond to test 2, while the next 7026 cycles correspond to test 3.



**Figure 23:** TB2= AEX5714, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve. The first 7026 cycles correspond to test 3, while the next 1296 cycles correspond to test 4.



**Figure 24:** TB1= LJA0401, (a) angle feature vector comparison analysis curve, (b) MAC feature vector comparison analysis curve. The first 8307 cycles correspond to test 6, while the next 3845 cycles correspond to test 7.

### 3.3 Analysis of results

The idea of performing a comparison analysis is to look at the dynamic behaviour of the tie bars when approaching failure through looking at the changing behaviour of the feature vectors over time. Thus, from the analysis of figures 16 to 24 several observations can be made. These observations are made considering the angle as comparison measure, while similar interpretations can be made using the MAC value as comparison measure:

1. A clear pattern when approaching failure can be appreciated in the comparison analysis curves corresponding to the change of extracted feature vectors from displacement signals of tie bars AET7119, LJA0399, LJA1440, and LJA0401, corresponding to analysis results of tests 3, 4, 6 and 7, respectively. This pattern emerges as a fast increment (resembling a monotonically increasing function) in the angle values between the reference feature vector

from an undamaged condition, extracted from a displacement cycle in an early stage of the test, and the feature vectors extracted from displacement signal cycles approaching the point of failure later in time. This pattern also can be observed as a markedly change in the slope of the curve representing the angle values between the reference feature vector and the feature vectors corresponding to displacement cycles near the point of failure. This change in slope is more drastic in tests 3 and 6 than the observed in tests 4 and 7.

2. The comparison analysis curves corresponding to the displacement signals for tie bars AET7119 and LJA1440 in tests 3 and 6 resemble step functions near the point of failure. On the other hand, the comparison analysis curves corresponding to the displacement signals for tie bars LJA0399 and LJA0401 in tests 4 and 7 resemble exponential functions near the point of failure.
3. By inspecting the angle feature vector comparison analysis curve, it can be said that, while the dynamic behaviour of the tie bars looks linear far from the point of failure, this behaviour shifts to a non-linear behaviour near the point of failure.
4. For the case of test 2 and tie bars AET7119 and BAH4263, no appreciable pattern can be appreciated in the comparison analysis results for the displacement signals. However, a high variability is observed on the angle values obtained when the comparison analysis is performed, meaning that a high variability exists in the feature vector values extracted from the displacement signals. In addition, it is noticed in this test that the extracted feature vectors, corresponding to the axial load cyclic signal, slowly but constantly change over time. This means that the amplitude of the applied axial load signal was also changing over the test.
5. For the case of test 5 and tie bars LK0034 and LK0046, even though none of them failed, a pattern is clearly appreciated in the feature vectors corresponding to the displacement signals of tie bar LK0046. This behaviour is similar to that observed in tie bars LJA0399 and LJA0401, where an exponential-like curve of the angle comparison analysis curve is appreciated. This may mean that tie bar LK0046 was near the point of failure when the test was stopped.
6. The same patterns explained in the points 1 to 4 above can be appreciated in the comparison analysis by tie bar plots.
7. It was noted that some of the tests had been interrupted and then reinitiated maintaining the same tie bars under tests. These interruptions are captured by the feature extraction algorithm and reflected in the form of transients in the displacement comparison analysis curves. When looking directly to the corresponding original data plots, it is observed that these results are obtained because the initial cycles of the restarted tests are distorted and, therefore, quite different from the reference cycles.
8. Although the dynamic behaviour of the tie bars near the point of failure has similarities, e.g., a monotonically increasing angle feature vector comparison analysis curve and non-linearity, their dynamic behaviour is quite different and particular to each tie bar, e.g. each one fails at a different number of cycles.
9. From the comparison curves, it is observed that small variations in the axial load have a big impact in the displacement signal response.

### 3.4 Validation of the SCFEA

In this section the proposed SCFEA is validated using an adaptive neuro-fuzzy inference system (ANFIS) network [Jang *et al.*, 1997]. The validation process consists in performing system identification of the unknown system, the analysed tie bar, employing an ANFIS network using a set of available input-output patterns, e.g. twist and axial load measurements as inputs and displacement measurements as output, as is illustrated in figure 25(a). The structure of the employed ANFIS network is shown in figure 25 (b). Once the ANFIS network has been trained, it

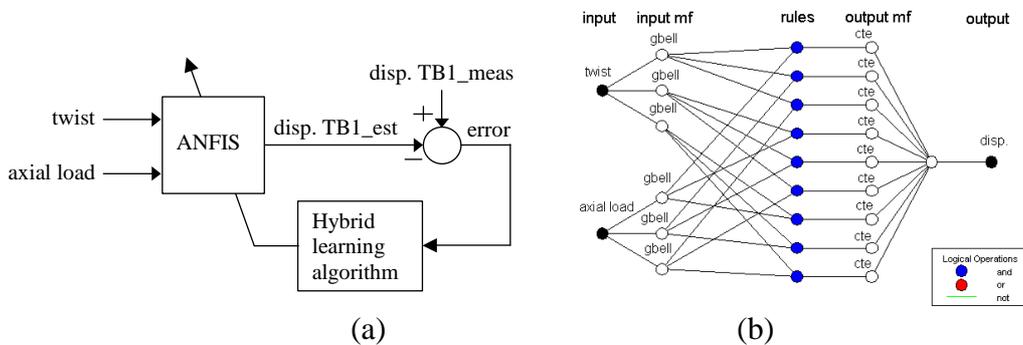
is used to predict the output of the system for the unseen input patterns. Therefore, the root mean square error (RMSE) value is calculated for every cycle of the input signals. These results are plotted to obtain a RMSE evolution curve. This curve and the angle evolution curve from the soft computing feature extraction approach are correlated in order to determine if they show a linear relationship. If this is the case, then the correlation coefficient will be near to 1 meaning that both approaches are consistent and this will validate the soft computing feature extraction approach.

The ANFIS network shown in figure 25(b) was implemented and simulated using the MATLAB/Simulink simulation environment. Measurement data from test 4 were used to train two ANFIS networks (one for each tie bar). Table II shows the number of cycles used as patterns for training and for validation, the number of epochs needed to train the ANFIS networks, the RMSE value obtained after training, and the RMSE values obtained after validation.

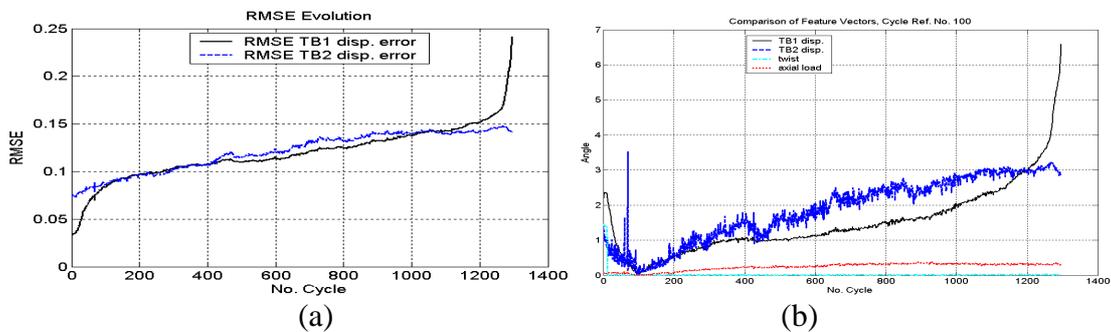
**Table II:** ANFIS RMSE training and validation results

	No. Cyc. Training	No. Cyc. Testing	Epochs Training	RMSE Training	RMSE Testing
Disp TB1	10 (10-19)	10 (20-29)	500	0.034509	0.03718
Disp TB2	10 (10-19)	10 (20-29)	500	0.074661	0.07469

The RMSE evolution curve and the angle evolution curve corresponding to each tie bar tested in test 4 are shown in Figures 26(a) and 26(b), respectively. Note in these figures that both curves are very similar in shape and trend. The correlation coefficients between the corresponding curves for tie bar 1 (LJA0399) and tie bar 2 (AEX5714) are  $r(RMSE\ TB1, ang.\ TB1\ disp) = 0.9792$  and  $r(RMSE\ TB2, ang.\ TB2\ disp) = 0.9653$ , respectively. Therefore, these values show that the corresponding curves for both tie bars are highly correlated, meaning that the results from two different approaches are similar, thus validating the SCFEA approach.



**Figure 25:** (a) System identification using an ANFIS network; (b) ANFIS structure.



**Figure 26:** (a) RMSE evolution curve; (b) angle evolution curve.

### 3.5 Tie bar critical degradation detection

In this section results of work carried out to develop a tie bar critical degradation detection algorithm based on pattern recognition of the features extracted by the SCFEA are presented. Two approaches have been considered, statistical process control (SPC) and monotonically increasing function detection.

SPC is a technique commonly used to monitor the manufacturing process in order to reduce variability and build quality into the product [Montgomery, 2005]. In the context of structural damage detection, the SPC method consists in monitoring the variability of some characteristic sensitive to damage in order to detect shifts or departures from an assumed state of health. For this technique, a control chart is built, which is a graphical display of the damage-sensitive feature that has been measured or computed versus the sample number or time. The control chart contains a centre line (CL), an upper control limit (UCL), and a lower control limit (LCL) defined based on the statistics of the monitored feature [Montgomery, 2005]. The CL represents the average value of the measured feature corresponding to a healthy state of the structure. The UCL and LCL are chosen so that if no damage is present in the structure, then nearly all of the measured features will fall between the control limits. Thus, as long as the feature measurements plot within the control limits, the structure is assumed to be in a healthy state. However, a measured feature that plots outside of the control limits is interpreted as evidence that the structure has evolved to a state of damage [Sohn *et al.*, 2000]. In the case of the tie bar, the feature assumed as sensitive to damage and thus the quantity to be monitored, is the angle between the reference feature vector and the remaining extracted feature vectors corresponding to the displacement signal.

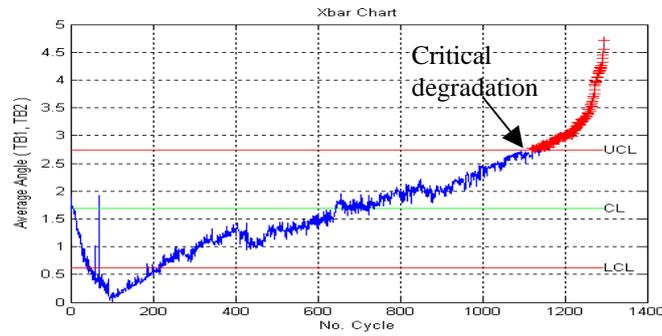
The SPC method is exemplified using the features extracted from data measurements corresponding to test 4 (see section 3). As tie bars are tested in couples it is assumed that there are two sets of extracted feature vectors corresponding to two sets of independent measurements. Thus, in order to build the control chart, the angle mean value denoted as  $\mu_{disp}$  and the angle standard deviation value denoted as  $\sigma_{disp}$  are calculated considering the angle value data obtained in section 3. The CL, which is given by the value  $\mu_{disp}$ , is calculated by first averaging the two angle measures for every cycle and then averaging all the obtained values. The standard deviation value  $\sigma_{disp}$  is calculated based on the averages of angle values obtained for every cycle. The UCL and LCL are then obtained based on the  $\sigma_{disp}$  value as:

$$UCL_{disp}, LCL_{disp} = \mu_{disp} \pm Z_{\alpha} \left( \frac{\sigma_{disp}}{\sqrt{q}} \right) \quad (61)$$

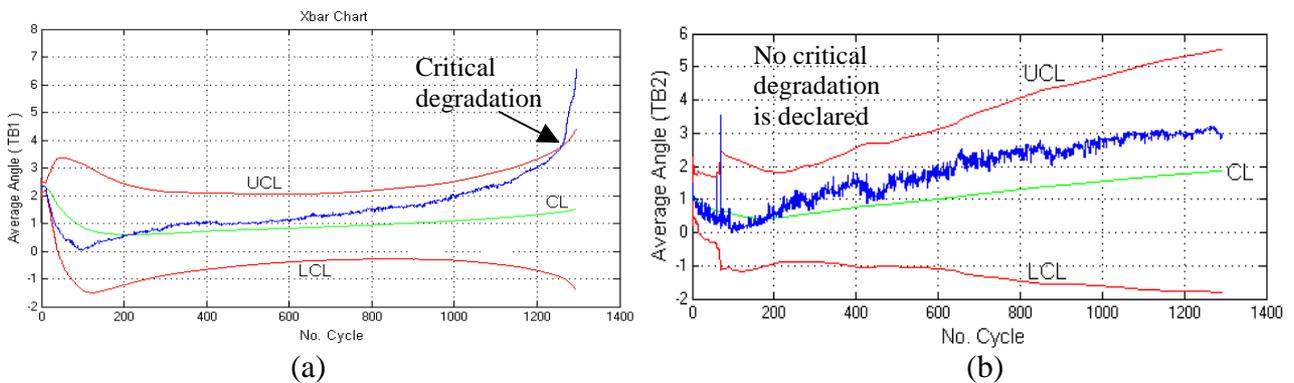
where  $Z_{\alpha}$  is the value of a standard normal distribution with zero mean and unit variance such that the cumulative probability is  $100(1-\alpha)\%$ , in this case  $Z_{\alpha} = 3$  which represent the 99.73% confidence;  $q$  is the size of each subgroup, here as there are two sets of measurements 2 groups are assumed. The generated control chart is shown in figures 27. Critical degradation detection is declared when the average measured angle value plots beyond the UCL line for more than two consecutive cycles. Note that values below the LCL do not indicate critical damage; in fact lower angle values indicate marginal change between feature vectors.

The control limits in the chart shown in figure 27 are fixed and determined considering the  $3\sigma$  limits from the statistics of the whole set of angle measurements derived by comparing the extracted feature vectors from the displacement signals from both tie bars. A second analysis can be performed considering each set of measurements separately. This can be achieved by building two control charts, one for each tie bar and corresponding extracted feature vectors. In this case it is proposed to calculate the CL, UCL and LCL in a dynamic manner. This is, the angle mean value

corresponding to each tie bar  $\mu_{disp}$  is calculated using the running mean, which means that the mean value is updated every cycle and so there are the control limits UCL and LCL. The resultant control charts for the two tie bars used in test 4 are shown in figure 28. Note that, as in the previous case, critical damage detection is declared when more than two consecutive points plot beyond the UCL.

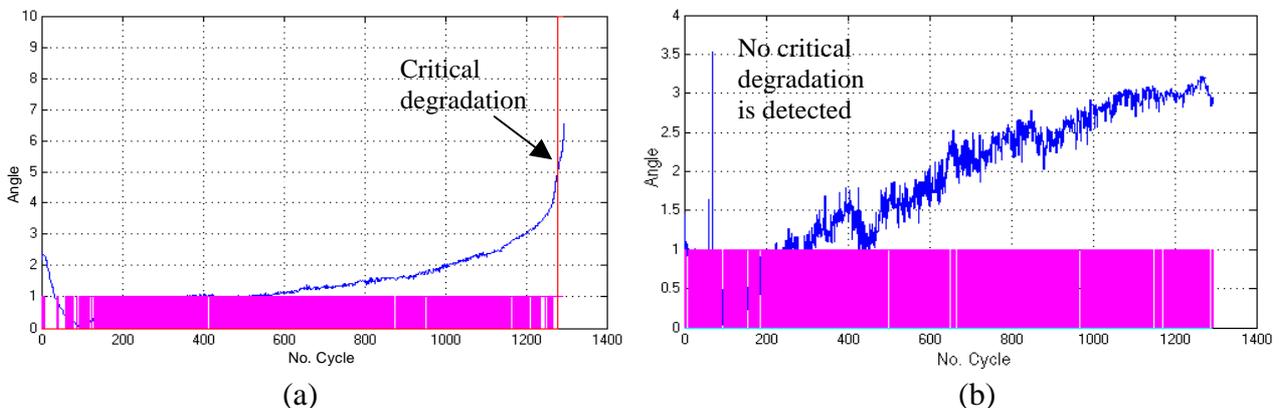


**Figure 27:** Control chart fixed control limits.



**Figure 28:** Control charts using dynamic control limits; (a) TB1, (b) TB2.

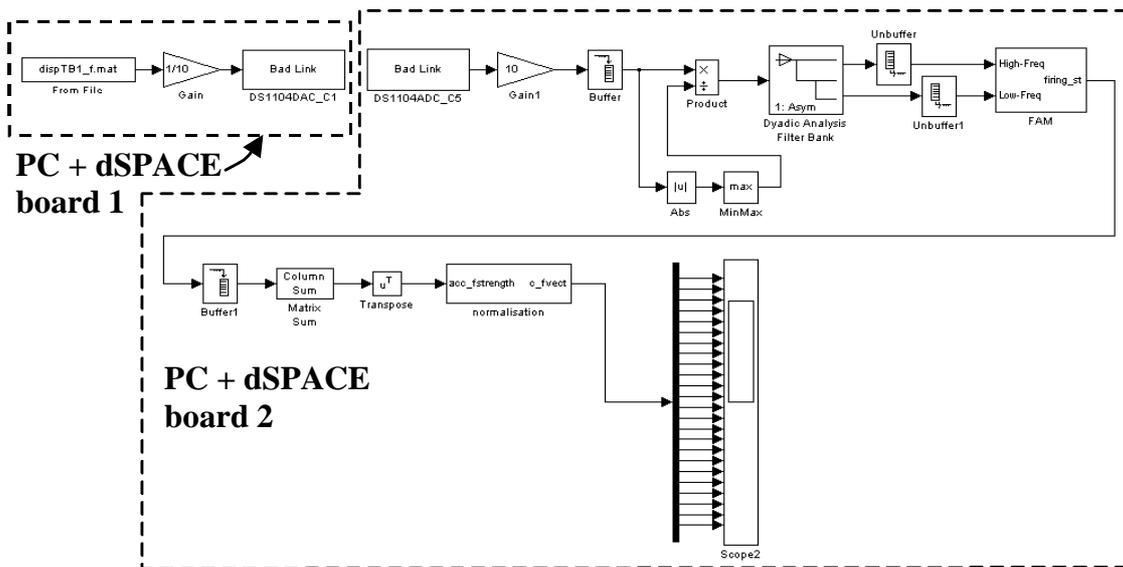
The second approach to detect critical damage detection is based on detecting if the angle feature vector comparison analysis curve represents a strictly increasing function. A function  $f$  is strictly increasing if, whenever  $x < y$ , then  $f(x) < f(y)$ . Thus the critical degradation detection algorithm uses the reference feature vector and compares it with the current cycle displacement feature vector to track and identify degraded tie bar condition. The angle magnitude resultant of this comparison is continuously monitored. If in time (as the cycle number increase) the angle feature vector comparison curve approaches a strictly increasing function for an interval of more than 20 cycles, then a critical degradation of the tie bar is declared. Figure 29 shows the result of this method applied to the extracted features from the tie bars used in test 4.



**Figure 29:** Critical degradation detection by detecting strictly increasing function; (a) TB1, (b) TB2.

### 3.6 Real time simulation

Simulation in real-time of the proposed SCFEA applied to tie bar data was carried out using the Simulink-dSPACE simulation environment (DS1104 R&D Controller Board) in order to investigate the size of memory requirements for algorithm implementation. The algorithm was implemented up to the stage of reporting the extracted feature vectors. In this simulation two PC's with dSPACE systems connected to them were used as is represented in figure 30. The first PC-dSPACE system was used to send-out, in real time, through a DAC, measured tie bar data. This simulated the process of obtaining measurements from an actual tie bar test. The second PC-dSPACE system was used to read the data through an ADC and run the feature extraction algorithm. Thus, only the SCFEA algorithm was compiled and loaded to the second dSPACE board. The size of the compiled object file .ppc (loaded to the dSPACE board) was 575 KB and the maximum processing time was 25  $\mu$ s.



**Figure 30:** Simulink model used for real time simulation of the SCFEA applied to tie bar data.

## 4 Soft computing feature extraction algorithm applied to pitch link data

In this section, results of feature extraction and critical wear detection algorithms developed to detect damage in the bearing system of pitch link (referred here to as pitch link system) part of the main rotor hub of a Lynx Helicopter are presented. The algorithm consists of two stages: feature extraction and statistical model development for feature discrimination. The feature extraction part of the algorithm is based on the SCFEA proposed in Section 2.3. The algorithm is applied to strain time-histories collected at the University of Bristol (UB) from experiments where two pitch links, one with negligible wear and the other noticeably worn, were subjected to a sinusoidal stress cycle of tension and compression with forces representative of those encountered in operation. Results show that undamaged (unworn) and damaged (worn) pitch links can be detected and successfully classified.

In the previous sections a feature extraction algorithm, referred to as SCFEA combining wavelet theory and fuzzy logic theory was proposed to monitor the state of degradation of tie bars. The structure of the proposed SCFEA is used in this section to extract damage sensitive features to detect damage in the bearing system of pitch link. In this development it is assumed that strain response signals from undamaged (unworn) and damaged (worn) systems are available for comparison and classification. Furthermore, it is assumed that the strain waveforms responses are cyclic signals with a known frequency.

## 4.1 Algorithm implementation

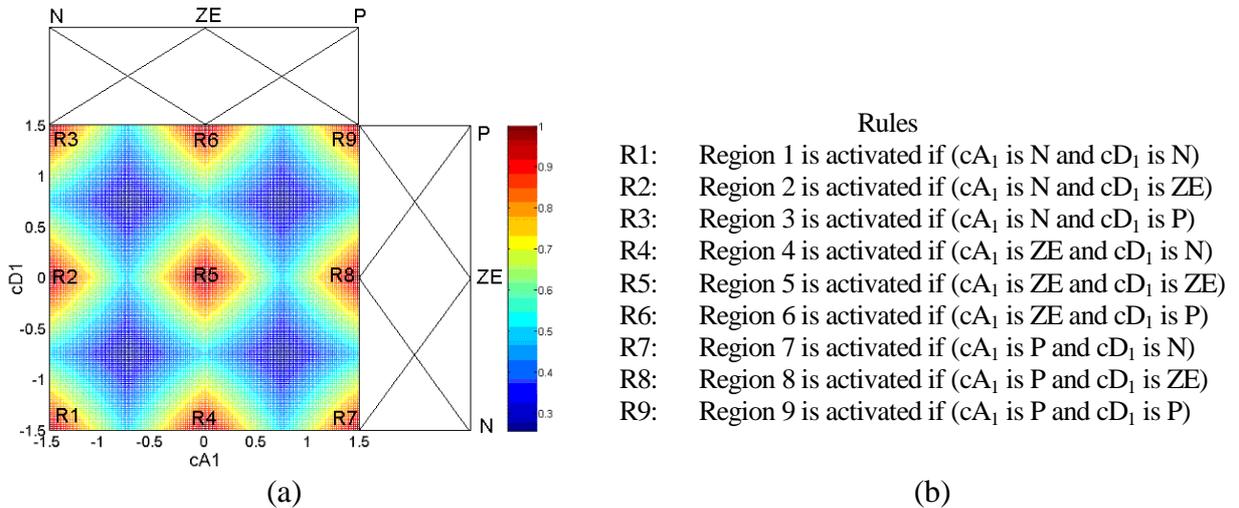
The SCFEA was implemented in the MATLAB/Simulink simulation environment and applied to pitch link data that was obtained as will be explained later. After normalisation, the DWT was implemented using the Haar wavelet decomposition filters. The approximation-detail data pairs  $(cA_1(t), cD_1(t))$  obtained from a single level DWT are used for feature extraction.

Two linguistic variables were defined  $x_A$  and  $x_D$  to represent the approximation and detail signals  $(cA_1(t), cD_1(t))$ , respectively. The corresponding base variables are denoted as  $cA$  and  $cD$  for  $x_A$  and  $x_D$ , respectively. The linguistic value sets  $T_A$  and  $T_D$  with the corresponding membership functions for  $x_A$  and  $x_D$  are given as follows:

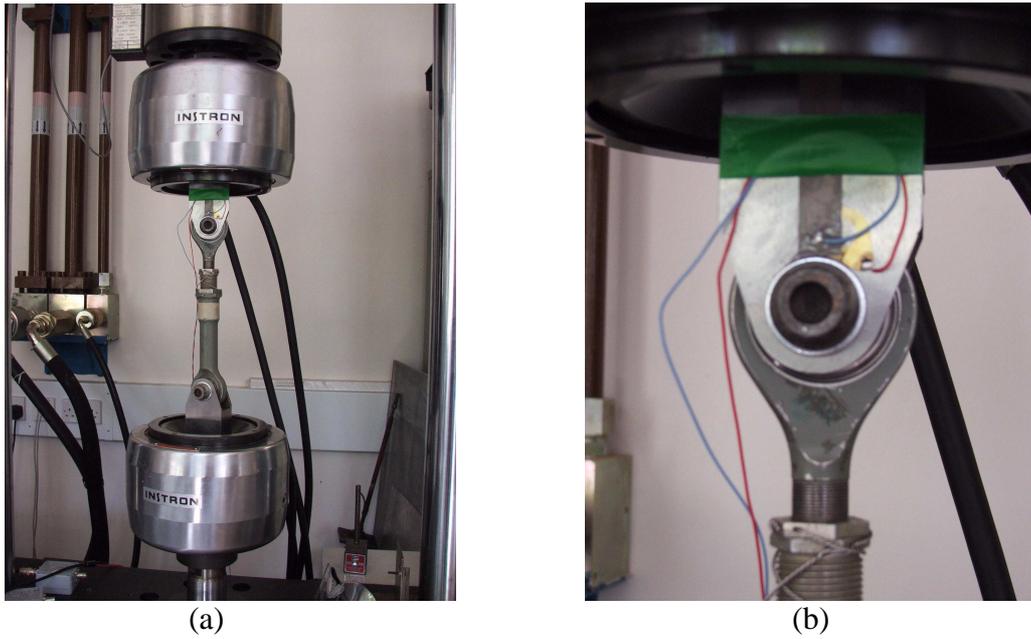
$$T_A = \begin{bmatrix} N \\ ZE \\ P \end{bmatrix} \text{ and } T_D = \begin{bmatrix} N \\ ZE \\ P \end{bmatrix}; \mu_A = \begin{bmatrix} \mu_N(cA(t)) \\ \mu_{ZE}(cA(t)) \\ \mu_P(cA(t)) \end{bmatrix} \text{ and } \mu_D = \begin{bmatrix} \mu_N(cD(t)) \\ \mu_{ZE}(cD(t)) \\ \mu_P(cD(t)) \end{bmatrix}$$

Thus, three fuzzy sets were defined for  $x_A$  and three fuzzy sets were defined for  $x_D$ . The fuzzy sets terms in both cases mean: N = Negative, ZE = Zero and P = Positive. Figure 31(a) shows the associated membership functions defined for  $x_A$  and for  $x_D$  together with the two-dimensional hyperspace generated by combining the fuzzy sets for the approximation and detail signals. Each one of the rules generated by combining the fuzzy sets for the detail and approximation signals is listed in figure 31(b).

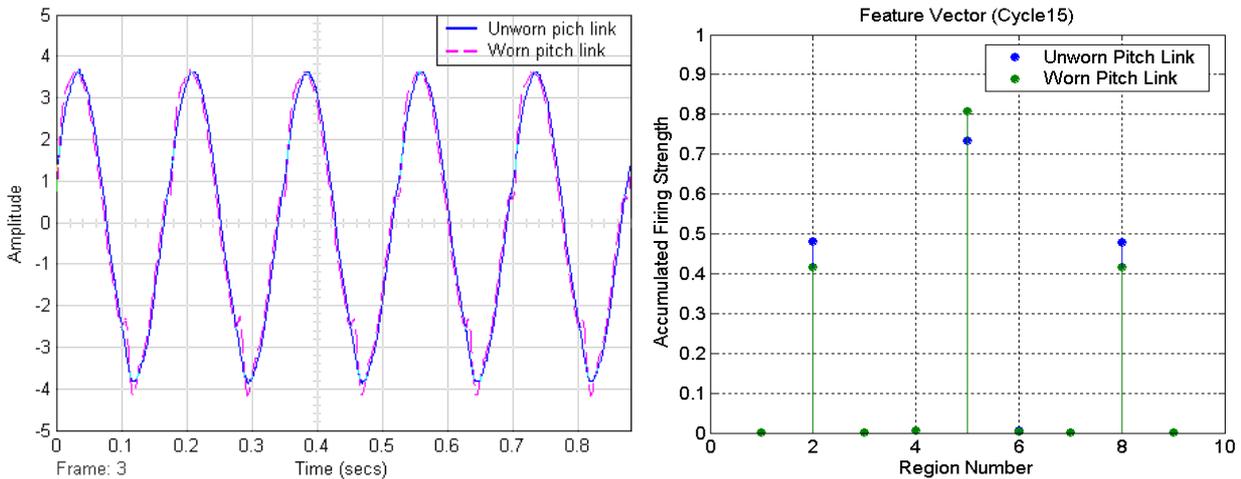
The algorithm was applied to strain time-histories collected from two experiments carried out at the University of Bristol (UB). In these experiments two pitch-links, one with negligible wear and the other noticeably worn, were set up in an Instron Machine as is shown in figure 32(a). The machine was set up to give a sinusoidal stress cycle of tension and compression, with forces representative of those encountered in operation. A piezo-ceramic patch sensor glued to the part holding the pin, see figure 32(b), was connected to a storage oscilloscope and the waveforms obtained from the worn and unworn pitch links were recorded. Examples of the strain response signals obtained are shown in figure 33(a). As the strain response signals present a cyclic behaviour, the SCFEA was applied to obtain a feature vector for every cycle of the recorded signals. Examples of the feature vectors obtained for an unworn and worn pitch links are shown in figure 33(b) for cycle 15.



**Figure 31:** (a) Membership functions for the fuzzy sets defined for  $x_A$  and for  $x_D$  together with the two-dimensional hyperspace generated by combining the fuzzy sets. (b) Rules generated by combining the fuzzy sets for the detail and approximation signals.



**Figure 32:** (a) Pitch link test set up in an Instron Machine; (b) Location of piezo-ceramic sensor.

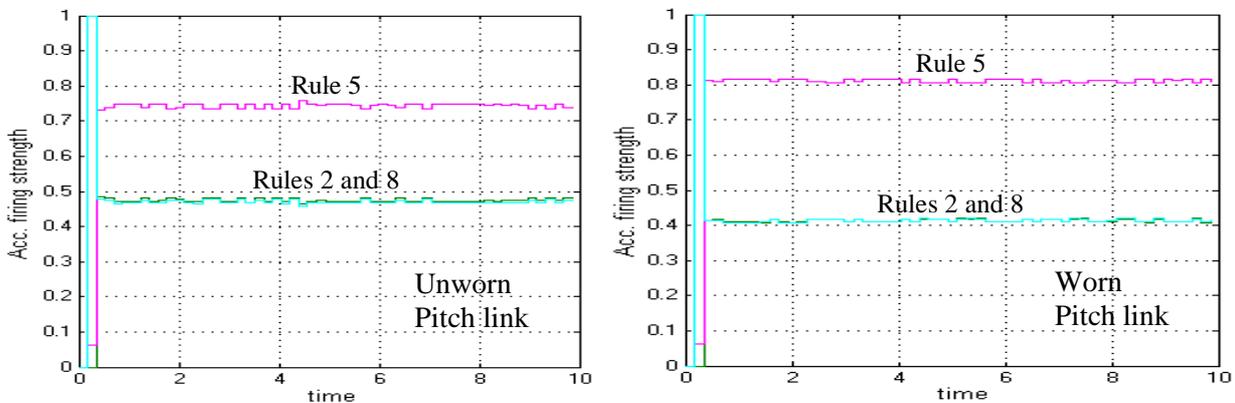


**Figure 33:** (a) Examples of the strain response waveforms obtained for unworn and worn pitch link systems; (b) Example of the feature vectors obtained for unworn and worn pitch link systems.

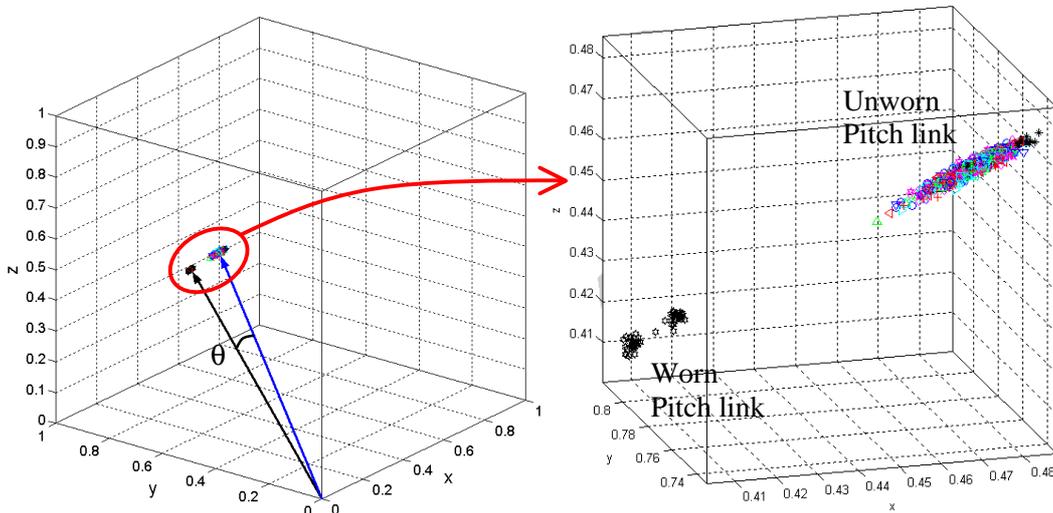
From experience of processing the tie bar data, it is known that not all the rules in the FAM are activated and sensitive to damage. Therefore, a feature (or rule) selection process is proposed based on the energy content of the rule (or fuzzy region) activation signals. Thus, by calculating the rule activation energies and comparing their magnitudes for unworn (undamaged) and worn (damage) cases, a subset of rules sensitive to damage can be selected. This means that the feature vector is reduced to contain the components with highest energy and sensitive to damage.

From the rule activation energy analysis it was detected that the rules R2, R5 and R8 have the highest content of energy while the remaining rules have negligible energy content. Therefore, the feature vector is reduced to contain only three values, the accumulated firing strengths corresponding to rules R2, R5 and R8. Once this vector is normalised to a unit vector, it becomes the output of the SCFEA. Note that this will reduce the computational burden required to process the algorithm due that, for real time implementation, the FAM will be formed only with these 3 rules (3 rules to be evaluated instead of 9 rules, in this case).

The time histories of the extracted feature vectors for undamaged (unworn) and damage (worn) pitch links strain response signals are shown in figure 34. In order to determine a reference feature vector, and as only the signal for one example of undamaged pitch link system was provided, simulation of undamaged systems was obtained by adding white noise (power spectrum density =  $1 \times 10^{-6}$ ) to the original undamaged signal. Ten undamaged pitch link systems using different random noise seeds were simulated. As each feature vector contains only three values (accumulated firing strengths for fuzzy rules 2, 5 and 8), it is possible to represent the feature vectors as points in a 3-dimensional (3-D) space. Figure 35 shows the 3-D representation of all the feature vectors corresponding to the unworn pitch link systems (obtained from the original signal plus the 10 obtained by adding white noise) and the corresponding feature vectors for a worn pitch link system. Note that a feature vector is obtained for every cycle of the signal (freq. approx. 5Hz), and this is carried out up to a length time of 10 seconds. Thus each point in figure 35 represents a feature vector extracted from each cycle of the respective strain response signal. From the same figure it is apparent that two classes can be clearly separated: unworn (undamaged) and worn (damaged) pitch link systems.



**Figure 34:** Feature vector time histories for unworn and worn pitch link systems.

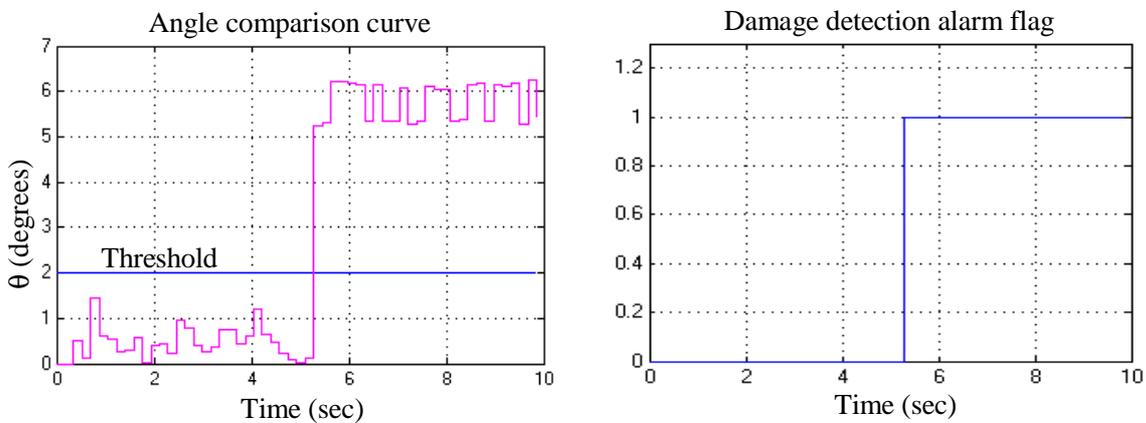


**Figure 35:** 3-D representation of the feature vectors obtained for the unworn and worn pitch link signals.

A reference or baseline feature vector was calculated by averaging the obtained feature vectors for all the unworn cases. The resultant reference vector was:  $R2=0.4719$ ,  $R5=0.7476$ ,  $R8=0.4672$ . Having available the reference feature vector, comparisons with the feature vectors for the worn and unworn signals were performed by calculating the angle  $\theta$  between them. This comparison was carried out in order to determine a threshold which separates damaged (worn) from undamaged

(unworn) systems. This threshold value was fixed at  $\theta = 2$  degrees (none of the unworn feature vectors when compared with the obtained reference feature vector went more than 2 degrees apart). Therefore, if an unknown state pitch link strain response signal is processed, then a damage detection alarm will be activated whenever the angle between the reference feature vector and the newly extracted feature vectors is equal to 2 or more degrees, and this comparison is performed cycle by cycle of the incoming signal. The damage detection alarm will indicate that the pitch link is critically worn.

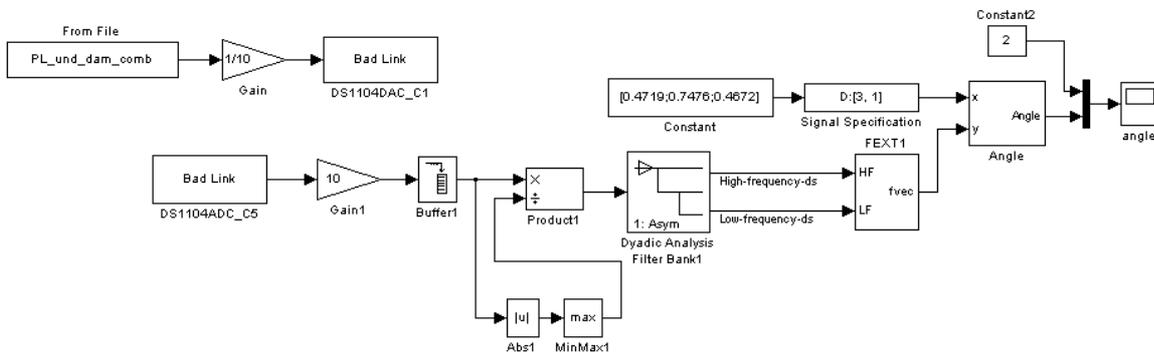
The effectiveness of the proposed feature extraction and damage detection algorithms was tested by processing the unworn and worn signals presented in figure 33(a). During the first 5 seconds of the simulation, the reference feature vector was compared with the feature vectors obtain for the unworn pitch link; while during the next 5 seconds of the simulation the reference feature vector was compared with the feature vectors obtain for the worn pitch link. Figure 36 presents the angle comparison curve and the damage detection alarm flag (0 = no damage, 1 = damage).



**Figure 36:** Damage detection results for the unworn and worn pitch link signals.

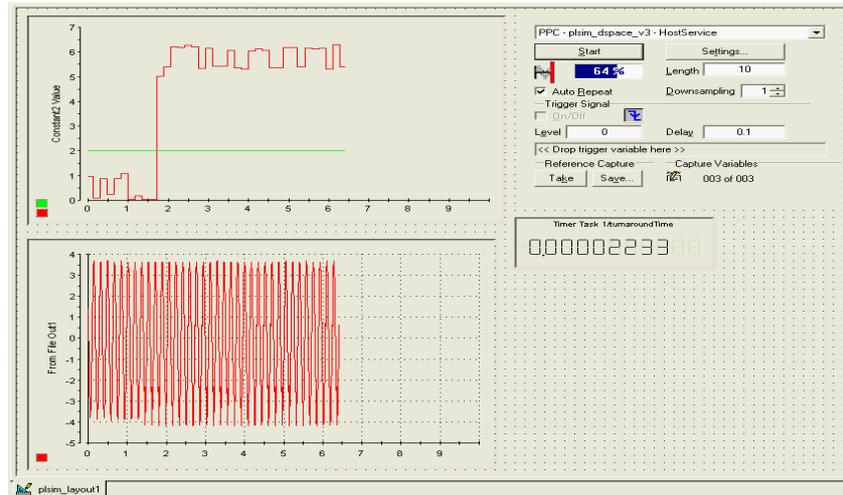
#### 4.2 Real time simulation

The damaged detection algorithm was simulated in real-time using the Simulink-dSPACE simulation environment (DS1104 R&D Controller Board). The implemented model is shown in figure 37, while the results are shown in figure 38.



**Figure 37:** Pitch Link damage feature extraction and damage detection model implemented in real time.

The time required to process the algorithm in real-time as reported by dSPACE (turn around time) is between 18 and 28 microseconds. Thus, the algorithm bandwidth is 36 KHz. The compiled object file (.ppc) loaded to the DSP processor in the DS1104 board has a size of 723 KB. Note that this file includes the data and the algorithm.



**Figure 38:** Damage detection results obtained from the real time simulation.

## 5 Conclusions and future work

In this document a feature extraction algorithm, referred to as soft computing feature extraction algorithm (SCFEA), has been proposed under the context of the WISD research project. The results of applying the SCFEA to data gathered from test performed on tie bar and pitch link components of the main rotor hub of a Lynx Helicopter have been presented. In the case of tie bar, results from data corresponding to six tests, where several tie bars were subjected to high level ground-air-ground (G-A-G) cyclic load testing until failure, have been presented. Comparison analyses of the extracted feature vectors, both by test and by tie bar, have been performed. Two measures were used to perform comparisons, the angle between vectors and the MAC value. From the comparison analyses curves, it is clear that a change can be appreciated in the dynamic behaviour of the tie bars when they are approaching failure. This change can be observed through the analysis of the variation of the extracted feature vectors from the displacement signals. The results presented indicate that a pattern can clearly be seen in the comparison analysis curve in 5 of the 6 tests. This demonstrates the applicability of the proposed approach for the task of feature extraction for Tie bar data. Additionally, two methods for pattern recognition and critical degradation detection of tie bar have been proposed. The first method uses techniques from statistical process control, while the second is based on detecting strictly increasing function. Both methods appear to detect critical degradation of tie bars. However, a robust analysis of these proposals is needed based on the statistical analysis considering a broader set of tie bars.

The SCFEA also was applied to strain data histories responses from unworn and worn bearing systems of pitch link. Results show that the angle between the feature vectors extracted from signals corresponding to an unworn pitch link and the feature vectors extracted from signals corresponding to a worn pitch link can be used to discriminate between these two classes. Based on the average feature vector for unworn pitch link and the average angle between this and feature vectors from worn pitch link a threshold angle value was selected. Thus, values beyond the fixed threshold were declared to come from a critically worn pitch link. Both off-line and real-time simulations of the proposed feature extraction and damage detection approaches have demonstrated their applicability.

Although promising results have been obtained, it is necessary to clarify that the available data corresponds to controlled experiments performed on tests rigs and not from the actual rotorcraft. A test rig for the pitch link system is currently under construction and it is expected to deliver data, which will be closer to the expected in a real rotorcraft environment. At the moment it is uncertain

if data from Tie bar system in the rotorcraft environment will be able to be measured. This means that the proposed feature extraction approach will necessarily need some adjustment, modifications or tuning to perform correctly in the real environment. The same may be said for the proposed pattern recognition (damage detection) approaches. Therefore, the development of the algorithms is an ongoing task, which means that the algorithms will need to be updated and tested each time new data are available.

## References

- Brown, M. and Harris, C.J. (1994). *Neurofuzzy adaptive modelling and control*, Prentice-Hall International, Hemel Hempstead, UK.
- Burrus, C.S., Gopinath, R.A., and Guo, H. (1998). *Introduction to wavelets and wavelet transforms a primer*, Prentice Hall, Upper Saddle River, New Jersey, USA.
- Coca, D. and Billings, S.A. (2001). *System identification using wavelets*, Department of Automatic Control and Systems Engineering, University of Sheffield, Research Report No. 785.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, **41**, pp. 909-996.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM.
- Douka, E., Loutridis, S. and Trochidis, A. (2003). Crack identification in beams using wavelet analysis, *International Journal of Solids and Structures*, **40**, pp. 3557-3569.
- Farrar, C.R. and Worden, K. (2005). An introduction to structural health monitoring, *to appear in Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*.
- Farrar, C.R., Doebling, S.W., and Nix, D.A. (2001). Vibration-based structural damage identification, *Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*, **359**, pp. 131-149.
- Farrar, C.R., Park, G., Allen, D.W. and Todd, M.D. (2006). Sensor network paradigms for structural health monitoring, *Structural control and health monitoring*, **13**, pp. 210-225.
- Gorton, G.F. (2006). *Further ground-air-ground (G-A-G) testing of tie bars*, Internal report, Westland, Document No. FFB/L1446/I02.
- Goupillaud, P., Grossmann, A., and Morlet J. (1984) Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, **23** (1), pp. 85-102.
- Grossmann, A., and Morlet, J. (1984). Decomposition of Hardy functions into square integrable of constant shape, *SIAM J. Math. Anal.*, **15** (4), pp.723-736.
- Haase, M. and Widjajakusuma, J. (2003). Damage detection based on ridges and maxima lines of the wavelet transform, *International Journal of Engineering Science*, **41**, pp. 1423-1443.
- Hong, J.-C., Kim, Y.Y., Lee, H.C. and Lee, Y.W. (2002). Damage detection using the Lipschitz exponent estimated by the wavelet transform: applications to vibration modes of a beam, *International Journal of Solids and Structures*, **39**, pp. 1803-1816.
- Jang, J. -S. R., Sun, C. -T., and Mizutani, E. (1997). *Neuro-fuzzy and soft computing*, Prentice-Hall, New York, USA.
- Kosko, B. (1992). *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*, Prentice-Hall International, Englewood Cliffs, N.J., USA.
- Kurata, N., Spencer, B. F. Jr., and Ruiz-Sandoval, M. (2005). Risk monitoring of buildings with wireless sensor networks, *Structural Control and Health Monitoring*, **12**, pp. 315-327.
- Li, C., Huang, J.-Y., and Chen, C.-M. (2004). Soft computing approach to feature extraction, *Fuzzy sets and systems*, **147**, pp. 119-140.
- Lynch, J. P., Sundararajan, A., Law, K. H., Kiremidjian, A.S., and Carryer, E. (2004). Embedding damage detection algorithms in a wireless sensing unit for operational power efficiency, *Smart Materials and Structures*, **13**, pp. 800-810.

- Mallat, S. G. and Hwang, W.L. (1992). Singularity detection and processing with wavelets, *IEEE Trans. Information Theory*, 38, pp. 617-643.
- Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **11** (7), pp. 674-693.
- Montgomery, D.C. (2005). *Introduction to statistical quality control*, 5<sup>th</sup> edition, John Wiley & Sons, USA.
- Peng, Z., He, Y., Chen, Z. and Chu, F. (2002). Identification of the shaft orbit for rotating machines using wavelet modulus maxima, *Mechanical Systems and Signal Processing*, **16**, pp. 623-635.
- Robertson, A.N., Farrar, C.R. and Sohn, H. (2003). Singularity detection for structural health monitoring using holder exponents, *Mechanical Systems and Signal Processing*, **17**, pp. 1163-1184.
- Sohn, H. and Farrar, C. R. (2001). Damage diagnosis using time series analysis of vibration signals, *Smart Mater. Struct.*, **10**, pp. 446-451.
- Sohn, H., Czarnecki, J.A. and Farrar, R. (2000). Structural health monitoring using statistical process control, *Journal of Structural Engineering*, November, pp. 1356-1363.
- Strang, G. and Nguyen, T. (1997). *Wavelets and filter banks*, Wellesley-Cambridge Press, Wellesley MA, USA.
- Tanner, N.A., Wait, J.R., Farrar, C.R. and Sohn, H. (2003). Structural health monitoring using modular wireless sensors, *Journal of Intelligent Materials Systems and Structures*, 14, pp. 43-56.
- Zadeh, L.A. (1965). Fuzzy Sets, *Information and Control*, 8, pp. 338-353.
- Zadeh, L.A. (1973). Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Systems Man Cybernet.* **SMC-1**, pp. 28-44.
- Zadeh, L.A. (1977). A theory of approximate reasoning, Memorandum No. UCB/ERL M77/58, Electronics Research Laboratory, College of Engineering, University of California, Berkeley.
- Zadeh, L.A. (1977b). Theory of fuzzy sets, Memorandum No. UCB/ERL M77-1, Electronics Research Laboratory, College of Engineering, University of California, Berkeley.
- Zadeh, L.A. (1994). Soft computing and fuzzy logic, *IEEE Software*, Vol. 11, No. 6, pp. 48-56.