We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Using Wavelets for Feature Extraction and Self Organizing Maps for Fault Diagnosis of Nonlinear Dynamic Systems

Héctor Benítez-Pérez, Jorge L. Ortega-Arjona and
Alma Benítez-Pérez

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50235

1. Introduction

Fault diagnosis has been established in two main approaches: model-based fault diagnosis and model-free fault diagnosis. Present paper focuses on the later, mainly as an extension of the approach proposed in [17]. The challenge here is to classify faults at early stages, with an accurate response. However, as the term model-free implies, a model for the plant is not available neither for fault-free nor for fault-present scenarios. The objective, thus, is to classify faults based on system's response and the related signal analysis, in terms of dilation and shift decomposition, as obtained by a wavelets approach. So, self-organizing maps (SOM) are proposed as a powerful nonlinear neural network to achieve such a fault classification.

Several strategies have been proposed for feature extraction using wavelets. For instance, [1] presents a wavelet packet feature extraction, based on the analysis and measure of a "distance" between the energy distribution of some signal classes and the proper classification by the use of fuzzy sets. Alternatively, [2] proposes the use of wavelets as a strategy of parametric system identification, giving prime emphasis to wavelet properties and parameter relations. The idea of using wavelets for fault classification is a powerful procedure for feature extraction of several scenarios, even in the case of frequency and power shifts. [3] and [4] have explored this approach for process system, in which practical results are satisfactory, regardless of the classification. Moreover, several other strategies using wavelets have been proposed for abnormal signal detection, like that presented in [5], in which a parasitic wavelet transform is proposed. Further research in the same direction is followed in [6], in which a cubic spline methodology is proposed for the boundary

problem, although the results of this approach tend to be just local, linear models. An alternative strategy for auto-correlation and signal discovery is proposed in [7], following some multi-class wavelet support vector machines. In all these methodologies, wavelets are used as a technique for feature extraction; however, none of the above have presented any enhancement for pattern classification. Here, an enhancement for pattern classification is proposed, in order to isolate different scenarios.

On the other hand, the use of neural networks for feature extraction only presents the disadvantage of inherent data uncertainties and large quantity of necessary data. Different proposals have previously explored similar strategies. For example, [8] proposes feature extraction using local parametric models, giving valuable results; however, there is a drawback of a bounded system response. This strategy for fault diagnosis integrates an ART2A network and a Kohonen neural network. The objective is to combine both strategies in order to generate two subsystems capable of overcoming glitches and redundant data representations [8]. Both subsystems, based on the ART2A topology and the Kohonen neural network, are used to perform a learning strategy. This strategy allows on-line fault diagnosis, with the inherent uncertainty of SOM variation due to the plasticity-stability dilemma. A fundamental work has been introduced in [9], in which an extended review is provided regarding topics related to sensors patterns and stability-plasticity trade-offs, inherent to an ART2A network. Interesting comments have been included about how a time window data can be monitored, in order to identify abnormal situations, as well as how data should be treated in terms of normalization, time scaling, and filtering, and their comparison prior to declare a winner selection. Further developments are addressed in [10], focusing on the use of a parallel ART2A network approach, based on a wavelet decomposition in which clustering is defined in the wavelet domain, although it is not proposed for a dynamical system.

Feature extraction for dynamical systems based on wavelets presents the advantage of scale decomposition, allowing several possibilities of fault detection depending on the scale of the fault. Similarly, fault detection can be easily engaged if a source of information is decomposed into several fruitful components. These components are taken as parameter vectors, where several signal conditions are highlighted depending on the resolution. Further, these components need to be combined in a fair strategy, in order to classify similar behaviors. To do so, the use of SOM is proposed, in which each vector is processed as a consecutive input. The result of this classification gives a number of selected patterns, depending on the learning rate, and regarding a time window. Nevertheless, using this technique, the plasticity-stability dilemma is still not solved.

As stated before, feature extraction presents an inherently extrapolated method to determine several characteristics (like geometry differences) related to the monitored signal, based on a scale factor. The most typical characteristics are those related to frequency and phase modification, which are multiplicative faults in terms of fault detection. Amplitude change is detected on the general modification from wavelets scales, and the consequent change on the

selected patterns. The importance of the methodology is the capacity for fault isolation at unknown scenarios, with enough time to pursue modification in terms of system safety. In this sense, time response is determined in terms of the sampled window and the classification of current analyzed data. It is necessary to establish a feasible relationship among the sampling window, the time taken to process information, and the accuracy to classify a particular scenario. This can be achieved by following a frequency analysis of some selected scenarios, in order to find such a relationship. However, this is a non-homogeneous strategy for any scenario. Further work needs to be done in terms of data analysis and sampling capabilities to recognize scenarios, either known or unknown.

Based on this extensive review, the current approach divides the process of fault isolation using SOM techniques into a two stages process. The first stage is a basic construction of the map, as pattern clusters, using SOM. The second stage is a labeling process that identifies scenarios of the system. Following this, the current approach proposes the classification of time-varying faults within a bounded time window, using the wavelet decomposition inherent response. The objective, hence, is to establish an approach for fault localization, based on feature extraction and clustering, by considering diverse fault-present and fault-free scenarios. The novelty of this approach is on signal classification for time varying scenarios, under unknown consideration. The proposed system is limited to certain signals conditions, such as coupled noise and frequency response. In this case, frequency dispersion should be bounded, regardless of time variance. The main advantage of this approach is related to fault isolation through signal decomposition, and classification in a bounded time response.

2. Background

2.1. The Wavelet Transform

Wavelet transform (WT) is an alternative method for processing transient, non-stationary signals simultaneously in time-scale domains [11]. Wavelets are used to decompose a signal into different *scale factors*. Wavelet approach provides a more natural description of the signal, in terms of a composition of a set of "typical signals", or wavelets. In fact, the WT is the correlation between a signal and a set of basic wavelets, proposed from a basic "mother wavelet" $h(t)$, chosen in order to analyze a specific transient signal of finite energy. Thus, a complete orthogonal set of "daughter wavelets" $h_{a,b}(t)$ is generated from $h(t)$ by two operations: a dilation a and shift b.

Formally, the dilation and shift, as wavelet coefficients of the signal, are defined by:

$$W_S(a, b) = \int_{-\infty}^{\infty} s(t) h_{a,b}^*(t) dt \qquad (1)$$

where $s(t)$ is the current signal, and the function $h_{a,b}(t)$ defined as:

$$h_{a,b}(t) = a^{-1/2} h\left(\frac{t-b}{a}\right) \tag{2}$$

The information used in this approach is based on both a and b, as expansion coefficients. Commonly, the mother wavelet is considered as a Daubechies signal. The wavelets automatically adapt to the different components of a signal, using a small window (large scale) to search for brief high-frequency components, and large window (low scale) to look for long lived, low-frequency components. The shape of low- and high-frequency components is determined by the mother wavelet. A further and deeper revision of the wavelet technique may be found in [12].

Wavelets can be represented using a function ψ, and the family $\Im$ of expanded and translated wavelets are expressed as :

$$\Im = \left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j}\right) : j,\, n \in \mathbb{Z} \right\} \tag{3}$$

which performs an orthogonal base $L^2 \not\subset$. Orthonormal wavelets are obtained by expanding this by a factor 2^j, allowing variations of the signal in a 2^{-j} resolution. The construction of these bases permits the study of multi-resolution of a signal. Formally, the approximation of a function with a resolution 2^{-j} is defined as an orthogonal projection over a space $V_j \subset L^2 \not\subset$

The space V_j groups every possible approximation with resolution 2^{-j}. Remember that an orthogonal projection from function f is a function $f_j \in V_j$ which is minimized $\| f - f_j \|$.

Definition. (Multi-Resolution) A family of closed subspaces $\{V_j : j \in \mathbb{Z}\}$ of $L^2 \not\subset$ is a multi-resolution approximation, if it satisfies the following properties:

i.   *For $j,\, k \in \mathbb{Z},\, f(t) \in V_j$ ▪ $f(t - 2^j k) \in V_j$*

ii.   *For $j \in \mathbb{Z}, V_{j+1} \subset V_j$*

iii.   *For $j \in \mathbb{Z},\, f(t) \in V_j$ ▪ $f\left(\frac{t}{2}\right) \in V_{j+1}$*

iv. $\lim\limits_{j \to +\infty} V_j = \bigcap\limits_{j=-\infty}^{+\infty} V_j = \{0\}$

v. $\lim\limits_{j \to -\infty} V_j = \bigcup\limits_{j=-\infty}^{+\infty} V_j = L^2 \not\subset$

vi.   *there is θ such that $\{\theta(t - n) : n \in \mathbb{Z}\}$ is a Riesz base of V_0 .*

Property (i) states that the subspace V_j is invariant in any translation proportional to scale 2^j Property (ii) is causal, since the resolution 2^{-j} owns the necessary information to calculate a

raw resolution 2^{-j-1}. If the functions are dilated in V_j by 2, then the details are amplified by a factor of 2. So (iii) defines an approximation of a raw resolution when the resolution 2^{-j} tends to be cero. (iv) implies that all the details have been lost, meaning that the projection of signal f over the space V_j when $j \rightarrow +\infty$ is zero:

$$\lim_{j \rightarrow +\infty} \| P_{V_j} f \| = 0.$$ (4)

On the other hand, if the resolution 2^{-j} tends to $+\infty$ property (v) forces that the approximation of the signal converges to the original signal:

$$\lim_{j \rightarrow -\infty} \| f - P_{V_j} f \| = 0.$$ (5)

Finally, the existence of a Riesz base $\{\theta(t-n) : n \in \mathbb{Z}\}$ of V_0 provides a discretization theorem. Function θ can be interpreted as a cell with unitary resolution. To compute several resolutions of signal, it is necessary then to compute the orthogonal components over different spaces $\{V_j : j \in \mathbb{Z}\}$ of $L^2 \not\subset$ According to the definition of a Riesz base, there are $A, B > 0$ that if $f \in V_0$, it may be decomposed as:

$$f(t) = \sum_{n=-\infty}^{+\infty} a[n] \theta(t-n)$$ (6)

with

$$A \| f \|^2 \leq \sum_{n=-\infty}^{+\infty} | a[n] |^2 \leq B \| f \|^2$$ (7)

This last expression guarantees that the expansion of the signal over $\{\theta(t-n) : n \in \mathbb{Z}\}$ is numerically stable.

The approximation of f in the resolution 2^{-j} is defined as the orthogonal projection $P_{V_j} f$ over V_j. To compute this projection, an orthogonal base should be find over the space V_j. Hence, the following theorem allows for an orthogonal Riesz base $\{\theta(t-n) : n \in \mathbb{Z}\}$, and it builds an orthogonal basis for each space V_j where dilation and transferring over a function ϕ, named *scale*.

Theorem (Goswami Jaideva, 1999). Being $\{V_j : j \in \mathbb{Z}\}$ an approximation of multi-resolution and ϕ a scale function, where the Fourier Transformation is:

$$\hat{\varphi}(\omega) = \frac{\hat{\theta}(\omega)}{\left(\sum_{k=-\infty}^{+\infty} | \hat{\theta}(\omega + 2k\pi) |^2 \right)^{1/2}}$$ (8)

Defining

$$\varphi_{j,n}(t)=\frac{2}{\sqrt{2^{j}}}\varphi\left(\frac{t-n}{2^{j}}\right) \tag{9}$$

Then the family $\{\theta_{j,n}:n\in\mathbb{Z}\}$ is an orthonormal base of $V_j \ \forall \ j\in\mathbb{Z}$

Demonstration

The objective is to build an orthonormal base, and therefore, a function $\varphi\in V_0$. Now, this function is expanded in terms of the Reisz base $\{\theta(t-n):n\in\mathbb{Z}\}$:

$$\varphi(t)=\sum_{k=-\infty}^{+\infty}a[n]\theta(t-n) \tag{10}$$

Computing the Fourier transform for this last function, it is obtained that:

$$\hat{\varphi}(\omega)=\hat{a}(\omega)\hat{\theta}(\omega) \tag{11}$$

where $\hat{a}$ is a Fourier series with period 2π and finite energy; $\hat{a}$ is expressed in terms of the orthogonal condition of $\{\varphi(t-n):n\in\mathbb{Z}\}$ into the Fourier dominion, being $\bar{\varphi}(t)=\varphi^{*}(-t)$ It is necessary that for any $n,\ p\in\mathbb{Z}$:

$$\langle\varphi(t-n),\ \varphi(t-p)\rangle=\int_{-\infty}^{+\infty}\varphi(t-n)\varphi^{*}(t-p)dt=\varphi*\bar{\varphi}(p-n)=\delta[n] \tag{12}$$

If the Fourier transform is obtained from this, it determines the following equation:

$$\sum_{k=-\infty}^{+\infty}\mid\hat{\varphi}(\omega+2k\pi)\mid^{2}=1 \tag{13}$$

On the other hand, the Fourier transform of $\varphi(t)*\bar{\varphi}(t)$ is $\mid\hat{\varphi}(\omega)\mid^{2}$, and therefore:

$$\hat{a}(\omega)=\frac{1}{\left(\sum_{k=-\infty}^{+\infty}\mid\hat{\theta}(\omega+2k\pi)\mid^{2}\right)^{\frac{1}{2}}} \tag{14}$$

To find an approximation of *f* over the space V_j, it is necessary to expand the function in terms of the orthogonal base from the scale function:

$$P_{V_j}f=\sum_{n=-\infty}^{+\infty}\langle f,\ \varphi_{j,n}\rangle\varphi_{j,n} \tag{15}$$

where the inner products are:

$$a_j[n] = \langle f, \varphi_{j,n} \rangle \tag{16}$$

giving a discrete approximation of scale 2^j. This last equation can be expressed as the convolution product:

$$a[n] = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \varphi\left(\frac{t - 2^j n}{2^j}\right) dt = f * \bar{\varphi}_j(2^j n) \tag{17}$$

with $\bar{\varphi}_j(t) = \sqrt{2^{-j}} \varphi(2^{-j} t)$. The energy of the Fourier Transform $\hat{\varphi}$ is typically concentrated in $[-\pi, \pi]$, and as a consequence, the Fourier Transform $\sqrt{2^j} \hat{\varphi}^*(2^j \omega)$ of $\bar{\varphi}_j(t)$ concentrates its energy in the interval $\lfloor -2^{-j}\pi, 2^{-j}\pi \rfloor$. Then, the discrete approximation is $a_j[n]$, as a low pass filter of function f, from sampling 2^j.

2.2. Self Organizing Maps (SOM)

The purpose of Kohonen's SOM is to capture the topology and probability distribution of some input data (Figure 1) [13][14]. First, a topology of SOM is defined as a rectangular grid [15].

Different types of grid may be used to represent data, although the one shown in Figure 2 presents a homogenous response suitable for noise cancellation.

The neighborhood function regarding a rectangular grid, such as this one, is based on a set of bi-dimensional Gaussian functions, as described by Equation 18.

$$h(i_1, i_2) = \exp\left(-0.5 * \frac{(i_1^{win} - i_1)^2 + (i_2^{win} - i_2)^2}{\sigma^2}\right) \tag{18}$$

where i_1 and i_2 represent the indices of each neuron, and σ is the standard deviation of each Gaussian distribution, which determines how the neighbor neurons of a winner neuron are modified. Each neuron also has a weight vector (w_i^j), which represents how the actual neuron is modified by an input updating. Thus, $h(i_1, i_2)$ is the Gaussian representation that allows for modifications of neighbor neurons of a SOM.
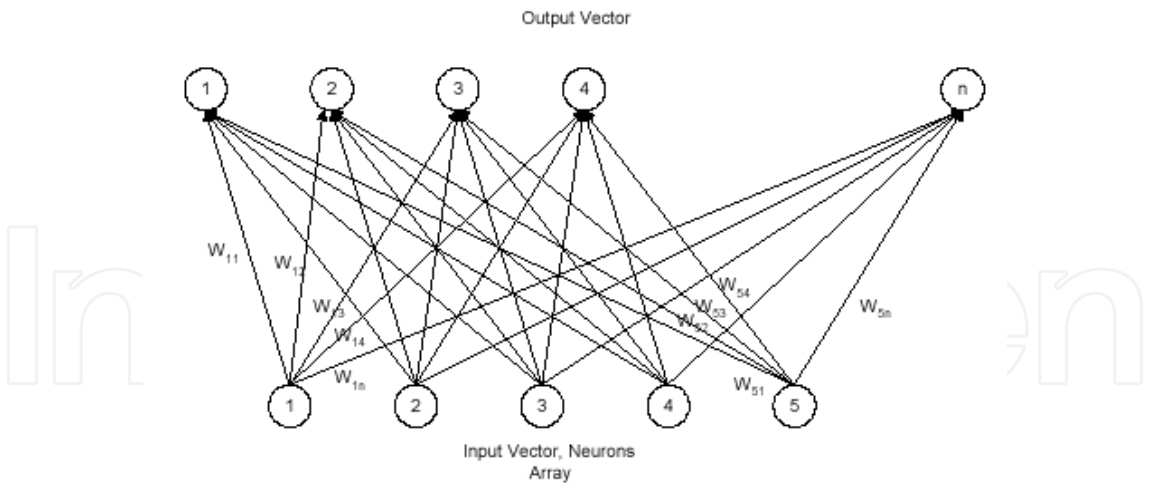
Figure 1. Topology Network of a SOM.

Equation 20 is the basis of the SOM. In this equation, the weight matrix is updated based on the bi-dimensional indexing, namely $h(i_1, i_2)$. This equation is used during training (offline) stage of the SOM. Moreover, this bi-dimensional function allows the weight matrix to be updated in a global way, rather than just to update the weight vector associated to a winner neuron.
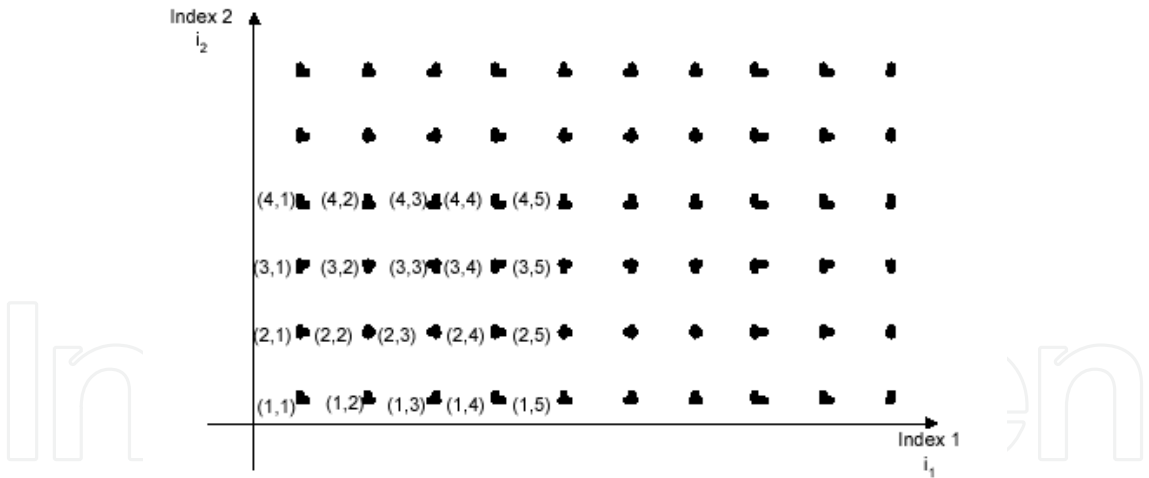

Figure 2. Index Grid for noise cancellation.

For updating the SOM, an inner product is performed between the weight matrix *W* and the input vector (*I*), in order to define a winner neuron. Having calculated this product, the maximum value is determined by the comparison between each scalar from resultant vector. This value is declared as the winner, just as in the technique known as "the winner take all". The related bi-dimensional index (Figure 2) is calculated in order to determine how the weight matrix is modified.

The updating process of the weight matrix is performed as shown by Equation 19.

$$w_j^{new} = w_j^{old} + \eta * h\left(i_1,\ i_2\right) * \left(I - w_j^{old}\right)$$ (19)

where η represents a constant value equals to 0.7. This parameter can be tuned as a learning parameter. Here, I represents the current input vector.

3. Main approach

The proposed organization for the actual model is shown in Figure 3. The model is divided into two stages: first, it is performed offline, where the SOM is trained; second, the fault diagnosis procedure is performed online.
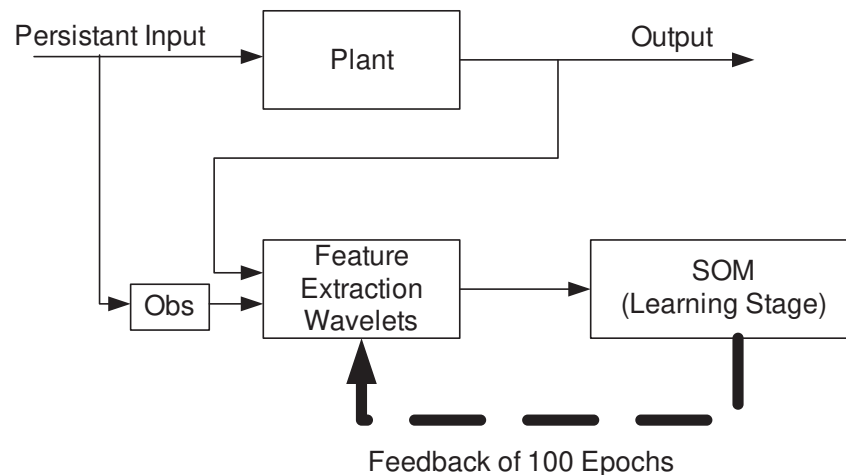
Figure 3. Organization for the model.

The SOM is trained offline, following the output signal, which is decomposed into several levels, as proposed by wavelets feature extraction scales. The observed dynamical system decomposition should perform this regarding its input, by considering a diversity of frequencies. The decomposition is proposed for several wavelets scales as feature extraction, where just one is chosen based on repeatability in similar scenarios, either fault-present or fault-free.

Several parameters need to be tuned for the SOM and wavelet feature extraction, such as:

• Length of the sampling window k.

• Number of wavelet decomposition levels.

• Learning value η.

• Vigilance threshold β.

During the offline stage, the SOM is trained based on known fault-present and fault-free scenarios with a local decomposition wavelet strategy, using a mother wavelet known as Daubechies 4. This means that four decomposition levels (scales) are produced. The Daubechies 4 is a wavelet particularly chosen here due to the case study response, as it is discussed later.

For fault-free scenarios, four decomposed levels have a particular powerful response, being different of the fault-present scenarios, as shown in the next section. In this case, similar patterns are stated to occur when a bounded fault scenario is presented. In both, fault-free and fault-present cases, and for four levels, Daubechies 4 presents a trustable response.

Current data has a pre-treatment as input/output responses from the plant. Input and output are locally normalized before they are processed by wavelets, to extract certain features. To take the advantage of this situation, it is necessary that the learning law of the SOM is suitable to perform an accurate response with respect to the case study.
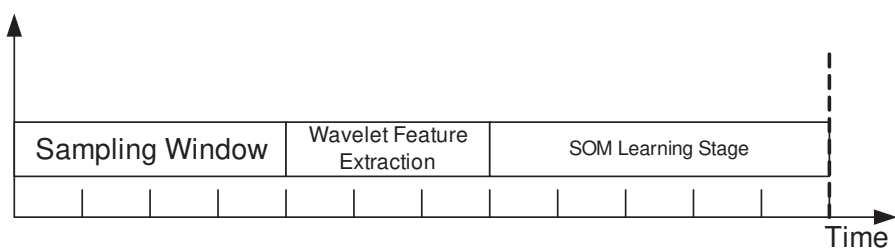
Figure 4. Three steps of the offline stage.
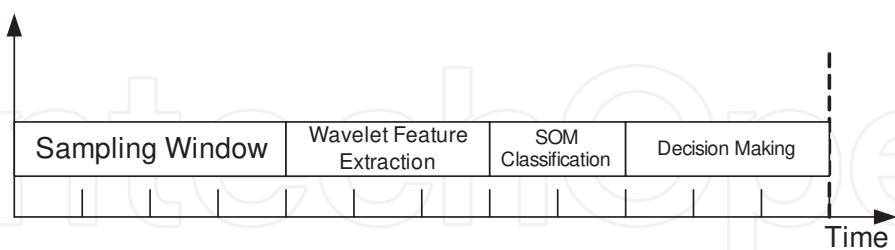
Figure 5. Three steps of the online stage.

Three steps are defined during the offline stage, as shown in Figure 4, in terms of data processing: first, a sampling window of input and output data is taken and normalized; next, this information is processed by the wavelet module, to perform feature extraction; finally, the local matrix is classified by the SOM in learning mode. In this case, it is necessary to provide with enough information and a diversity of scenarios (fault-present and fault-free) to the SOM, in order to ensure a suitable fault identification.
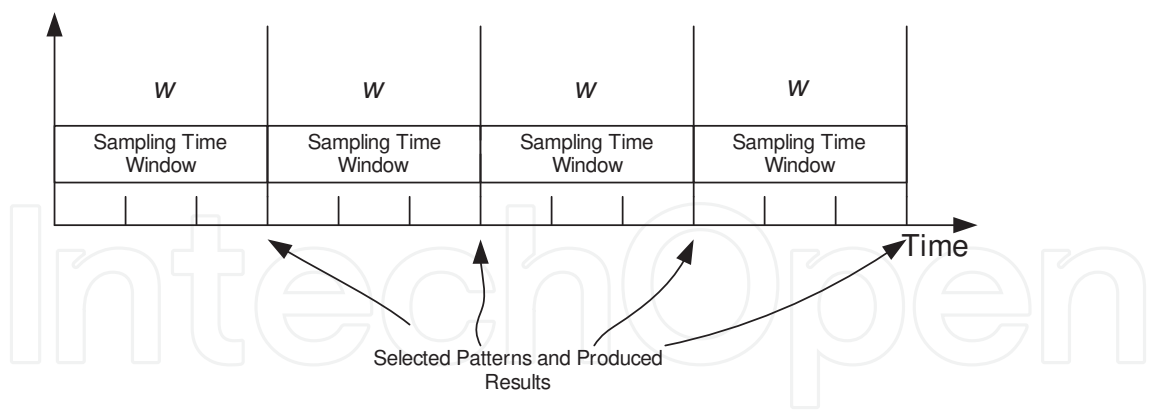
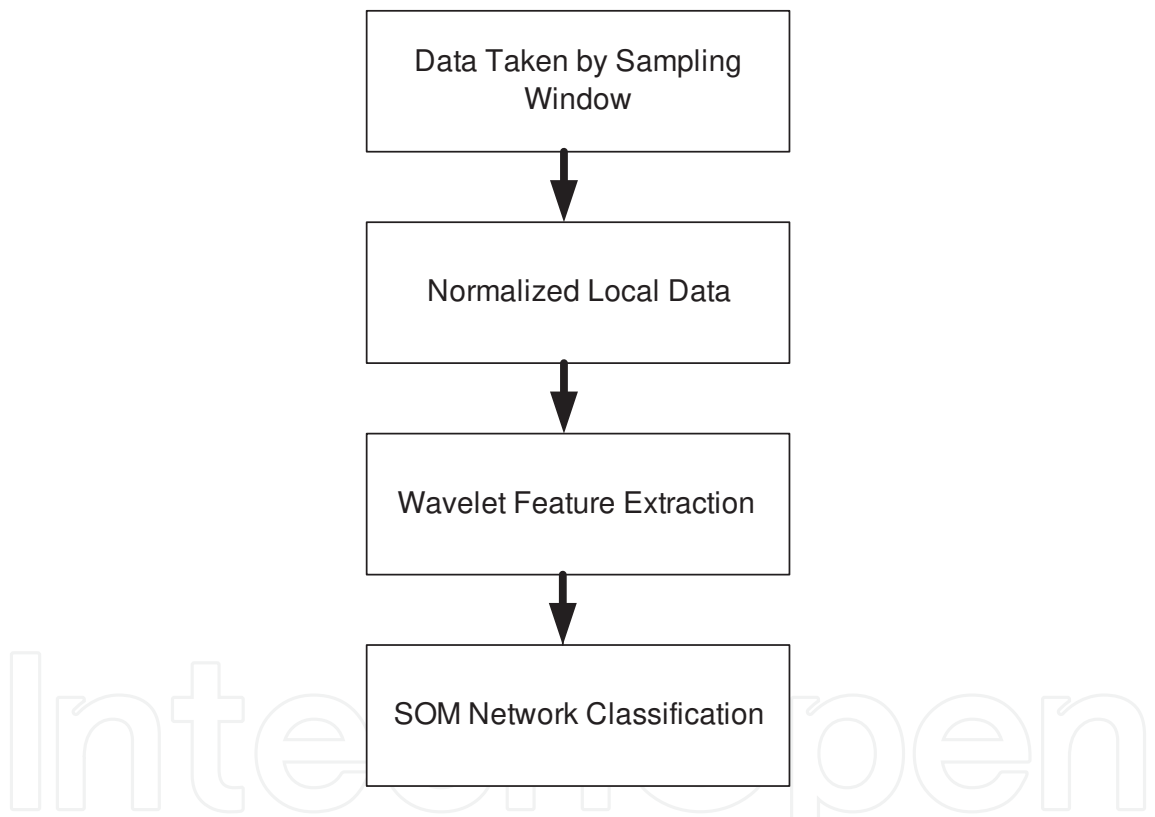**Figure 6.** Results produced during the online stage.

Figure 7. Proposed main procedure for the model-free fault diagnosis.

During the online stage, the SOM does not perform any learning procedure. The comparison between already classified patterns against current scales vectors, produced by wavelets, is performed by an inner product, where the minimal value obtained amongst all patterns is defined as the winner, and compared with β. If the winner value is smaller than β, this is a correct winner; otherwise, the SOM is not capable of performing this comparison. The value β is called *error*, for the purposes of model evaluation. Similarly to the offline stage, the online stage takes four major steps, as shown in Figure 5, in order to produce a result. The main difference, thus, is that here SOM does evaluate the resultant

local model, generated by wavelet feature extraction. The decision making module inherent in SOM determines if this classification is valid or not. Notice that for the current purposes, time consumption is neglected.

The results of the online stage are produced every w sampling time windows, which is the time taken by the application in order to produce a result (Figure 6). Such a time window w becomes crucial in terms of the frequency of the plant. Further, the sampling window is a multiple of inherent period of the system. In fact, the plant response during fault-present and fault-free scenarios need to be bounded to this parameter, in order to guarantee a reliable response.

The main procedure followed here is shown in Figure 7, in which the normalized, local preprocessing stage is executed per local sample data; after that, the wavelet feature extraction is performed as data decomposition. The results related to different levels are processed by the SOM per level, where just one would be the winner, and learned by the SOM.

4. The case study

The case study is integrated by a system, as it is shown in Figure 8. It consists of a multiple-input, single-output (MISO) system, with a PID controller, and a switching fault injection procedure. As it may be noticed, the dynamics of this plant tends to be quite slow in comparison with the occurrence of faults, according to the dynamics of the plant and the fault scenarios. This characteristic is crucial for the construction of the local model and the feature extraction, in order to produce a fruitful fault diagnosis procedure. Case study is linear and modeled through model-based techniques; however, when a fault is present, its dynamics becomes nonlinear. This nonlinear behavior tends to be extremely difficult to be diagnosed by classical strategies, like for example, unknown input observers [8]. This scenario has been presented for local system identification techniques and for a global classification strategy, in order to have an accurate fault diagnosis strategy [16]. However, the strategy is dependent on the persistent excitation of certain frequency responses. Alternatively, the proposed strategy here overcomes such a frequency dependency, since the only obvious dependency is the sampling period. This strategy is based on the sensibility of the feature extraction strategy, which uses Daubechies 4 wavelets (db4).

The dynamics of the case study are expressed as vectors of the state space representation:

k1= [0.1; -20.000];

e=[0; 0.001];

b=[1.8000 -2.1000; 0.9000 0.8600];

A=[1.1000 0; 0 2.1000];

c=[0 1];

where A, b, c, and e belong to the case study model, and k1 to the PID controller. The types of faults are:

Fault 1: Backlash of 0.01, followed by a dead zone between {-0.1 0.1}

Fault 2: Backlash of 0.91, followed by a dead zone between {-0.3 0.4}

For this case study, the parameters to be tuned (such as β, η, a, b and k) are bounded to specific ranges:

$0.01 \beta 0.99$
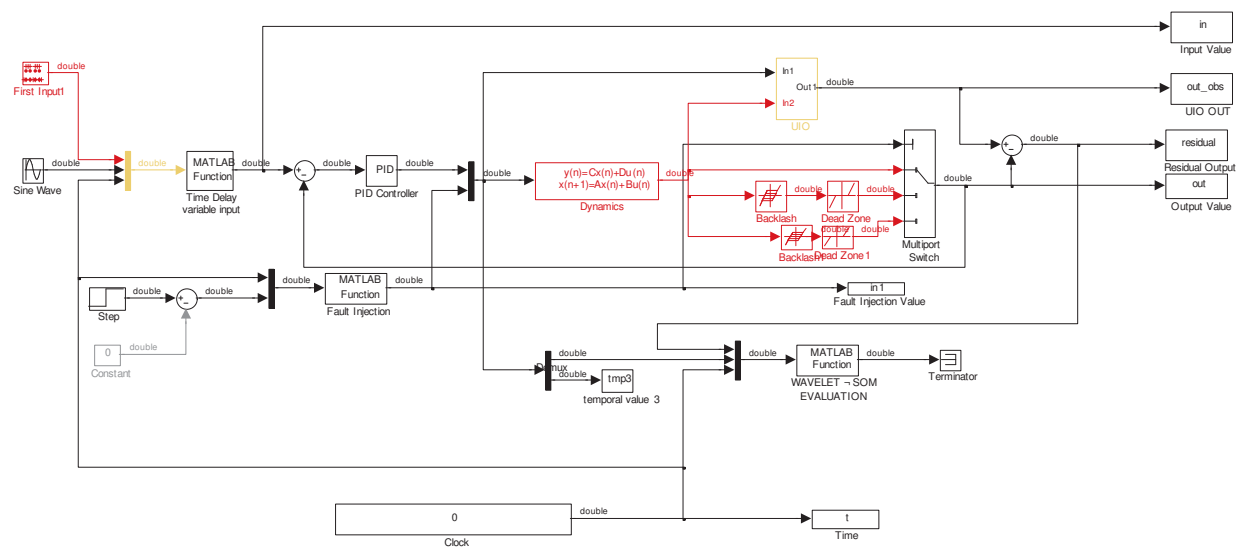
$0.01 \eta 0.99$ where η =0.1952 for this case study

$2a, b6$

$k = 10$

Figure 8. Schematic diagram of the system for the case study.

These ranges are arbitrarily fixed, since there is no further information regarding these values. In order to obtain certain selection according to fault presence and dynamic system response, a testing (response) from the tuples integrated by the combination of these values is performed, as shown in the next section. Notice that the pattern selected by the SOM is a representation of the most suitable approximation of current feature extraction, that is, a model of the current response.

5. Case study results

The present results are referred in terms of pattern construction and feature extraction, considering several known fault-present and fault-free scenarios. Regarding this, Figure 9 shows the output response of the benchmark during a fault-free scenario. The response of the system is inherently stable, due to the inner local control. In this case, the sampling time window has a duration of 10 seconds, representing a period of system response. This re-

sponse produces the level decomposition presented in Figure 10, in which four wavelet levels are used to decompose the data. Since response is fairly stable, the main differences amongst levels are not significant neither regarding to power, nor regarding to frequency selection. Based on this approximation, the selected patterns should be similar, equidistant, and close related; otherwise, an unknown response has appeared.
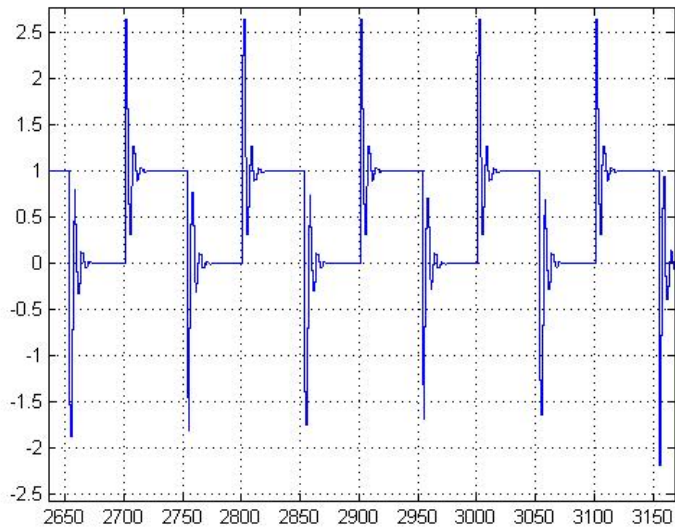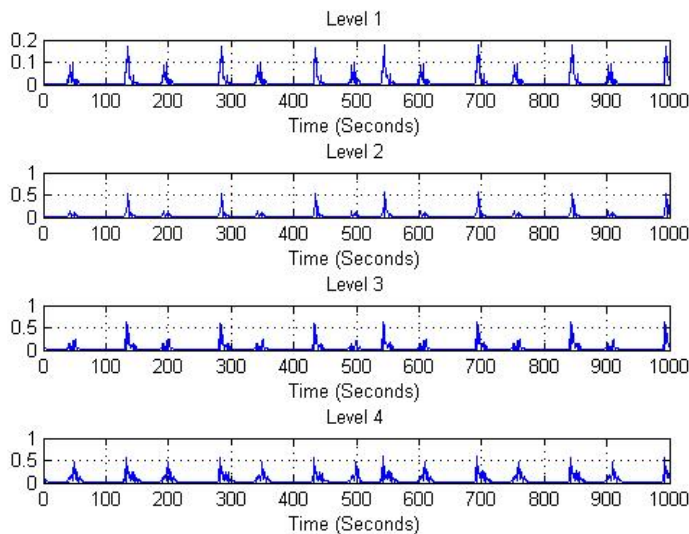

Figure 9. Fault-free response for the current example.

Figure 10. Wavelets feature extraction for fault-free scenario.
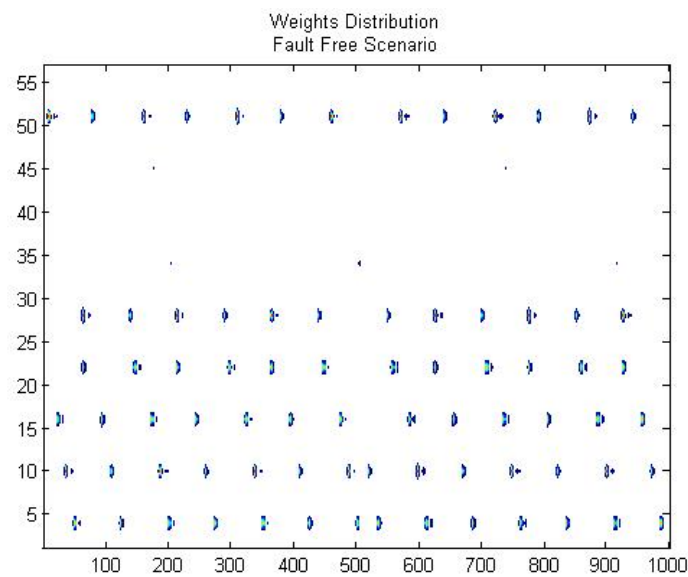
Figure 11. SOM for the fault-free scenario.
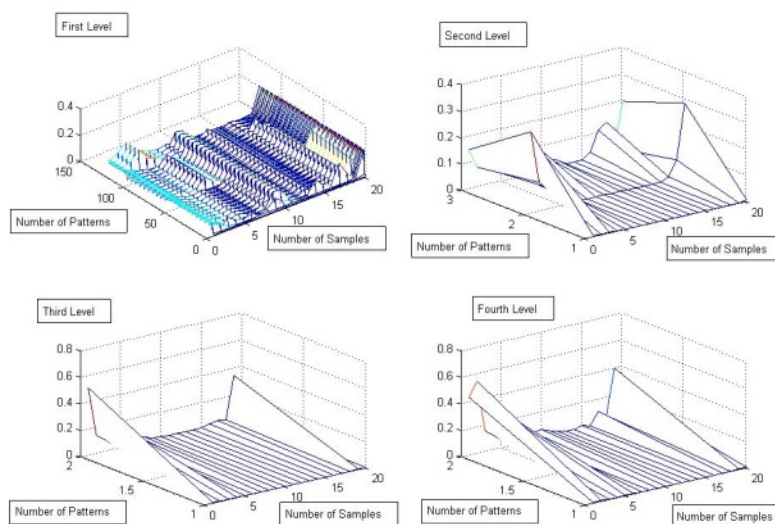
Figure 12. Selected patterns per level for the fault-free scenario. The winner patterns are at the first level.

The fault-free scenario shows the response of a damped system, in which four decomposed levels tends to be similar in terms of power values. For this scenario, the patterns are classified by the SOM as shown in Figure 11. The total number of patterns is 125, from which 55 patterns are "zoom" (enlarged) patterns with a similar power response, given the wavelets feature extraction. Remember that Daubechies 4 has been selected for feature extraction, since this wavelet is stable enough in terms of feature extraction repetition for same scenarios. In Figure 11, the amplitude of 55 patterns reflects a similar response amongst them. Moreover, there is a region around patterns 30 to 45, whose amplitude is close to zero. No-

tice that the pattern response tends to be similar through time. In Figure 12, the related patterns per level are classified by the SOM for this fault-free scenario. In this case, level 1 presents most of the variations with respect to the observer scenario. This is reflected by the number of created patterns regarding the rest of the levels: level 1 proposes 125 patterns, while the rest of the levels propose mostly 2.

Figure 13 presents the response of the system and the selected patterns that vary from 0.1 to 0.8, in magnitude terms, presenting a close relation between the winner patterns. It is interesting to observe that the most selected pattern has a magnitude of 0.6.
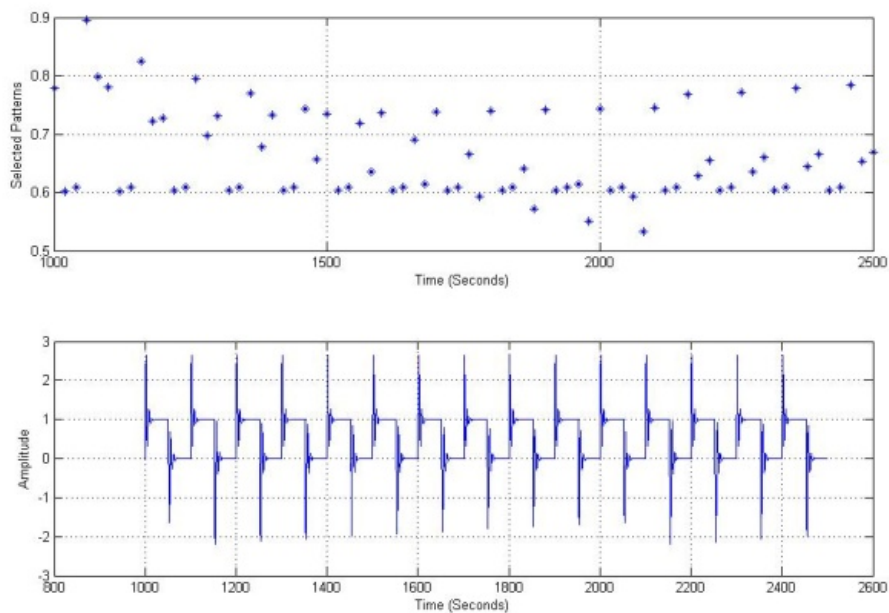
Figure 13. Fault-free scenario and selected patterns for the case study.

For the fault-present scenario, two types of faults are injected to the system (see Section 4). First, let us consider Fault 1, which is an increment of sudden amplitude, as shown in Figure 14. In this case, the fault scenario is presented from 400 to 800 seconds, and it is related to a 10% increment of the amplitude. A detail of the response for this scenario is shown in Figure 15. Observe that the system response presents just an oscillation, and not a clear stage of fault condition. However, the expected response is not desirable according to system dynamics.

Figure 16 shows the features extraction results for this scenario. In this case, power values are modified around 500 seconds, in comparison with the fault-free response (Figure 10). This difference is a clear modification of phase around the 500 seconds. This slight phase modification is the result of the increment of amplitude for this fault scenario. Regarding the patterns that are classified by the SOM for this fault scenario, the number of patterns simply does not augment, but the feature extraction results are different. This is shown in the amplitude of the patterns and the selected patterns, especially in pattern number 15,

which clearly exposes the presence of the fault. In Figure 16, the first level presents a very noticeable difference between the two responses of the whole scale presentation. Moreover, the response in terms of patterns classified by the SOM tends to increase, but it is still stable (Figure 17).
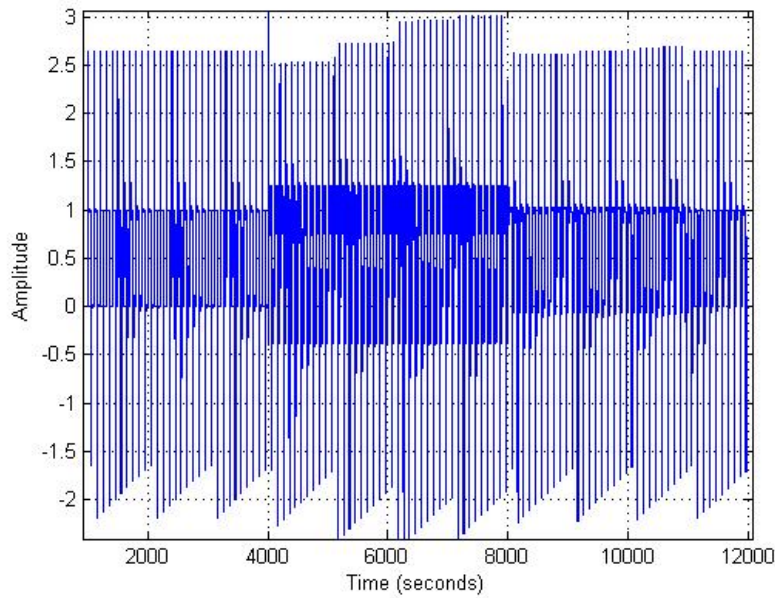

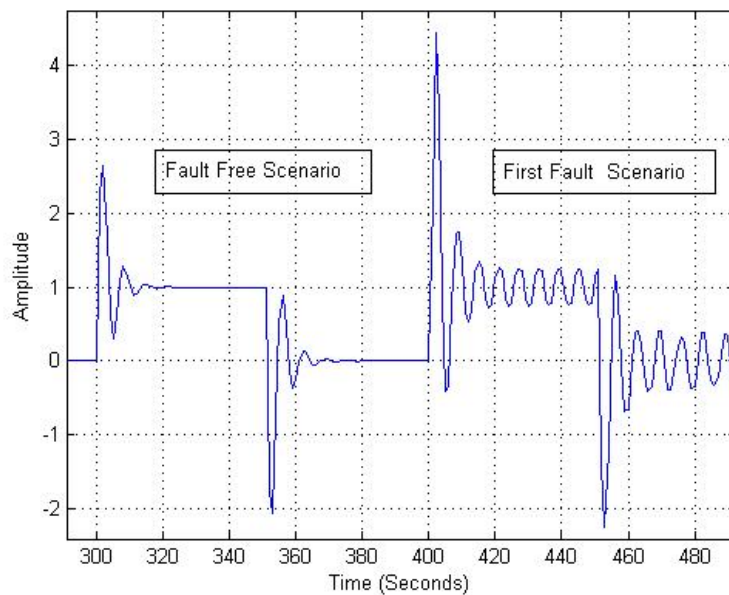**Figure 14.** System response during fault present scenario.

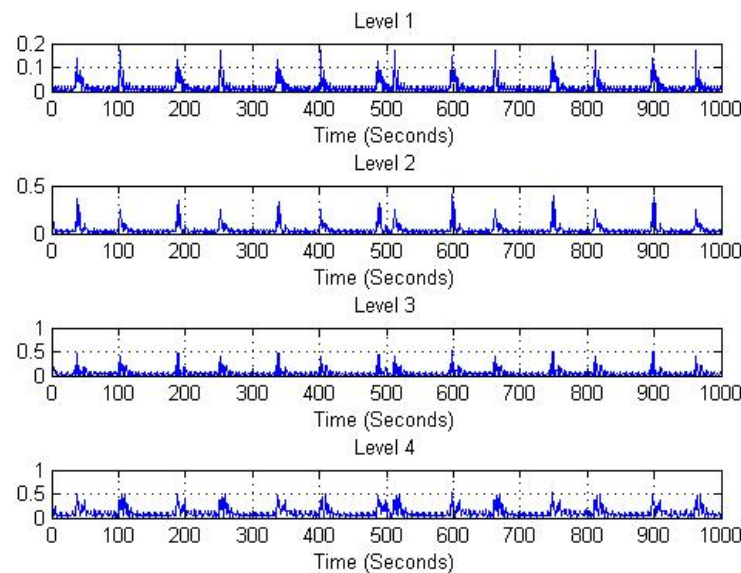**Figure 15.** Detail of the response for the first fault present scenario.

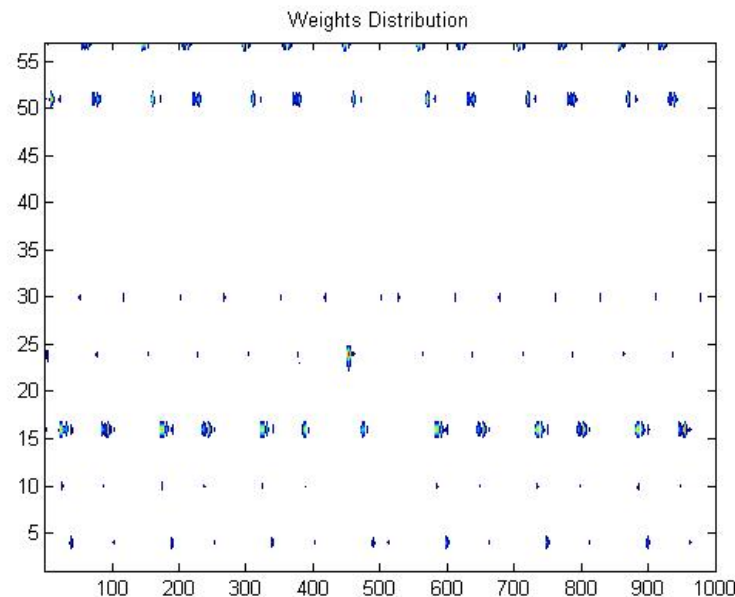**Figure 16.** Wavelets feature extraction for the fault present scenario.

Figure 17. SOM for the fault present scenario.

For the case second fault scenario, Fault 2, Figure 18 shows a detail of the output response of the system, in which an oscillation tends to be larger than that present in the first fault scenario. Although, this behavior is quite erratic, the case study is still measurable, as shown in Figure 19. In this figure, the patterns are selected and quite defined in terms of the presence

of the fault in fast response. There is a small magnitude variation around patterns 2 to 3, at the beginning of the fault scenario, which is resolved at the fifth sampling window. The values of the selected patterns in terms of magnitude are quite defined around 3, which is clearly different in comparison with the other two previous scenarios.
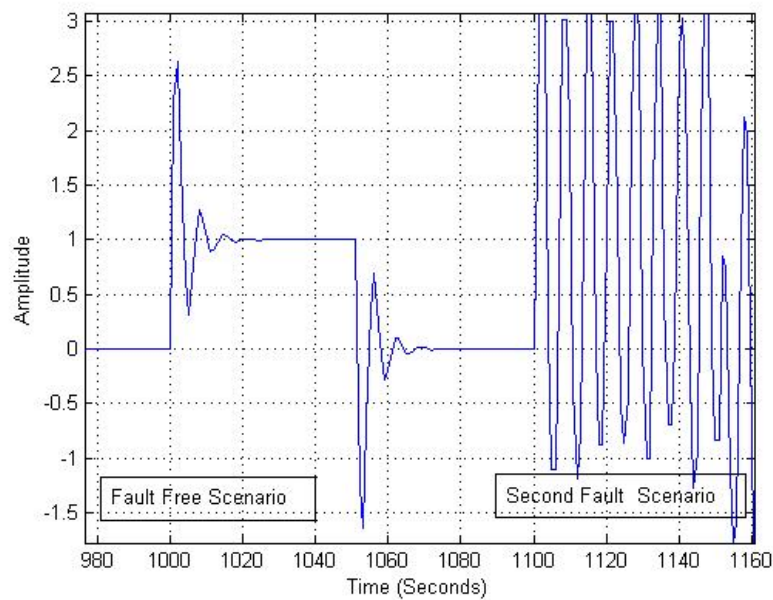

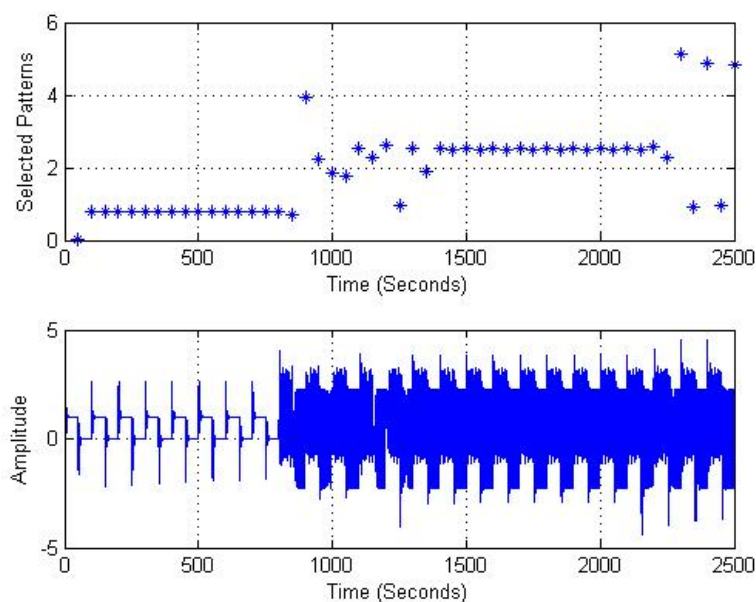**Figure 18.** Detail of the response for the second fault present scenario.

Figure 19. General response for the second fault present scenario and the selected pattern.

Figure 20 shows the selected patterns per level, which present an increment in the number of patterns. Now, the first level has 200 patterns, while the rest of three levels have around 80 patterns. Moreover, the amplitude of learned and winner patterns are around 0.15 for the first level, and close to zero for the second, third, and fourth levels, in the majority of learned patterns (red patterns).
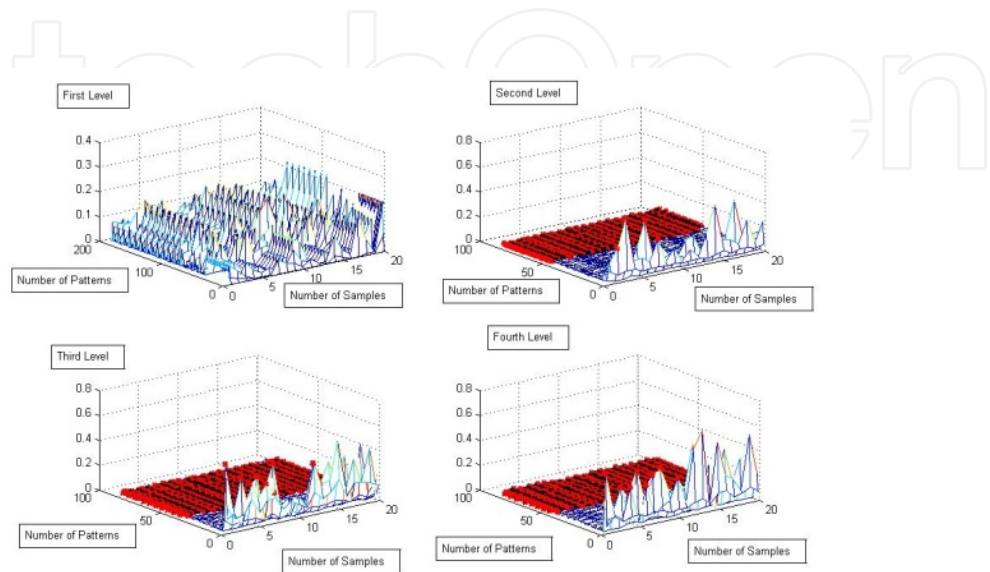

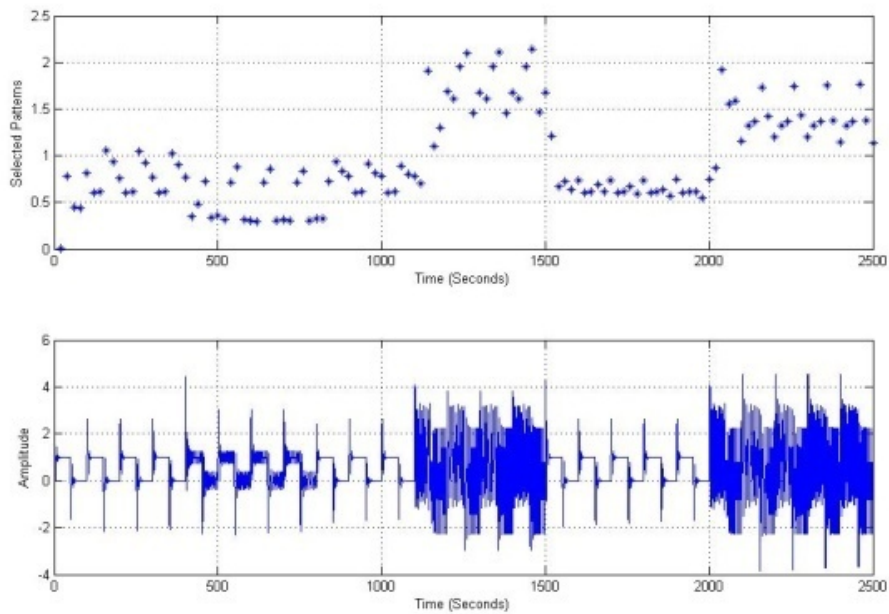**Figure 20.** Selected patterns per level for the second fault present scenario.

Figure 21. Current output response for the fault-present and fault-free scenarios.

Figure 21 shows a combination of the three scenarios: the fault-free scenario from time 0 to 400 seconds, the first fault scenario from 400 to 700 seconds, and the second fault scenario from 1100 to 1500 seconds, and again from 2000 to 2500 seconds. From this combination, several patterns have been selected, as shown by the upper diagram of Figure 21. These selected patterns, depicted with their amplitude, show a clear definition of the three fault scenarios: the fault-free scenario is related to the patterns measured at 0.7, the first fault-present scenario is related with the two patterns measured at 0.4 and 0.9, respectively, and the second fault-present scenario is related with the three patterns whose values are 1.4, 1.6 and 2.0. The richness is quite obvious in this scenario, due to the oscillation present in it.
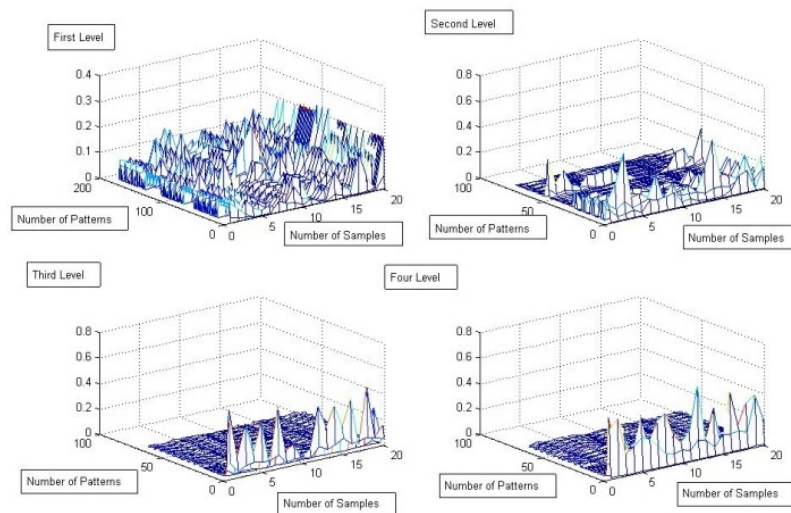

Figure 22. Selected patterns per wavelet levels for the three scenarios.

Finally, Figure 22 shows the levels of the four wavelets, depicted as learned patterns, in which the first level has 190 winner patterns, while the other three levels have 60 winner patterns. Observe that the richness of the combined winner patterns determines a clear classification of the patterns, as shown in Figure 21. Moreover, the current amplitude is quite similar and rich for the first level, as presented in the three scenarios.

6. Concluding remarks

The present work shows a strategy for fault diagnosis based on local feature extraction and global system classification, in which different parameters play an important role in this condition monitoring proposal. The number of samples with respect to sampling window, the learning and vigilance parameters, and the number of extracted features has been tuned by an extensive review of every possible scenario. This approach enhances the capabilities of the simple use of a neural network for pattern classification. This strategy has shown an alternative for classification of abnormal situations, with no information from current case

study as state space response. Furthermore, pattern databases have been constructed, based on several selected scenarios, which have been obtained offline per scenario. This initial information is basic in order to obtain an accurate model of the system under study. This strategy shows how model-free techniques can be implemented for fault diagnosis, and the need of a large amount of data, as well as the extensive review of multiple parameters. The main contribution here is the use of several pre-processing data stages, in order to conform suitable and accurate information to be processed by the neural network. In addition, a database of several scenarios needs to be used for a trustable fault diagnosis strategy. In fact, it is necessary to define a heuristic index that allows the differentiation between scenarios. Here, the use of one specific time frequency distribution approach over the rest of the current algorithms is pursued. Moreover, this strategy could address a proper dynamic non-linear system for online classification of unknown scenarios.

Acknowledgements

Author details

Héctor Benítez-Pérez[1*], Jorge L. Ortega-Arjona[2] and Alma Benítez-Pérez[3]

*Address all correspondence to: hector@uxdea4.iimas.unam.mx

1 Departamento de Ingeniería de Sistemas Computacionales y Automatización, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, México

2 Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, México

3 CECYT 11 "WILFRIDO MASSIEU", Instituto Politécnico Nacional, México

References

[1] Li, D., Pedrycz, W., & Pizzu, N. (2005). Fuzzy Wavelet Packet based Feature Extraction Method and Its Application to Biomedical Signal Classification. *IEEE Transaction on Biomedical Engineering*, 52(6), 1132-1139.

[2] Billings, S., & Wei, H. (2005). A New Class of Wavelet Networks for Nonlinear System Identification. *IEEE Transaction on Neural Networks*, 14(4), 862-874.

[3] Wang, X., Chen, B., Yang, S., & Mc Greavy, C. (1999a). Application of Wavelets and Neural Networks to Diagnostic Systems Development, 2, an Integrated Framework and its Application. *Computers and Chemical Engineering*, 23, 945-954.

[4] Chen, B. H., Wang, X. Z., Yang, S. H., & Mc Greavy, C. (1999). Application of wavelets and neural networks to diagnostic system development, 1, feature extraction. Computers & Chemical Engineering , July 1999, 23(7), 899-906.

[5] Zhan, Z., Ikeuchi, H., Saiki, N., Imamura, T., Miyake, T., Toda, H., & Horihata, S. (2008). Fast Wavelet Instantaneous Correction and its Application to Abnormal Signal Detection. *International Journal of Innovative Computing Information and Control*, 4(10), 2697-2710.

[6] Cheng, Y., Junxian, L., Feng, B., & Guan, W. (2009). Hermite Cubic Spline Multi-Wavelet Natural Boundary Element Method. *ICIC Express Letters*, 3(2), 213-217.

[7] Xing, Y., Wu, X., & Zhiliang, X. (2008). Multiclass Least Squares Auto-Correlation Wavelet Support Vector Machines. *ICIC Express Letters*, 2(4), 345-350.

[8] Benítez-Pérez, H., & Benitez-Perez, Alma. (2009). The use of ARMAX strategy and Self Organizing Maps for Feature Extraction and Classification for Fault Diagnosis. *International Journal of Innovative Computing, Information and Control, IJICIC*, 5(2), 1-ISIII08-025.

[9] Whiteley, J., Davis, A., & Mehrotra, Ahalt S. (1996). Observations and Problems Applying ART2 for Dynamic Sensor Pattern Interpretation. IEEE Transactions on Systems, Man, and Cybernetics- Part A: Systems and Humans, 26, 4, July.

[10] Moisen, C., Benítez-Pérez, H., & Medina, L. (2008). Non-contact Ultrasound for Flaw Characterisation using ARTMAP and Wavelet Analysis. DOI: 10.1504/IJMPT. 2008.022517, *International Journal of Materials and Product Technology*, 33(4), 387-403.

[11] Abbate, A., Koay, H., Frankel, J., Schroeder, S., & Das, P. (1997). Signal Detection and Noise Suppression Using a Wavelet Transform Signal Processor: Application to Ultrasonic Flaw Detection. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control , January 1997, 44(1), 14-26.

[12] Frazier, M. (1999). *An Introduction to Wavelets Through Linear Algebra*, Springer-Verlag.

[13] Kohonen, T. (1989). *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, Germany.

[14] Hassoum, H. (1995). *Fundamentals of Artificial Neural Networks*, Massachusetts Institute of Technology, USA.

[15] Nelles, O. (2001). *Non-Linear Systems Identification*, Springer-Verlag, Berlin, Germany.

[16] Blanke, M., Kinnaert, M., Lunze, J., & Staroswicki, . (2003). *Diagnosis and Fault Tolerant Control*, Springer-Verlag.

[17] Benítez-Pérez, Héctor, & Benitez-Perez, Alma. (2010). The use of WAVELET strategy and Self Organizing Maps for Feature Extraction and Classification for Fault Diagnosis. *International Journal of Innovative Computing, Information and Control, IJICIC*, 6(11), 4923-4936.