Hydroelectric power plant management relying on neural networks and expert system integration

J.M. Molina*, P. Isasi, A. Berlanga, A. Sanchis

Departamento de Informática, Universidad Carlos III de Madrid. Avda. de la Universidad 30, 28911 Leganés, Madrid, Spain

Abstract

The use of Neural Networks (NN) is a novel approach that can help in taking decisions when integrated in a more general system, in particular with expert systems. In this paper, an architecture for the management of hydroelectric power plants is introduced. This relies on monitoring a large number of signals, representing the technical parameters of the real plant. The general architecture is composed of an Expert System and two NN modules: Acoustic Prediction (NNAP) and Predictive Maintenance (NNPM). The NNAP is based on Kohonen Learning Vector Quantization (LVQ) Networks in order to distinguish the sounds emitted by electricity-generating machine groups. The NNPM uses an ART-MAP to identify different situations from the plant state variables, in order to prevent future malfunctions. In addition, a special process to generate a complete training set has been designed for the ART-MAP module. This process has been developed to deal with the absence of data about abnormal plant situations, and is based on neural nets trained with the backpropagation algorithm. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Predictive maintenance; Neural networks; ART; LVQ; Power plants; Expert systems

1. Introduction

The management systems of power plants continuously monitor different features of several subsystems: energy distribution (located outside the power plant), energy generators, motors, turbines, etc. The features of each subsystem are related to a set of variables that define the current situation in the subsystem. The evaluation of these variables gives some guidelines to human operators to detect abnormal situations in power plants. However, only a small set of variables can be observed and analyzed, to give useful information to the human observer.

On the other hand, simple automatic monitoring systems are, in general, able to analyze all the input values and generate alarms. Simple monitoring systems warn when the numerical value of a variable is outside the range set by human experts. More sophisticated systems include the possibility of forecasting future events, in order to anticipate malfunction situations. The main advantages to be gained by using predictive modules are the optimization of productivity, operating safety, and the protection of equipment against damage caused by malfunctions. These predictive systems are especially useful in distributed control centers, since the general philosophy is to model the plant and predict future situations instead of merely taking precautionary measures (Evans et al., 1994).

The development of more complex management systems, for predictive automatic surveillance, needs more advanced techniques. Some Artificial Intelligence (AI) techniques have been suggested in the literature for such systems (Raghavan and Simon, 1993; Loskiewicz-Buczak et al., 1993). These techniques give additional information for the identification of potential future failures and their causes. They analyze the relations

---

* Corresponding author. Tel.: +34-91-624-9116; fax: +34-91-624-9129.

E-mail address: molina@ia.uc3m.es (J.M. Molina).

between items of data coming from vital components of a power plant, thereby making it possible to perform effective preventive maintenance.

Knowledge-based systems (KBS) are able to provide valid ranges for variables, and to predict future abnormal situations from these values, using the expert knowledge of plant operators (Kim et al., 1994). In order to reduce the dependence of a final KBS on the knowledge-acquisition phase, some machine learning techniques (Yoshikawa et al., 1995) have also been applied.

In particular, the use of neural networks (NN) is a good approach to predictive control systems because of the highly non-linear relationships between the data and the abnormal situations (Isasi et al., 1996; Yonghong and van Cauwenberghe, 1995). While an expert system tries to mimic the response of a human operator, analyzing the same variables as the human, the neural networks overcome these limits and try to analyze non-linear relations between many different signals.

The man–machine communication is simple with an expert system, because the system could explain failures and also indicate why some decisions have been taken. This communication is hard with an NN, because the output depends non-linearly on different signals. So a complete integration of these two techniques could improve the global management system. Both the NN and the expert system information has been taken into account in the final decision. The former considers all the signals measured in the power plant, and makes a complex decision that considers the non-linear relations. The latter analyzes only a subset of the signals but can generate an explanation about the predicted events to support the human decision-making.

In this work, a hybrid predictive system named MAPAIS (an abbreviation of the Spanish for an advanced system for predictive maintenance incorporating audio and video) has been developed to forecast alarms in a hydroelectric power plant, and is included in a complete monitoring system named HYPERVISION. All the information that is processed to monitor the plant could be divided in three clusters: internal state variables, audio signals and video signals. The internal state variables are related to physical measures (temperature, pressure, water level, oil level, revolutions, etc.) taken in several devices (valves, pipes, rotors, switches, etc.). The audio signals are recorded in turbines in order to capture the sound of the rotors and the water in different power situations, from normal to critical regimes. The video cameras record images from external power-plant components. All the recorded information is shown to human operator through the HYPERVISION multimedia interface, and is used to evaluate the performance of MAPAIS.

The predictive system (MAPAIS) relies on advanced artificial intelligence techniques, audio and video signal processing, expert knowledge systems and NN modeling. Within the predictive modules of the system, the main goal of the NN maintenance module is to take advantage of the data recorded at every point in the hydroelectric plant. In this way, the predictive modules are acting as predictors of future incidents in the sense that they learn to detect situations of potential conflict.

Inside the predictive system, two different NN have been developed: the first receives the values of all the internal state variables (NN Predictive Maintenance-NNPM), and the other takes the spectral values of different acoustic signal measures in the turbines (NN Acoustic Prediction-NNAP). The NNPM developed here is an ART-MAP that predicts the future situation in the power plant on the basis of the internal state variables. The NNAP is an LVQ NN that classifies audio spectrums into normal or abnormal acoustic noise.

The use of NN using internal state variables when dealing with real situations (not simulations) adds an important problem: the acquisition of a complete set of data corresponding to abnormal situations is not possible, due to security and productivity restrictions in the plant. In this work, a special procedure for generating complete training sets has been developed. The non-linear relationships between the internal state variables have been obtained in a Multilayer Perceptron (MLP), trained using a backpropagation algorithm.

The system is currently working in the hydroelectric plant named Villalcampo I (IBERDROLA), which is located in the Zamora province of Northwest Spain. This plant has also been used for implementing and further testing of the MAPAIS project.

2. System architecture

In Fig. 1, a general view of the graphical supervisory system (HYPERVISION) is presented. HYPERVISION shows a schematic image of the electricity generating machine group, constituted by three independent generators named group 1, group 2 and group 3, respectively. The plant operator is able to obtain information about any system parameter by descending a hierarchical hyperlink structure.

The system could be decomposed into three levels (Fig. 2).

- Acquisition: this level records values of the input variables (digital/analog values).
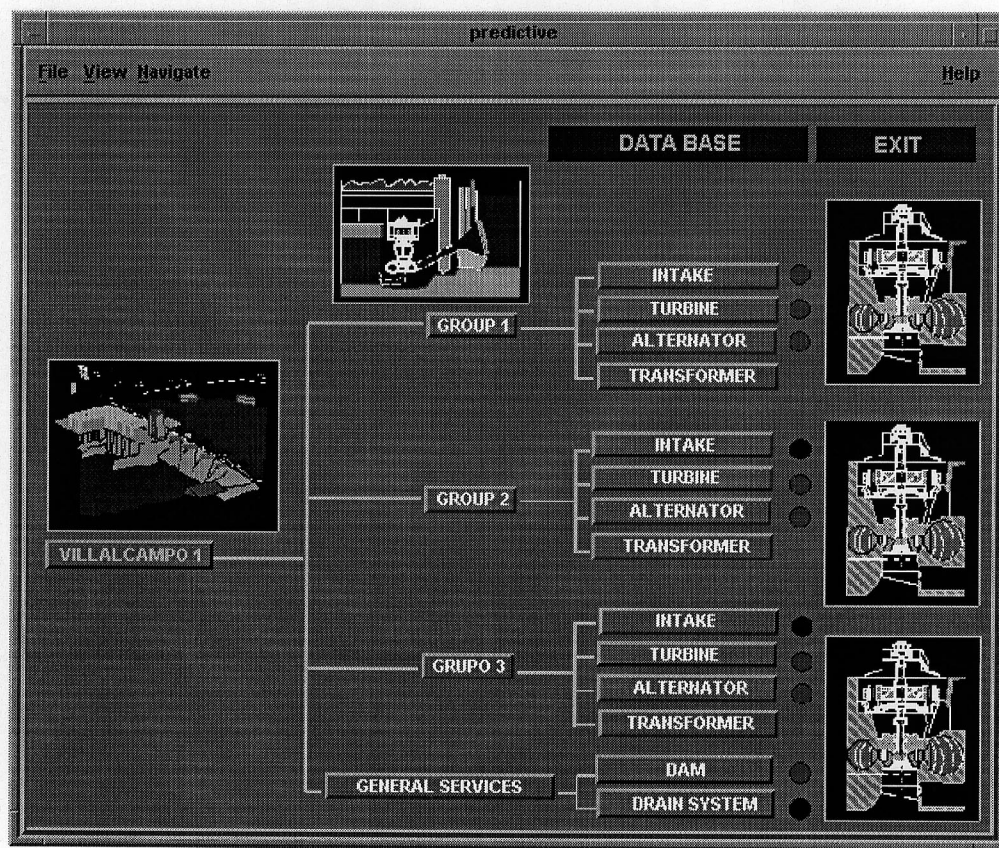- Prediction: the main part of the work, composed of several modules.

Fig. 1. Monitoring system interface.

● Presentation: where all the power-plant constituents are displayed in graphical interfaces to the plant staff.

Three different acquisition systems are installed in the power plant: SCADA, to measure internal state variables; COMPASS, to record acoustic signals; and TV

Fig. 2. Predictive system architecture.

cameras, to record external images of the power plant. The data needed to predict future situations are obtained from this information.

● SCADA: the pressure, temperature, levels, etc. of each device.
● COMPASS: the audio spectrum in the generating group.
● TV cameras: images of the electric switches that transfer the energy out of the plant.

The predictive maintenance system can be decomposed into several modules; each one is independent and predicts individually. All the modules are integrated into a distributed architecture. A brief description of modules is as follows.

● The NN predictive maintenance module produces predictions about future malfunctions, using a model of the plant. It relies on a NN model generated from past data and present events. The main drawback to this approach is that abnormal incidents seldom occur, and it is therefore difficult to isolate the main variables involved in every process.
● The image-processing module was designed to reinforce the measurements obtained from several electrical transceptors. Thus, image processing is
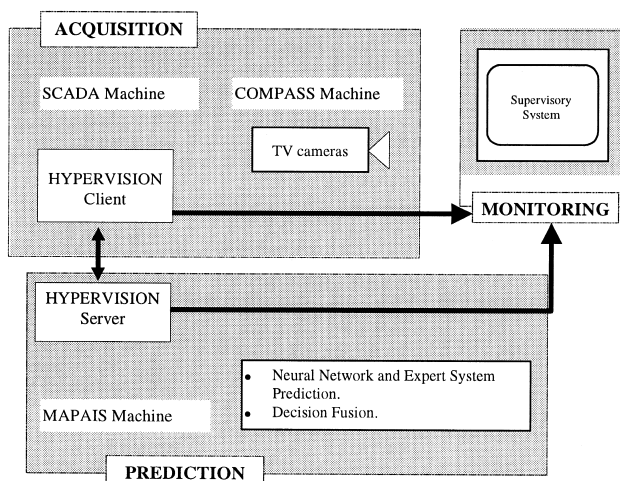
included to assess the functioning of the power switching systems.

- The rule-based diagnostic module will collect all the information available from human expert controllers, in order to predict problems.
- The NN acoustic prediction module takes advantage of audio data recorded at different points in every energy group, to generate a signal warning of future abnormal situations.
- Finally, the data fusion and decision module collects all the available data to integrate the information by means of Dempster–Shafer theory (Gonzalez and Dankel, 1993).

A general data-flow scheme of the system is shown in Fig. 3.

The graphic interface to the predictive modules is a subtree of the monitoring system. The normal and abnormal situations are represented as traffic lights in the interface: green (normal) and red (alarms). When a module generates an alarm, it activates the interface panel (red lights) shown in Fig. 4; at the same time, a red light turns on in the graphic interface of Fig. 1. In

this case, the human operator could navigate following the hyperlink structure to find the variable that has produced the alarm, and the predictive module that has generated it. This information is included in the interface shown in Fig. 4. Each variable has two lights, the left hand one related to the decision taken by the expert system, and the right hand one linked to the NNPM decision. The possible combinations of lights are as follows.

- Only the left light on: the expert system detects an alarm, the NNPM does not detect an alarm, and the Data Fusion and Decision Module does not detect an alarm.
- Only the right light on: the NNPM detects an alarm, the expert system does not detect an alarm, and the Data Fusion and Decision Module does not detect an alarm.
- Both lights on: the Data Fusion and Decision Module detects an alarm although neither the NNPM nor the expert system detect an alarm.

The results of the NNAP are presented independently in the interface, and are not integrated in the fusion
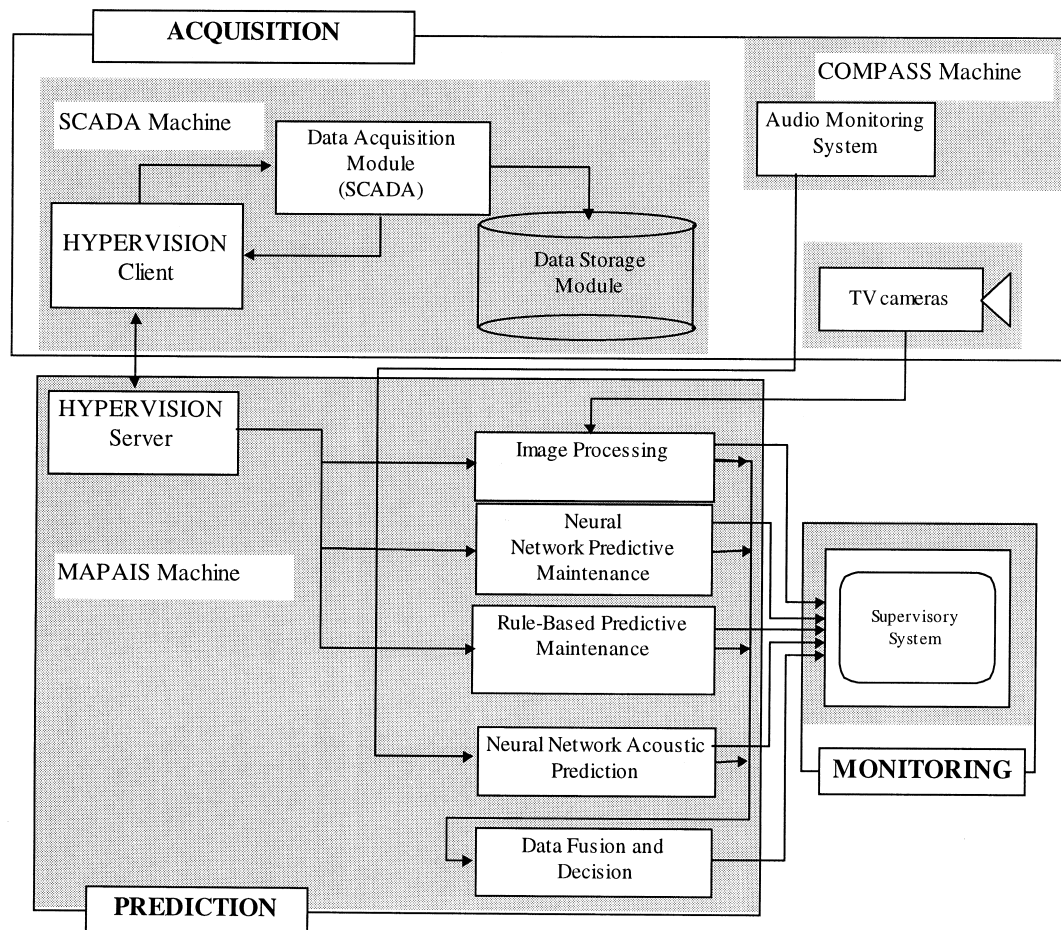

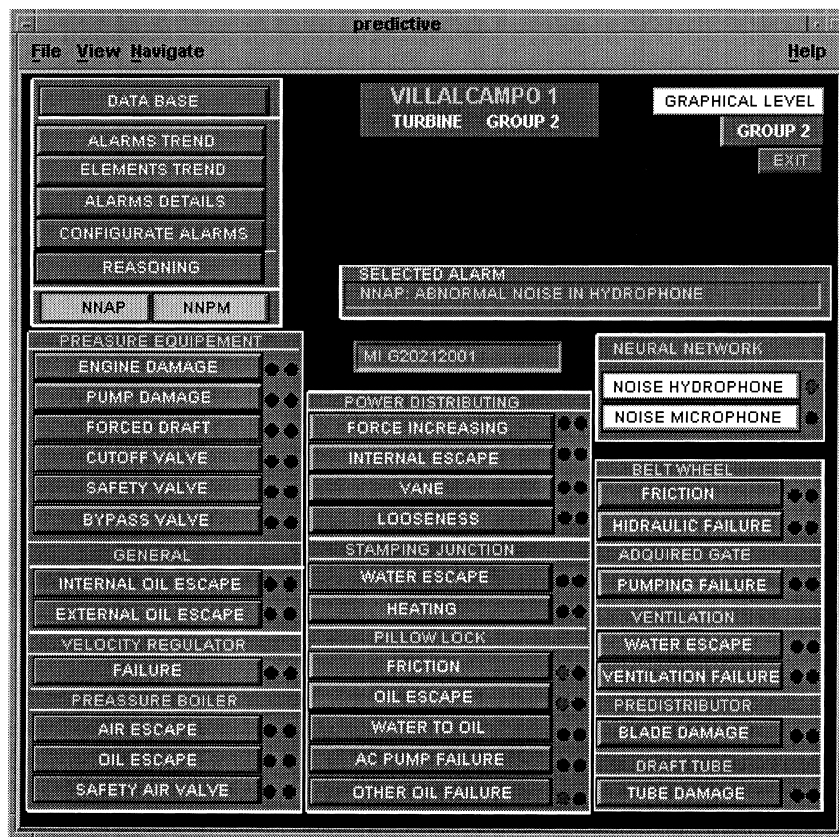Fig. 3. Diagram of the data flow between modules.

Fig. 4. Predictive system interface.

process. The NNAP handles different data from the expert system and NNPM; therefore, the results do not give any information about the internal state variables. The acoustic data that is acquired is recorded from the electricity-generating machine group, and is only analyzed in the NNAP. The NNAP does not receive any other data from the power plant; therefore no decisions are obtained about the other subsystems.

Besides monitoring the actual situation in the power plant, the tool allows historical data to be accessed. The past data is used to improve the decision process. These data are used in two ways: to improve the expert system, by comparing the decisions of a human operator and the system, and to re-train the NN in order to adjust the net to new situations that have never happened before.

3. ART and LVQ equations

3.1. ART

Adaptive Resonance Theory (ART) was first introduced by Carpenter and Grossberg (Carpenter and Grossberg, 1986a, 1986b, 1987a) and extended by analog inputs (Carpenter and Grossberg, 1987b, 1987c) as

an attempt to solve the plasticity–stability problem. The NN designed in this way is stable in face of irrelevant new inputs, while adapting to relevant ones. Two layers, called F_A and F_B, compose the topology of the network (Fig. 5). The cells in the F_A layer accept inputs from the environment, and are fully connected (bottom–up connections) to the F_B layer. Each cell in the F_B layer represents a pattern class. The F_B has also connections to all the cells in the F_A layer (top–down) connections. Each cell in the input layer actually consists of six units working in synchrony (Fig. 6). The
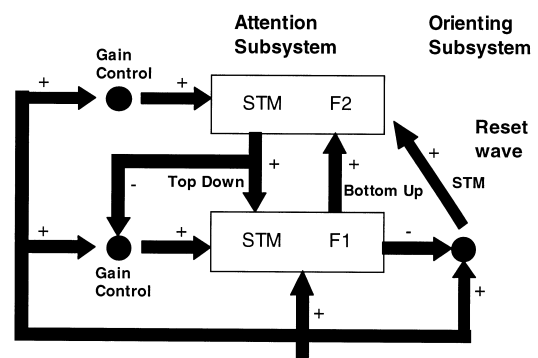

Fig. 5. ART architecture.

objective of this division is to normalize the input, and to store the patterns for fair comparisons.

The complete operational scheme of the system is as follows.

(a) The kth input signal $A_k = (a_1^k, \ldots, a_n^k)$ is introduced in the F_A layer.

(b) The input is normalized inside each F_A cell through the $a_i \rightarrow b_i \rightarrow d_i \rightarrow c_i \rightarrow g_i$ sequence, and the final signal (g_i) is sent to the F_B layer through the bottom–up connections.

(c) In the F_B layer a competition takes place, until only one cell remains active.

(d) The winning cell sends its signal back to the F_A layer through the top–down connections.

(e) This signal arrives to the macro-cells in F_A and an output c_i is produced through the $g_i \rightarrow e_i \rightarrow d_i \rightarrow c_i$ sequence.

(f) This c_i signal is compared with the g_i signal produced in step (b):

(f1) If the difference between the two signals exceeds some pre-determined value (vigilance parameter) the winning cell is not the real representative of the input. This cell is temporarily removed from the F_B layer and the process jumps to step (b).

(f2) If the difference between the two signals does not exceed the vigilance parameter, the winning cell is said to be the real representative of the input. Then the input is learned in the bottom–up and top–down connections of the winning cell.

The output calculations for macro cells in the F_A layer are shown in Eqs. (3)–(8).
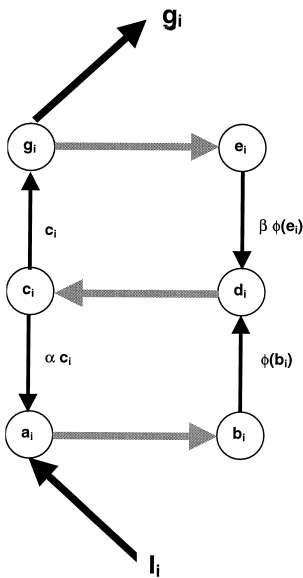
$$a_i = a_i^k + \alpha \times c_i \tag{1}$$

Fig. 6. F_A macro cell architecture.

$$b_i = \frac{a_i}{\gamma + \|A\|} \tag{2}$$

$$c_i = \frac{d_i}{\gamma + \|D\|} \tag{3}$$

$$d_i = \Phi(b_i) + \beta \times \Phi(e_i) \tag{4}$$

$$e_i = \frac{g_i}{\gamma + \|G\|} \tag{5}$$

$$g_i = c_i + \sum_{j=1}^{p} \varphi(h_j) \times v_{ji} \tag{6}$$

where $0 < \alpha < 1$, $\gamma > 0$, $\beta > 0$, v_{ji} is the connection from the jth cell in F_B to the ith cell in F_A, $A = (a_1, \ldots, a_n)$, $D = (d_1, \ldots, d_n)$, $G = (g_1, \ldots, g_n)$, $\Phi()$ is a sigmoid threshold function, $\varphi()$ is a delta function, h_j is the output of the cells in F_B layer, is calculated by a linear combination of g_i and the weights.

The learning process is guided by Eqs. (9)–(12).

$$\frac{dv_{ji}}{dt} = \varphi(h_j) \times (g_i - v_{ji}), \quad \text{top-down connections} \tag{7}$$

$$\frac{dw_{ji}}{dt} = \varphi(h_j) \times (g_i - w_{ji}), \quad \text{bottom-up connections} \tag{8}$$

$$\frac{\rho}{\gamma + \|F\|} > 1, \quad \text{reset condition} \tag{9}$$

$$f_i = \frac{c_i + \xi \times g_i}{\gamma + \|C\| + \xi \times \|G\|}, \quad \text{each component of } F \tag{10}$$

ART models have shown a special ability in pattern classification, even in complex domains (Carpenter and Grossberg, 1988).

3.2. LVQ

LVQ was developed by Kohonen (1986) as a method for the supervised clustering of a space into mutually exclusive classes. This is accomplished by defining a small set of centroids, and yields a Voronoi tessellation of the space (that is boundaries are perpendicular bisector planes of lines joining pairs of neighboring prototype vectors). An input vector $\underline{x}$ is said to belong to the class represented by the centroid, c_i, if and only if:

$$|c_i - \underline{x}| \le |c_j - \underline{x}| \quad \forall j, \tag{11}$$

where $|.|$ denotes Euclidean distance.

In the LVQ1 training phase, the centroid with the minimum distance to the presented pattern is said to

be the "winn", c_{winn}, and is updated using the following rule:

$$\Delta c_{\text{winn}} = \begin{cases} +\alpha(t)(\underline{x} - c_{\text{winn}}) & \text{if the class is correct} \\ -\alpha(t)(\underline{x} - c_{\text{winn}}) & \text{if the class is incorrect} \end{cases}$$

$$(12)$$

The overall behaviour of the algorithm is to place the centroids in the mass centers of the distributions. Stable solutions can be obtained if a sufficient number of centroids are available and the density functions are well behaved.

4. Data acquisition

In the power plant, as was commented previously, three different types of information are gathered: internal state variables, sound spectra and external images. The predictive system is based on the first and second types of information, which are used as the input of each subsystem. The data acquisition follows different methods according to the type of information that must be captured.

4.1. Internal state data acquisition

The power plant has an electricity-generating machine group, comprising three independent generators, named group 1, group 2 and group 3, respectively. These generators are constantly monitored. Every group can be sketched as in Fig. 7, where part "A" represents the alternating current generator and "B" represents the hydraulic turbine (obviously, this part is underwater). Many sensors, located at different points in the machinery, serve to monitor the group.
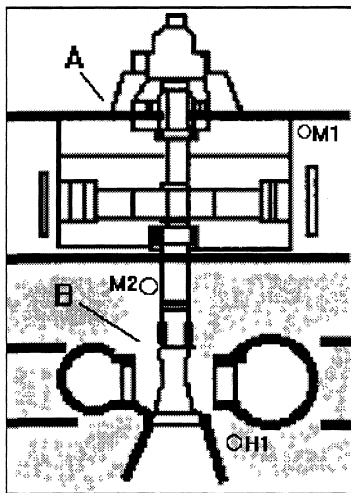

Fig. 7. Sketch of a generating group, comprising an alternator "A" and a turbine "B". Acoustic data acquisition points are denoted as M1, M2, H1.

A generating group is decomposed into functional units. Each unit is divided into different elements. The decomposition of each generating group is shown in Table 1. The data-acquisition equipment, which was already working in the plant, is a SCADA system. This system is able to register the values of a set of variables in an element.

4.2. Acoustic data acquisition

The data-acquisition equipment (COMPASS Machine Monitoring System Type 3540), that was already working in the plant was considered flexible enough to perform all the preprocessing. A baseband spectrum mode was selected with a frequency span of 1 Hz to 5 kHz, which is considered to be adequate for the purpose of this work. The inclusion of higher frequencies would yield no improvement. Indeed, a decrease in performance of the algorithm could be observed due to extra high-pass noise introduced in the system.

Hanning windowing was used, and the FFT spectrum was computed using an averaging mode spanning 10 epochs in order to produce an output every 10 min. This time lap was considered sufficient as changes in groups tend to be slow. A typical spectrum collection is depicted in Fig. 8, corresponding to the microphone in alternator 3.

The placement of microphones M1 and M2 and hydrophone H1 is marked with circles, representing microphones in the alternator and the turbine, as well as a hydrophone in the turbine in Fig. 7.

In order to obtain the initial data (normal and abnormal) needed to train the prediction module, each group was forced to work at several different power production rates (Table 2). An expert was able to classify each situation using lateral information obtained from other sources. In Table 2, the following situation is described: group 1 is working at a fixed rate and group 3 works at various different rates. group 1 is generating background noise, which corrupts the measurements of group 3. This background noise is unavoidable, and the system therefore has to cope with it. These data are complex, and involve both the produced power and an expert-based classification of the situation. It is important to mention that this knowledge-based classification is carried out with the concurrent expert observation of many other variables. Acoustic information is never used, exclusively, to justify expert decisions.

Every spectrum can be considered as a point in an N-dimensional space, corresponding to either "safe" situations or future abnormal situations. In this way, the problem is reduced to a classification task in an N-dimensional space. Nevertheless, some special characteristics arise when analyzing the data. First of all, the variance in FFT samples is not evenly distributed in

Table 1
Decomposition of a generating group

Functional unit	Element	Example of variable
Water intake	Intake gate	*Gate_Intake_Gate*
	Chain drive	*Motor_Pressure*
Turbine	Pressure equipment	*Level*
	Velocity regulator	*Deposit_Level*
	Pressure boiler	*Level*
	Power distributing	*Position*
	Stamping junction	*Temperature*
	Pillow lock	*Oil_Temperature*
	Belt wheel	*Temperature*
	Bailing bomb	*On*
Alternating current generator	Rotor	*Velocity*
	Stator	*Intensity_Generator*
	Stimulation	*Regulator_Position*
	Equipment	
	Push pillow lock	*Metal_Temperature*
	Low pillow lock	*Low_Metal_Temperature*
	High pillow lock	*High_Metal_Temperature*
	Brakes	*Normal_Pressure*
	Cooling system	*Air_Temperature*
Current transformer		*Differential_Protection*
Dammed up waters		*Closed*
Exhaust system		*Fired*

the frequency range, as can be observed in Fig. 6. This is due to the time-varying humming noise from the water when traveling through the pipes, plus a lowpass filtered component of noise from adjacent groups (data is usually recorded when several groups are working together). Nevertheless, despite the non-uniform variance, the distribution of every FFT sample can be considered as locally gaussian. It is also important to note that vectors are not evenly distributed between the two classes, because "normal functioning"
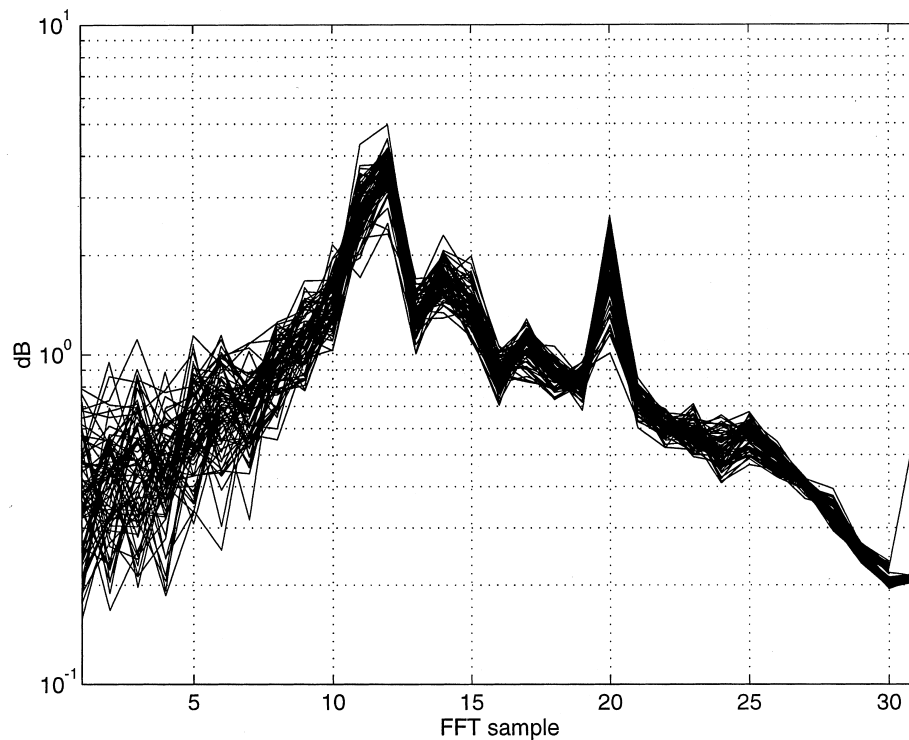

Fig. 8. Ensemble of FFT spectra representing one of the normal regimes in group 3.

will be the predominant situation, and "potential conflicts" will seldom arise. At the same time, the two events do not have the same importance, because wrongly classifying a situation as conflictive is less problematic than failing to predict a potentially dangerous situation. A similar problem arises in radar detection where a false alarm is not as problematic as the non-detection case (Loskiewicz-Buczak et al., 1993). In this case, this idea is reinforced by the fact that the data fusion and decision module has been designed to consider every sound-module warning. The sound module is therefore used as a warning system that is very sensitive to alterations in sound information (conservative behaviour); when activated, it forces the system to take account of information from other variables (prediction modules). Thus, the system is operating outside the typical Bayesian arena, because the most frequent events are not the most relevant ones. This issue has been taken into account.

5. Predictive system: expert and neural system

Both the expert system and the NNPM produce predictions about future malfunctions, by using a model of the plant. The inputs to these systems are the variables recorded by means of the SCADA system. The number of variables, for a generating group, is 106 and these are stored to obtain time series data that represent the normal power-plant functioning over several years.

All the variable values are inputs to the predictive NN. However, to decide the situation in the power plant, a human operator is able to analyze only 15 variables. Like the human operator, the expert system developed here treats only these 15 variables as inputs, and uses the range of each variable and the relations among them to take decisions.

In both cases, the output of each system is an alarm that represents a future abnormal situation of the power plant. An alarm points out a future abnormal functioning of one of the elements. The final decision is taken in the fusion module, where the information about alarms is fused, using the belief and plausibility

Table 2
Power situations used to obtain initial acoustic data

Generated power (Mw)	Expert's classification
5	Abnormal regime
9	Abnormal regime
13	Normal regime
20	Normal regime
25	Normal regime
30	Normal regime
31	Abnormal regime

degree of each decision. The ultimate goal of the system is to carry out predictive maintenance. This task could be reduced to a classification task by introducing each different set of input signals in a different category. Each class can be related to a different prediction of the power plant's behavior. The acquisition system is able to record 106 variables, v_i, and the classified vector comprises a value of each variable, $(v_1, v_2, v_3, \ldots, v_{106})$.

When dealing with NN, a complete set of vectors is needed for the training phase. The set must include vectors representing normal situations, and vectors for each possible alarm. Relations between power plant states and alarms are highly dynamic and non-linear. Thus, NN provide a good way of approximating those relations.

It is interesting to outline some specific features of these kinds of systems.

1. It is not possible to know in advance how many classes will be needed.
2. The dynamic nature of input data necessitates the "on-line" creation of new classes if none of the existing ones are appropriate.
3. The training set is reduced to the stored data from real situations. Stored data correspond to situations of either normal working or an alarm situation.

ART-NN models (Carpenter et al., 1991) are the best candidates for such systems due to the first and second characteristics. They are unsupervised and are continuously learning, and as a consequence can produce new categories while the system is working (classifying).

In order for the NN (ART) to produce a generalized behavior, a complete training set is needed. The generation of the training set above mentioned is not possible with real data from the power plant, so it is necessary to generate what can be called an "artificial training set". To generate this artificial training set, a special process is considered. The special generation process is based on an expert system and a set of neural networks, NN_i, one for each variable, v_i, considered by the expert system.

The expert system generates alarms, based on expert human knowledge. This system is composed of a small number of rules (two or three) for each possible alarm. The expert system does not relate more than four variables in each rule, and globally only considers 15 variables. The expert system produces the alarm category and the variables responsible for each alarm. As the expert system only knows the relations among 15 of the 106 system variables, it is impossible to generate a complete vector, defined by all variable values representing this alarm.

The generation of the values of the remaining variables in abnormal situations is based on the relations between these variables and the variable that

produces the alarm. This relation is only known in normal situations and is non-linear, so a NN is proposed to obtain the complete vector, (v_1, v_2, v_3,...,v_{106}), from the variable value. The NN$_i$ (i = 1–15) is trained with a backpropagation algorithm using stored normal working data.

The stored data is recorded continuously, and supplies a good database with which to train any system that does not require data from malfunction situations. The main problem of this database appears when the learning process needs information about alarms in the past, as in the case of NN, and this information needs to be generated for training. A special procedure is necessary to obtain a complete training set. The training set generation procedure for the ART model is as follows.

1. The expert system generates an alarm, A_k, and the variable that produces it, v_i.
2. The variable, v_i, is introduced with its value in its corresponding, NN$_i$, (previously trained for that variable).
3. The output of NN$_i$ is a complete vector, (v_1, v_2, v_3,...,v_{106}), that represents all the variable values for that alarm, A_k.
4. This vector is introduced in the training set, labeled with the name of the alarm, A_k.
5. The previous four steps are repeated until the training set is complete.
6. The real stored data (normal situations) are added to the training set with the label "no-alarm".

A schematic description of the process is shown in Fig. 9. The training set built in this way has been used to train an ART-MAP NN.

6. Neural-network acoustic prediction

The next step was to construct a classifier for the sound signals, NNAP, produced by the above-mentioned working regimes (see Table 2). Several well-known classification methods were considered, and the one that best fits the requirements has been implemented in the sound module.

An MLP is very appropriate when handling high-dimensional problems because it uses hyperplanes very efficiently (Rognvaldsson, 1993). Nevertheless, these capabilities deteriorate when dealing with sparse data, which are seen to occur in the application. Other possible methods, e.g., the k-NN ("k" nearest neighbours) method, would incur in high computational and storage requirements. The ART algorithm (Grossberg, 1987) allows on-line training as new patterns are presented, but in this work such capability is of limited utility as no information is available about newly formed categories, and thus one would lose generalization capabilities. These considerations led to the selection of LVQ (Kohonen, 1986) as the technique with the best simplicity–performance ratio. Some researchers (Bartal et al., 1994) have also been pointed out some advantages of this method. This algorithm has been revisited in Section 3.

As mentioned in Section 3, not every situation has the same importance. In order to design a conservative system (which will produce an alarm even when presented with a new sound), it is necessary to increase the probability of false alarms when predicting future incidents, in order to reliably predict a real incident which could produce a disastrous effect. The solution can be achieved by constructing a "risk function" for every decision (incident/safe) using conditional prob-
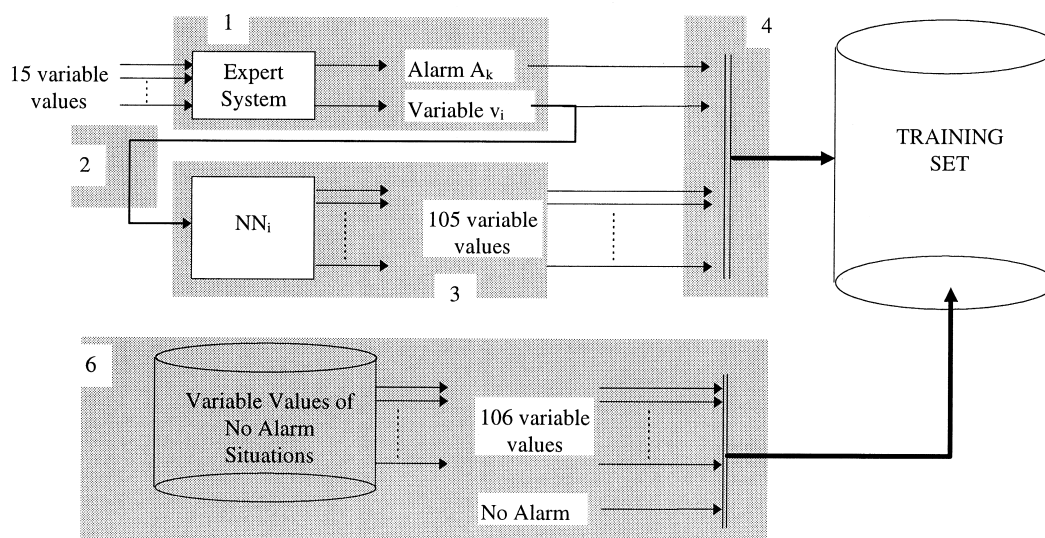

Fig. 9. Special training process description.

ability functions when a pattern $\underline{x}$ is presented (Pao, 1989):

$$R_{\text{safe}}(\underline{x}) = l_{11}P(safe|\underline{x}) + l_{12}P(\text{incident}|\underline{x}) \tag{13}$$

$$R_{\text{incidence}}(\underline{x}) = l_{21}P(\text{safe}|\underline{x}) + l_{22}P(\text{incident}|\underline{x}). \tag{14}$$

In this case, $l_{12} = k\, l_{21}$ $(k > 1)$, and $l_{11} = l_{22} = 0$, yield a biased decision boundary when detecting future incidents:

"future incident" $\Leftrightarrow R_{\text{incident}}(\underline{x}) < R_{\text{safe}}(\underline{x})$

$$\Leftrightarrow P(\text{"incident"}|\underline{x}) > (1/k)P(\text{"safe"}|\underline{x}). \tag{15}$$

The same concept can be directly translated to the evaluation of distances, by multiplying the distance to the centroids of class "future incident" by k.

$$d(c_i, \underline{x}) = \begin{cases} |c_i - \underline{x}| & \text{if} \quad c_i \in \text{"safe"} \\ k|c_i - \underline{x}| & \text{if} \quad c_i \in \text{"future incident"}. \end{cases} \tag{16}$$

This new distance measure achieves a conservative system, and thereby extends the region of future risk with respect to conventional Bayesian decision boundaries.

Nevertheless, such a modification can lead to the misclassification of certain patterns because LVQ1 tends to place the centroids in the mass centers. A solution is obtained by using a modification of this basic algorithm, called LVQ2 (Kohonen, 1990), which is closer to Bayes Decision Theory, as it places the centroids near the decision boundary, thereby reducing the risk of misclassification. In this case, the two nearest centroids are considered, and are only updated if the input vector is misclassified by the winning centroid, the next nearest neighbor has the correct class, and $\underline{x}$ falls into a predefined window:

$$\Delta c_{\text{winner}} = -\alpha(t)(\underline{x} - c_{\text{winner}})$$

$$\Delta c_{\text{runner-up}} = +\alpha(t)(\underline{x} - c_{\text{runner-up}}). \tag{17}$$

In general, LVQ techniques require extensive off-line training time, and are unable to incorporate new data without complete re-training. These characteristics might be critical for certain applications but are acceptable in the case studied here because the time lapses are long enough to allow re-training, and also because new situations seldom arise (only during an initial testing phase work an expert assessment will be available and new, correctly-classified data, be incorporated).

7. System performance on a hydroelectric plant: final conclusions

This paper has described a practical method to assess the control operation of a power plant based on a dynamic predictive module. The proposed strategy provides key improvements over human control strategies. The proposed method has been applied to the Predictive module of the Villalcampo I hydroelectric plant in Zamora, Spain.

During a period of a week, tests were carried out on the maintenance system under the strict supervision of the technical staff of the plant. During the test period, final tuning of the system was needed due to the exceptional weather conditions under which the system had been tested (heavy rainfall that overloaded the dam capacity, so that the floodgates had to be opened to allow the excess water to run off). This unusual situation allowed the generalization capacity of the system to be evaluated because the NNs had to deal with unusual data.

The good performance of the ART model, developed for forecasting, proves the ability of the training-set generation procedure to estimate variable values related to future abnormal situations.

Three main incidents happened, two of them caused by data from hydrophones at turbines in groups 2 and 3, and the other one caused by the microphone at a turbine in group 3. The extracted conclusions are as follows: generalization of the data recorded at the hydrophones is poor because of the high variability in sound spectra even in non-changing environments, due to an excess of noise. The incident produced by data from the microphone in turbine at group 3 was identified by experts as a "normal situation", and was therefore included as new data for re-training, together with some data collected while developing the NNAP. This phenomenon was explained as follows. Due to the long period of time between the initial data collection and the final testing, some special minor variations occurred in the working conditions, and hence such re-training was considered as normal final tuning.

After that, the system was considered to have reached an ideal working behavior. It was finally included in the monitoring system on January 1996 and it is currently being used as a decision-making aid.

The system has been designed with the probability of detecting a true positive problem equal to one. The security restrictions of the power plant have imposed that design specification and level of performance should be observed in all the tested systems.

A comparison between the predictions made respectively by humans and by the predictive system was subsequently carried out, analyzing the number of false alarms, in order to evaluate the results of the proposed system (Kobayasli et al., 1998). Each day, a number of possible alarms are generated by the operator or by the predictive system. These tentative alarms could be either a real alarm or a false one. The percentage of false alarms in 1996, over a month, is shown in Figs.
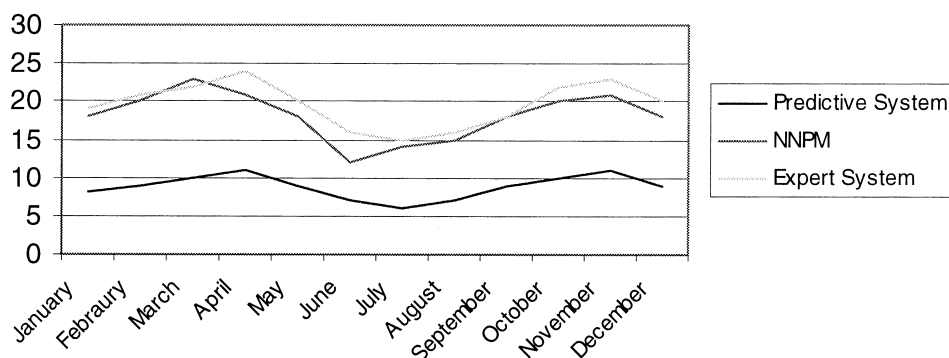
Fig. 10. False alarms in a hydroelectric power plant: predictive system, NNPM and expert system.
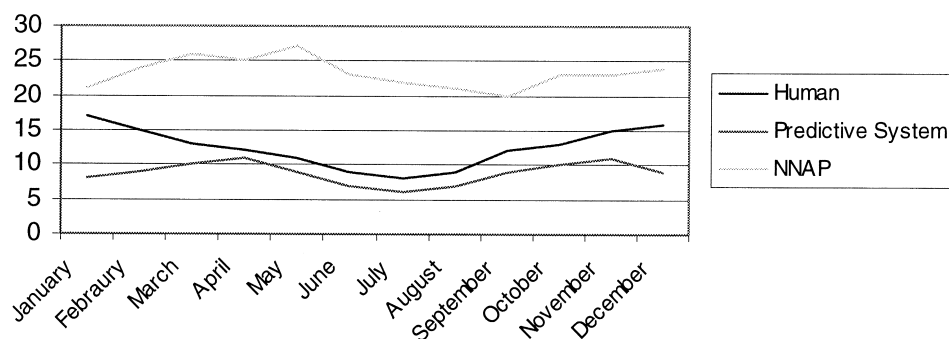

Fig. 11. False alarms in a hydroelectric power plant: human operator, predictive system and NNAP.

10 and 11. In Fig. 10 the results of the operator are compared with those of the predictive system (expert system–NNPM).

It can be seen that the human operator obtains worse results than the predictive system but better results than those achieved individually by each of the modules that compose the predictive system. The predictive system has obtained best results; that is, it has detected all true positive alarms, with the lowest production of false alarms.

Fig. 11 shows the false alarms predicted by the three systems: human, predictive and NNAP. The percentage of false alarms of the NNAP duplicates the percentages of predictive system and the human operator.

In Fig. 12, the results of the full system are summarized.

NN predictive maintenance gives the most accurate predictive control since there are no maintenance considerations involved. The accuracy of NNAP is lower than any of the other systems due to the variability of acoustic signals. The results show the necessity for integrating partial results from each subsystem.
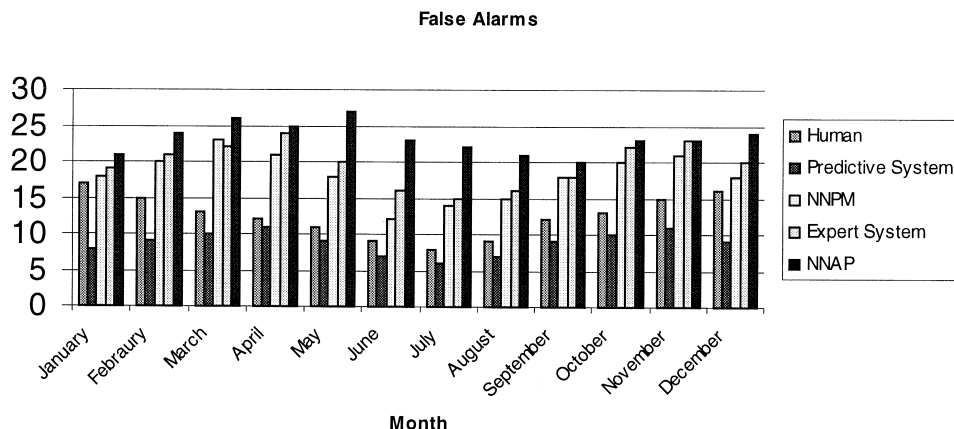

Fig. 12. False alarms in a hydroelectric power plant.

Acknowledgements

References

Bartal, Y., Lin, J., Uhrig, R.E., 1994. Nuclear power plants transient diagnostics using LVQ or some networks don't know that they don't know. International Conference on Neural Networks 6, 3744–3749.

Carpenter, G., Grossberg, S., 1986a. Absolutely stable learning of recognition codes by a self-organizing neural network. In: Proc. of AIP Conference: Neural Networks for Computing, 77–85.

Carpenter, G., Grossberg, S., 1986b. Adaptive resonance theory: stable self-organization of neural recognition codes in response to arbitrary lists of input patterns. In: Proc. of 8th Annual Conference of the Cognitive Science Society, 45–62.

Carpenter, G., Grossberg, S., 1987a. A massive parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics and Image Understanding 37, 54–115.

Carpenter, G., Grossberg, S., 1987b. ART2: self-organization of stable category recognition codes for analog input patterns. In: Proc. of the IEEE First International Conference on Neural Networks, vol II, 727–736.

Carpenter, G., Grossberg, S., 1987c. ART2: self-organization of stable category recognition codes for analog input patterns. Applied Optics 26, 4919–4930.

Carpenter, G., Grossberg, S., 1988. The ART of adaptive pattern recognition by a self-organizing neural network. Computer, March, pp. 77–88.

Carpenter, G.A., Grossberg, S., Reynolds, J., 1991. ARTMAP: supervised real-time learning and classification of non-stationary data by a self-organizing neural network. Neural Networks 4, 565–588.

Evans, J.T., Gomm, J.B., Williams, D., Lisboa, P.J.G., 1994. Online modeling and predictive control using a neural network. In: Proc. IASTED International Conference, Modeling, Simulation and Control in the Process Industry, 83–86.

Gonzalez, A.J., Dankel, D.D., 1993. The Engineering of Knowledge-Based Systems. Englewood Cliffs, NJ, Prentice Hall.

Grossberg, S., 1987. Competitive learning: from interactive activation to adaptive resonance theory. Cognitive Science 11, 23–63.

Isasi, P., Molina, J.M., Navia, A., 1996. Hidroelectric power plant predictive maintenance relying on neural network acoustic module. In: Proc. International Conference on Neural Information Processing, 1175–1180.

Kim, H.G., Choiç, S.S., Kang, K.S., Chang, S.H., 1994. Development of an on-line expert system for integrated alarm processing in nuclear power plants. Transactions of the American Nuclear Society 71, 121–123.

Kobayasli, T., Tani, T., Abe, N., Miyamoto, S., 1998. Comparison between human supervisory control and hierarchical control system based on human's knowledge in petroleum plant. In: Proc. of World Congress on Computational Intelligence, Anchorage, Alaska, pp. 200–204.

Kohonen, T., 1986. Learning vector quantization for pattern recognition. Technical Report, TKK-F-A601, Helsinky University of Technology, Department of Technical Physics, Laboratory of Computer and Information Science.

Kohonen, T., 1990. In: Eckmiller, R. (Ed.), Statistical pattern recognition revisited. Advanced Neural Computers, Elsevier, Amsterdam.

Loskiewicz-Buczak, A., Alguindigue, I.E., Uhrig, R.E., 1993. Vibration analysis in nuclear power plant using neural networks. In: Proc. 2nd ASME JSME Nuclear Engineering Joint Conference, 2, 43–51.

Pao, Y.H., 1989. Adaptive Pattern Recognition and Neural Networks. Addison-Wesley, Reading, MA.

Raghavan, R., Simon, B.H., 1993. Advanced plant maintenance and surveillance system for the nuclear power plants of the next century. In: Proc. 2nd ASME JSME Nuclear Engineering Joint Conference, 2, 693–697.

Rognvaldsson, T., 1993. Pattern discrimination using feedforward networks: a benchmark study of scaling behaviour. Neural Computation 5, 483–491.

Yonghong, D., Van Cauwenberghe, A., 1995. Use of neural networks for non-linear predictive control. In: Proc. 3rd European Control Conference vol. 3, 2760–2765.

Yoshikawa, S., Saiki, A., Ugolini, D., Ozawa, K., 1995. Nuclear power plant monitoring and fault diagnosis methods based on the artificial intelligence technique. In: Proc. SMORN VII, a Symposium on Nuclear Reactor Surveillance and Diagnostic, 1(4.9), 1–9.