

## Towards a Neuro-Fuzzy System for Time Series Forecasting in Maintenance Applications

M. El-Koujok, R. Gouriveau, N. Zerhouni

*FEMTO-ST Institute, UMR CNRS 6174 - UFC / ENSMM / UTBM,  
Automatic Control and Micro-Mechatronic Systems Department,  
24, rue Alain Savary, 25000 Besançon, France  
e-mail: {mohamed.elkoujok, rgourive, zerhouni@ens2m.fr}*

---

**Abstract:** In maintenance field, industrial and research communities take a growing interest in the "understanding of the degradation phenomenon". Within this frame, the general purpose of the paper is to explore the way of defining a prognostic system that is able to approximate and predict the degradation of equipment. Prognostic is defined and the evolution of developments on forecasting methods is studied. A neuro-fuzzy system for failure prediction based on the ANFIS model is closely studied and a pre-treatment of data is proposed to perform short term accurate and mid-term reliable predictions.

---

### 1. INTRODUCTION

The high costs in maintaining complex equipments make necessary to enhance maintenance support systems and traditional concepts like preventive and corrective strategies are progressively completed by new ones like predictive and proactive maintenance (Lung *et al.*, 2003, Muller *et al.*, 2008). Thus, prognostic reveals to be a very promising activity as it should allow avoiding inopportune maintenance spending. That said, prognostic must obviously not be the single maintenance task as its applicability, cost and effectiveness can vary from an industrial plant to another. Within this frame, the objective of the work partially reported here is to define a prognostic system that is able to take into account the dynamic of the equipment. A central problem can be pointed out from this: the accuracy of a prognostic system is related to its ability to approximate and predict the degradation of equipment. In other words, starting from a "current situation", a prognostic tool must be able to forecast the "future possible situations". From the research point of view, many developments exist to support this forecasting activity (De Gooijer and Hyndman, 2006). However, in practice, choosing an efficient technique depends on classical industrial constraints that limit the applicability of the tools: available knowledge (model, past experiences...), evolution of the monitored system (linearity, periodicity...), decision criteria (precision, reasoning...), real implementation requirements (flexibility, complexity, computation time...).

In this context and according to the maintenance field, the purpose of the work is to point out a predictive tool and to study the applicability of it. In this way, the paper is organized in three parts. First, the concept of "prognostic" is positioned within the maintenance strategies and an overview of prediction approaches is given (section 2). The evolution of forecasting systems based on artificial intelligence is then more closely studied (section 3). A specific class of methods emerges from it: the neuro-fuzzy systems. One of them is also presented in section 4, the ANFIS, and an improvement

of this tool to perform accurate mid term predictions is proposed and illustrated on an academic example.

### 2. FROM THE PROGNOSTIC PROCESS TO THE CHOICE OF A FORECASTING APPROACH

#### 2.1 *The emergence of prognostic in maintenance policies*

In the maintenance area, one usually speaks about fault detection and diagnostic, response development and scheduling of these actions. Briefly these steps correspond to the need, of "perceiving" phenomena, of "understanding" them and of "acting" consequently. However, rather than understanding a phenomenon which has just appeared like a failure (a posteriori), it's convenient to try to "anticipate" its manifestation in order to consequently and, as soon as possible, resort to protective actions. This is what could be defined as the "prognostic process". Also, industrials show a growing interest in this thematic which becomes a major research framework; see recent papers dedicated to condition-based maintenance (Ciarapica and Giacchetta, 2006, Jardine *et al.*, 2006, Vachtsevanos *et al.*, 2006). Prognostic is recognized as a key feature in maintenance strategies as it usefully completes maintenance tasks. However, this concept isn't well defined. So, let's have a closer look on it.

#### 2.2 *Prognostic and forecasting*

In scientific literature, "prognostic" is sometimes related to fracture and fatigue (Farrar *et al.*, 2003). Some authors define it as a process whose objective is to predict the remaining useful life before a failure occurs given the current machine condition and past operation profile (Jardine *et al.*, 2006). Prognostic is also seen as the "chance" that a machine operates without a fault / failure up to some future time (Lin and Makis, 2003). Thus, a salient characteristic of prognostic can be pointed out: "prognostic" is mostly assimilated to a "prediction process" (a future situation must be caught).

The ISO comity defines prognostic as "the estimation of time to failure and risk for one or more existing and future failure modes" (ISO 13381-1, 2004). More generally, definitions of "prognostic" are grounded on the failure or fault notion, which implies that "prognostic" is associated with a degree of acceptability. Thus, it should be based on assessment criteria, whose limits depend on the system itself and on performance objectives. Consider Fig. 1: the predicted situation at time " $t+dt$ " can be considered as a critical one because of the degradation limit. Without this limit, there is no way to conclude on the predicted situation, and thus, it is impossible to dress an appropriate maintenance.

Finally, prognostic could be split into 2 sub-activities: a first one to predict the evolution of a situation at a given time (forecasting process), and a second one to assess this predicted situation with regards to a referential. In practical situation all this must be supported by operational tools (hardware and software technology devices). In this paper, the "assessment characteristic" of prognostic is leaved aside and the developments here after deal with "forecasting".

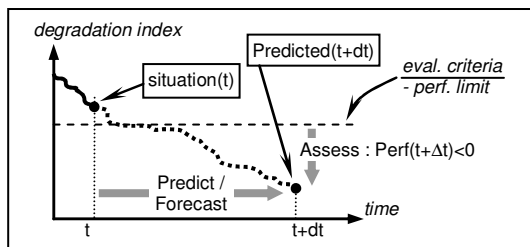


Fig. 1. Prognostic as a forecasting and assessment process

### 2.3 Prediction / forecasting approaches overview

According to some authors, the methods presented in this section are sometimes labeled as "prognostic techniques" (§ 2.2). However, most of them refer to what, in this paper, is called "prediction / forecasting".

Various techniques to prognostics have been developed. These methods can be classified as being associated with one of the following two approaches: model-based and data-driven (Byington, 2002, Chiang *et al.*, 2001).

*Model based approaches.* The model-based methods assume that an accurate mathematical model can be constructed from first principles. As an example, physics-based fatigue models have been extensively employed to represent the initiation and propagation of structural anomalies. The main advantage of these approaches is their ability to incorporate physical understanding of the system to be monitored. But this closed relation with a mathematical model may also be a strong weakness: it can be difficult even impossible to catch the system behavior (comprehensiveness, non linearity).

*Data driven approaches.* Data-driven approaches derive directly from routinely monitored system operating data. In many applications, measured input/output data is the major source for a deeper understanding of the system degradation behavior. According to the scientific literature, data-driven

approaches can be divided on two global categories: statistical techniques and artificial intelligent (AI) techniques. The strength of data-driven techniques is their ability to learn from examples and capture subtle relationships among the data even if the underlying relationships are unknown or hard to describe.

*Forecasting tool choice.* Real systems are complex and their behavior is often non linear, non stationary. These considerations make harder a modeling step, even impossible. Yet, a forecasting computational tool must deal with it, and, according to precedent remarks, artificial intelligence should be valorized to support it. In fact, AI techniques have been increasingly applied and have shown improved performances over conventional "prediction" approaches (Wang *et al.*, 2004, Yam *et al.*, 2001, Zhang *et al.*, 1998). More precisely, adaptive networks are well suitable for forecasting. This point is more widely discussed in next section.

## 3. ADAPTIVE NETWORKS: TOWARDS A NEURO-FUZZY SYSTEM FOR FORECASTING

### 3.1 Neural Networks – a fitted forecasting technique

Artificial neural networks (ANNs) are a special case of adaptive networks that have been extensively explored in literature because of the following aspects. ANNs can perform nonlinear modeling without a priori knowledge: they are able to learn complex relationships among "inputs and outputs". Moreover, from the computational point of view, ANNs are quick processes.

ANNs have two typical connection architectures: feedforward (like the multi layers perceptron MLP or the radial basis function network RBF) and recurrent networks (like the Elman architecture or the recurrent radial basis function network). Both have been employed in system behavior forecasting. Yet, there is no unique way of designing an accurate ANN for "prediction".

### 3.2 From ANNs to neuro-fuzzy systems for forecasting

The idea of using ANNs for forecasting is not new: in 1964, Hu used the Widrow's adaptive linear network to forecast the weather. However, due to the lack of a training algorithm at the time, the research was quite limited and ANNs were leaved aside. Since the 80<sup>th</sup>, research in the area is growing up (see also Fig. 2).

*First step: feedforward networks.* One of the first successful application of ANNs in forecasting is reported by Lapedes and Farber (1987) who designed a feedforward ANN that can accurately mimic a chaotic series (Zhang *et al.*, 1998). In general, feedforward ANNs (PMC, RBF) trained with the backpropagation algorithm (introduced at this time) have been found to perform better than classical autoregressive models for the trend prediction of non linear time series (Yam *et al.*, 2001).

*Second step: learning and parameterization improvement.* There are many factors that can affect the performance of

ANNs (number of inputs and outputs nodes, number of layers, activation functions, learning algorithm, training sample...). Thus, building a neural network forecaster is a non trivial task. Since the 90<sup>th</sup> many works are led to improve the accuracy of predictions while reducing time processing. The developments are based on the research of guidelines for the optimization of ANN's architecture and the development of new training algorithms (Hippert *et al.*, 2001).

*Third step: backward networks.* In order to explicitly take into account the time in forecasting tools, backward networks architectures were also developed and recurrent ANNs were compared with some of the well known methods for the prediction of non-linear time series. The results indicated that RNNs have a better forecasting performance than the classical methods and are even better than the feedforward type ANNs (Yam *et al.*, 2001).

*Fourth step: towards neuro-fuzzy systems.* ANNs have successfully been used to support the prediction process, and research works emphasize on the interest of using it. Nevertheless, some authors remain skeptical: 1) the design of an ANN is more of an art than a science, 2) ANNs are "black-boxes" which imply that there is no explicit form to explain and analyze the relationships between inputs and outputs. According to these considerations, recent works focus on the interest of hybrid systems: many investigations aim at overcoming the major ANNs drawback (lack of knowledge explanation) while preserving their learning capability. In this way, neuro-fuzzy systems are well adapted (next section).

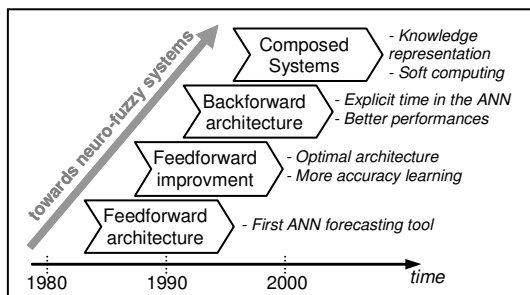


Fig. 2. From ANNs for forecasting to neuro-fuzzy systems

### 3.3 Neuro-fuzzy systems for forecasting

Both neural networks and fuzzy inference systems (FISs) are dynamic and parallel processing systems that estimate input output function. The interest of their integration follows from their complementary features. FIS uses linguistic rules for system behavior forecasting. It starts from highly formalized insights about the dynamic behavior of the system, and then formulates expert knowledge in fuzzy IF-THEN rules in order to ensure the forecasting task. However, FISs lack learning capability: sometimes, it's difficult to properly determine the fuzzy rules and to optimize the membership functions of premises. A solution to overcome these disadvantages is to use ANNs to train the fuzzy structure and the parameters. Moreover, the rules of neuro-fuzzy systems (NFSs) are transparent, enabling validation and manipulation of knowledge by experts (James *et al.*, 2003). At last,

compared to ANNs, NFSs seem to be a promising tool if the volume of available data is limited (Mahabir *et al.*, 2006).

Actual developments confirm the interest of using NFSs to forecast non linear time series (Wang *et al.*, 2004, Yam *et al.*, 2001, Zhang *et al.*, 1998): NFSs overpass others methods in both forecasting accuracy and training efficiency. As a result, an adequate tool to support the forecasting phase of the prognostic process should be an NFS. Work of (Wang *et al.*, 2004) demonstrates that the adaptive neuro-fuzzy inference system (ANFIS) is a very reliable and robust machine health condition predictor. Next part is devoted to it.

## 4. ANFIS: TOWARDS A MID-TERM RELIABLE FORECASTING SYSTEM

### 4.1 ANFIS: adaptive neuro-fuzzy inference system

*Principle.* ANFIS is a class of adaptive networks introduced by (Jang and Sun, 1995) and exploited recently to support forecasting tasks. It can be seen as a feed-forward neural network structure where each layer is a neuro-fuzzy system component. It simulates Sugeno fuzzy rule where the consequent part of the rule is a linear combination of input variables and a constant. The final output of the system is the weighted average of each rule's output.

*Architecture, learning and inferring process.* Consider Fig. 3 to describe the architecture of ANFIS and briefly explain the learning and inferring process, and see (Jang and Sun, 1995, Wang *et al.*, 2004) for further explications. The ANFIS network is composed of five layers excluding input layer. If two membership functions are assigned to each input variable (like in Fig. 3), then 16 rules are formulated for the inference operation ( $2^4$ ). These rules are represented in general form in the following manner:

$$R_j : \text{if } (x_1 \text{ is } A_1^j) \text{ and } (x_2 \text{ is } A_2^j) \text{ and } (x_3 \text{ is } A_3^j) \text{ and } (x_4 \text{ is } A_4^j) \\ \text{then } y_j = c_1^j x_1 + c_2^j x_2 + c_3^j x_3 + c_4^j x_4 + c_5^j \quad (1)$$

where  $j=1, \dots, 16$  is the rule number.

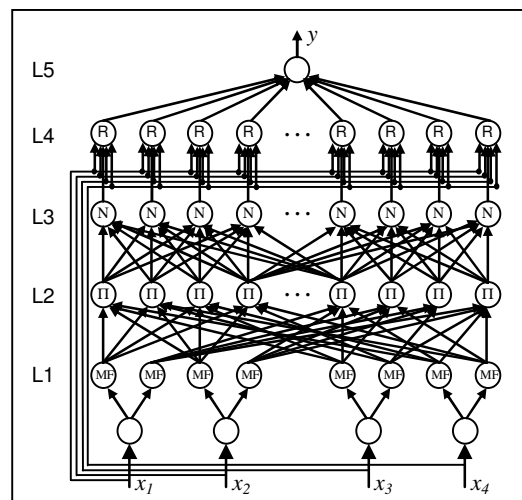


Fig. 3. Architecture of an ANFIS with 4 inputs

Layer 1 is the fuzzyfication layer in which each node represents a membership value to a linguistic term like Gaussian or sigmoid functions, etc. In this way, two Gaussian function memberships can be assigned to each input variable:

$$\mu_{A_l^j}(x) = \exp(-0.5[(x - m_l^j)/b_l^j]^2) \quad (2)$$

where  $l$  is the number of the input  $\{1, \dots, 4\}$  and  $\{m_l^j, b_l^j\}$  is the parameter set. Parameters in the layer are referred to as premise parameters.

In Layer 2, each node performs a fuzzy T-norm operation. Outputs of this layer are the rule firing strength. Max-product operator is generally used:

$$\mu_j = \mu_{A_1^j}(x_1) \cdot \mu_{A_2^j}(x_2) \cdot \mu_{A_3^j}(x_3) \cdot \mu_{A_4^j}(x_4) \quad (3)$$

In layer 3, all the rule firing strengths are normalized. After the linear combination of the input variables in Layer 4, the predicted output  $y$  is obtained in Layer 5 via weighted average (4).

$$y = \frac{\sum_{j=1}^{16} \mu_j (c_1^j x_1 + c_2^j x_2 + c_3^j x_3 + c_4^j x_4 + c_5^j)}{\sum_{j=1}^{16} \mu_j} \quad (4)$$

This above stated NF system has 96 parameters to be optimized (16 premise MF parameters and 80 consequent parameters) which is the aim of the learning phase. Hybrid learning algorithms are used to be deployed; especially the combination of the gradient descent and least squares estimate algorithms. In fact, works showed that such an approach reduces the complexity of the algorithm while increasing the efficiency of learning (Jang and Sun, 1995, Li and Cheng, 2007, Wang *et al.*, 2004).

*Performances and limits.* Different works (see § 3.3) show that ANFIS can capture the behavior of a system quickly and accurately, even better than other methods (included artificial neural networks). Moreover, contrary to an ANN, an ANFIS can be interpreted by an expert. That said, in practice, the modelling of an adapted ANFIS architecture run up against similar problems as for ANNs. In fact, the exactness of forecasting obtained through this tool depends on many factors: inputs selection, number and type of membership functions, learning algorithm. As an example, the same ANFIS as proposed here above but with 3 membership functions for each input (instead of 2) would have 431 parameters to be optimized (instead of 96). This obviously would be time expensive. More over, this would impair the transparency of the underlying model... Nevertheless, better results can be performed without increasing the complexity of the model and this is the purpose of next section.

#### 4.2 Inputs selection

*Classical inputs.* In forecasting applications, the inputs of the ANFIS are directly extracted from the time series data. As an

example, (Wang, *et al.*, 2004) construct a predictor with 4 inputs  $x_{t-3}, x_{t-2}, x_{t-1}, x_t$  to forecast the output  $x_{t+1}$ . However, this type of "input selection" doesn't enough exploit the specificity of ANFIS inference...

*Inputs selection.* ANFIS approximates the output variable by decomposing the whole inputs space into several partial fuzzy spaces and representing the output space with a linear equation (see (1) and (4)). There is a way to take unfair advantage of this characteristic to improve the ANFIS forecasting results. Consider Fig. 4 to argue this assumption. At a given time  $t$ , the measure  $x_{t+r}$  can be approximated as follows (linear approximations):

$$x_{t+r} = x_{t-3} + (x_{t-2} - x_{t-3}) + (x_{t-1} - x_{t-2}) + (x_t - x_{t-1}) + A_r \quad (5)$$

From this and considering the learning capability of ANFIS, the predicted output should be well approximated at any time by modifying the classical inputs data space partitioning and retaining the following set of inputs:  $\{x_{t-3}, x_{t-2}-x_{t-3}, x_{t-1}-x_{t-2}, x_t-x_{t-1}\}$ . This pre-treatment allows the taking into account of the dynamic of the signal, and the "learning" of its evolution instead of the data series.

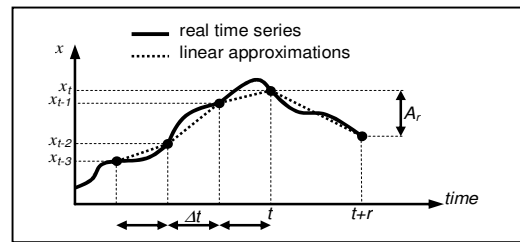


Fig. 4. Linear approximation of a data series with ANFIS

As with classical inputs, the choice of a membership function type (Gaussian...) aims at reaching a compromise between performance, number of parameters and computing time.

*Tests data.* The chaotic Mackey-Glass time series data is used to validate the effect of the proposed inputs selection. This time series is a benchmark problem extensively used: it's a non periodic and non convergent time series. Considering our final applicative objective (the prognostic of failures), to be capable to carry out predictions on such a signal is of good omen: real systems are complex and have generally a non-stationary and non-linear behaviour, what makes difficult a modelling phase. Working with this data series constitutes a first step to the specification of a prognostic system able to take into account the dynamic of real systems.

*Simulation.* The conditions of simulation are resumed in Table 1. In order to extract more solid conclusions from the comparison results, the same training and testing data sets were used to train and test both models (Table 2). Predictions were made at " $t+1$ ", at " $t+10$ ", and at " $t+50$ " in order to measure the stability of results in time. The prediction performance was assessed by using the root mean square error criterion (RMSE), the mean absolute percent error (MAPE) and the maximum of percent error (max\_PE).

**Table 1. Conditions of tests**

Inference system	
number of inputs .....	4
input partitioning .....	2 memberships per input
membership functions .....	Gaussian
inference type .....	linear – Sugeno
Learning scheme	
training/test data samples ...	500/500 and 250/750
training algorithms .....	LSE + Gradient Descent
training epoch number .....	10
error goal .....	0
initial step size .....	0,01
step size decrease rate .....	0,9
step size increase rate .....	1,1

4.3 Discussion

*Short-term predictions.* In short-term forecasting ( $t+1$ ) the ANFIS with the "selection of inputs" (ANFIS2) does not provide very higher results than the classical model (RMSE and MAPE are quite the same than with classical inputs). However, it is a little more accurate.

*Mid-term predictions.* With multi-step-ahead prediction ( $t+10$  and  $t+50$ ), the advantage of selecting the inputs is much obvious: the second ANFIS predicts in a more accurate manner. Both RMSE and MAPE are divided by 2.

*Error spreading.* The maximum of the percent error in forecasting is always smaller with the second model. This is all the more significant in the multi-step-ahead prediction (up to 2 times smaller). See Fig. 5 for an illustration. From the decisional point of view in maintenance applications, this aspect is very important: ANFIS2 enables to have a more robust forecasting tool as the variation of error is not as big as in the classical model. Thus, the confidence on the predictions is all the more greater.

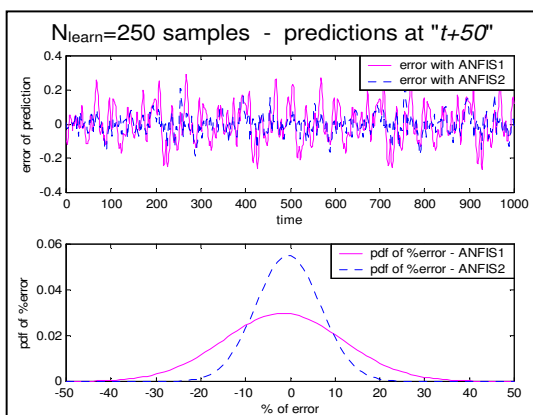


Fig. 5. Errors of predictions and pdf of percent errors

*Learning data set.* The ANFIS with "inputs selection" trained with 250 data samples overpass the performance of the classical model trained with 500 learning data samples: at " $t+50$ ", predictions of ANFIS2 with 250 learning samples are much more accurate and have a much smaller error spreading

than that of the classical ANFIS with 500 training samples. Moreover, these predictions are quite the same than that of the normal ANFIS at " $t+10$ " with  $N_{learn}=500$ . This is an encouraging result: classically, one says that the number of learning samples must be more than five times the number of parameters to be updated, which is often prohibitive in real applications. Yet, both models have 96 parameters.

In a word, the modification of the proposed inputs selection is an efficient way to improve the effectiveness of ANFIS predictions, notably in mid-term forecasting. Thus, this lets hope that it will be possible to build a prognostic system whose error of prediction would be mastered. The consequent maintenance decisions will be all the more better.

**Table 2. Simulation results**

$N_{learn} = 500$ $N_{test} = 500$		Classical inputs	With inputs selection
t+1	RMSE	0,0012	0,0012
	MAPE	0,09 %	0,09 %
	max_PE	8,34 %	7,26 %
t+10	RMSE	0,0512	0,0277
	MAPE	4,57 %	2,27 %
	max_PE	37,62 %	12,52 %
t+50	RMSE	0,1024	0,0529
	MAPE	9,79 %	5,8 %
	max_PE	57,62 %	30,88 %
$N_{learn} = 250$ $N_{test} = 750$		Classical inputs	With inputs selection
t+1	RMSE	0,0013	0,0012
	MAPE	0,11 %	0,10 %
	max_PE	8,15 %	7,63 %
t+10	RMSE	0,0549	0,0339
	MAPE	5,04 %	2,96 %
	max_PE	27,23 %	13,90 %
t+50	RMSE	0,1084	0,0581
	MAPE	10,26 %	5,05 %
	max_PE	59,19 %	32,31 %

4.4 The inside effect of inputs selection

The improvement of ANFIS predictions is not due to an increased complexity of the model, nor to a different learning algorithm. Also, the training and test data sets are rigorously the same in comparison results...

Consider two ANFIS: the first one (ANFIS1) with classical inputs  $x_1$  and  $x_2$  defined on  $[a, b]$ , and the second one (ANFIS2) with selected inputs  $x_1$  and  $x_1-x_2 \in [a-b, b-a]$ . Consider also two data samples for each ANFIS:  $\{X_1, X_2\}$  and  $\{X_1, X_2'\}$  for the first one, and  $\{X_1, X_1-X_2\}$  and  $\{X_1, X_1-X_2'\}$  for the second one, where  $X_2' \approx X_2$ . These two samples are very similar. However, ANFIS2 "considers" them as "more similar" than ANFIS1 (Fig. 6):  $\Delta\mu_{ANFIS2} < \Delta\mu_{ANFIS1}$ . Thus, during the learning phase, ANFIS2 has a lower correction of parameters. In other words, the inputs selection enables to have a compromise between approximation and generalization. This also explains that the learning phase can be performed with less learning samples.

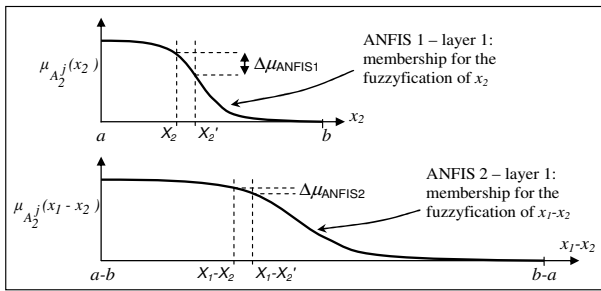


Fig. 6. Example of fuzzyfication in ANFIS1 and ANFIS2

## 5. CONCLUSION AND WORK IN PROGRESS

The high costs in maintaining complex equipments make necessary to enhance maintenance support systems. However, researchers are just beginning to address the concept of prognostic and real effective prognostics systems are scarce. In this context, the work reported here aims globally at drawing the way of implementing an efficient prognostic tool in industrial applications.

The concept of "prognostic" has been positioned within the maintenance strategies. This revealed that the accuracy of a prognostic system is related to its ability to approximate and predict the degradation of an equipment. Thus, a study of prediction approaches has been led in order to identify a potential technique. According to the global requirements that can be expected from a forecasting tool (accurate prediction with non linear data, knowledge representation, flexibility, computation time...) the ANFIS adaptive network has been chosen. An improvement of this technique has also been proposed. It enables the forecasting to be a more robust task, without increasing complexity of treatments. However, a deeper consideration of prognostic items should be taken into account (functioning modes, degradation laws...).

The work is still in progress and the developments are at present extended in two principal ways. First, the multi-step-ahead prediction is been more closely investigated, which is a critical task for real prognostic systems. Secondly, the predictive ANFIS tool is been modified in order to integrate multi-dimensional data (complex systems can't be supervised by a single monitoring index).

## REFERENCES

Byington, C., M. Roemer, G. Kacprzyński and T. Galie (2002). Prognostic Enhancements to diagnostic Systems for Improved Condition- based maintenance. In. *2002 IEEE Aerospace Conference*, Big Sky, USA.

Chiang, L.H., E. Russel and R. Braatz (2001). *Fault detection and diagnosis in industrial systems*. Springer-Verlag, London.

Ciarapica, F.E. and G. Giacchetta (2006). Managing the condition-based maintenance of a combined-cycle power plant: an approach using soft computing techniques. *Journal of Loss Prevention in the Process Industries*, **19**, 316-325.

De Gooijer, J.G. and R.J. Hyndman (2006). 25 years of time series forecasting. *International Journal of Forecasting*, **22**, 443-473.

Farrar, C.R., F. Hemez, G. Park, A.N. Robertson, H. Sohn and T.O. Williams (2003). A Coupled Approach to Developing Damage Prognosis Solutions. In: *Damage Assessment of Structures - The 5th International Conference on Damage Assessment of Structures (DAMAS 2003)*, Southampton, UK.

Hippert, H.S., C.E. Pedreira and R.C. Souza (2001). Neural Network for short term load forecasting: A review and evaluation. *IEEE Transaction on Power Systems*, **16**, 44-55.

ISO 13381-1 (2004). *Condition monitoring and diagnostics of machines – prognostics – Part1: General guidelines*. International Standard, ISO.

Iung, B., G. Morel and J.B. Leger (2003). Proactive maintenance strategy for harbour crane operation improvement. *Robotica*, **21**, 313-324.

James, W.F.C., D.A. Linkens, M.F. Abbod, M. Chen, J.L. Burton, K. Feeley and F.C. Hamdy (2003). Artificial Intelligence in predicting bladder cancer outcome: A comparison of Neuro-Fuzzy and Artificial Neural Networks. *Clinical Cancer Research*, **9**, 4172-4177.

Jang, J-S.R. and C-T. Sun (1995). Neuro-Fuzzy modeling and control. *Proceedings of the IEEE*, **83**, 378-406.

Jardine, A.K.S., D. Lin and D. Banjevic (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, **20**, 1483-1510.

Li, C. and K.H. Cheng (2007). Recurrent neuron-fuzzy hybrid-learning approach to accurate system modeling. *Fuzzy Sets and Systems*, **158**, 194-212.

Lin, D. and V. Makis (2003). Recursive filters for a partially observable system subject to random failure. *Advances in Applied Probability*, **35**, 207-227.

Mahabir, C., F. Hicks and F.R. Fayek (2006). Neuro-Fuzzy river ice breakup forecasting system. *Cold regions science and technology*, **46**, 100-112.

Muller, A., M.C. Suhner and B. Iung (2008). Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial sytem. *Reliability Engineering and System Safety*, **93**, 234-253.

Vachtsevanos, G., F.L. Lewis, M. Roemer, A. Hess and B. Wu (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. Wiley & Sons, Hoboken, New Jersey.

Wang, W.Q., M.F. Goldnaraghi and F. Ismail (2004). Prognosis of machine health condition using neuro-fuzzy systems. *Mechanical Systems and Signal Processing*, **18**, 813-831.

Yam, R.C.M., P.W. Tse, L. Li and P. Tu, (2001). Intelligent predictive decision support system for condition-based maintenance, *International Journal of Advanced Manufacturing Technology*, **17**, 383-391

Zhang, G., B.E. Patuwo and M.Y. Hu (1998). Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, **14**, 35-62.