Use of an ANFIS Network for Relative Humidity Behaviour Modelling on the South Region of Jalisco, México

Marco Gamboa[1], Marco Castro[2], and Francisco Herrera[3]

[1] Instituto Tecnologico de Ciudad Guzman, Departamento de Sistemas y Computacion, Av. Tecnologico #100 Cd. Guzman, Jalisco, Mexico
mgamboah@hotmail.com

[2] Instituto Tecnologico de la Paz, Departamento de Sistemas y Computacion, Boulevard Forjadores de Baja California 4720 La Paz, Baja California Sur, Mexico
ma_cl@hotmail.com

[3] Universidad Central "Marta Abreu" de las Villas, Facultad de Ingenieria Electrica, Carretera Camajuani S/N, Santa Clara, Cuba
herrera@fie.uclv.edu.cu

Abstract. This paper describes the application of an adaptive network using a neuro-fuzzy inference system called ANFIS for the identification of a relative humidity prognosis model on the south region of the state of Jalisco, México. The algorithm uses measurements obtained from meteorological stations as input-output signals and human expert's considerations on an if-then fuzzy knowledge base. The obtained model represents the system's dynamics. Figures showing the obtained model are presented.

1 Introduction

Currently, fuzzy modelling along with other techniques, such as neural networks, are being applied to identify systems not amenable to classical modelling techniques, due to the lack of precise, formal knowledge about them, their highly nonlinear behaviour, their degree of uncertainty, or their dynamic characteristics. One of the reasons for this is the capability of fuzzy systems to integrate information from different sources, such as physical laws, empirical models, or measurements and heuristics [1], and their ability to deal with qualitative information rather than solely relying on quantitative data.

Fuzzy modelling systems can be viewed as a special kind of Knowledge-Based Systems that uses a set of IF-THEN rules based on linguistic labels that represent the expert's knowledge about the system [2]. The part of the system responsible for interpreting the rules on the fuzzy knowledge base to, given the past input values, obtain the predicted output values is called an inference system. The proposed

© A. Gelbukh, R. Monroy. (Eds.)
Advances in Artificial Intelligence Theory
Research on Computing Science 16, 2005, pp. 75-84

model to determine relative humidity behaviour has its origins on the work by Takagi and Sugeno on fuzzy systems identification [3].

For a system with m inputs and one output on each instant of time t, it is possible to obtain a sample of the input and output data on that instant z(t) given by:

$$z(t) = [y(t)\ u_1(t)\ \ u_2(t)\ .\ .\ .\ \ u_m(t)]. \qquad (\mathbf{1})$$

The model's size and complexity depends on the amount of input and output variables employed, and it is called the Order of the model.

Assuming a model structure defined by:

$$y(t) = g(\varphi(t),\theta). \qquad (\mathbf{2})$$

Where y(t) represents the system's output on the instant t and the regressor vector φ(t) contains the ordered values of the input and output variables on t and on past instants (t-1, t-2….t-k.)

$$\varphi(t) = [\ y(t\text{-}1) \ldots y(t\text{-}k)\ \ u_1(t) \ldots u_1(t\text{-}k_1) \ldots \ldots \ldots u_m(t) \ldots u_m(t\text{-}k_m)\ ]. \qquad (\mathbf{3})$$

The fundamental hypothesis is the existence of an unknown computable predictive function *g(.)* that applied to a set of regressor vectors allows us to determine the value of the future system outputs. Parametric model estimation techniques have been used to find a g'(t) function that can approximate within a certain precision degree using the available data so that:

$$y(t) = g'(t) + \ \text{'}t. \qquad (\mathbf{4})$$

whereas v't represents the noises and effects produced by all those causal variables whether quantitative, qualitative, visible or invisible that where not included on the regressor set.

These estimation techniques are based on the assumption of a pre-established nonlinear form (Artificial Neural Networks and Fuzzy Rule Based Models) with excellent results. However, in prognosis work, the main focus of the scientific community has been on the use of connexionist systems (Artificial Neural Networks). And less attention has been paid to the exploration of adaptive fuzzy inference systems such as [4].

The present article describes the application of the ANFIS model over a series of meteorological data obtained on several stations located on the south portion of Jalisco State, and the way the identification algorithm is applied.

2 Structure of ANFIS

In classic set theory, a subset *S* on the universe *U* can be defined as a function that relates each element *x* of *U*, with an element of the discrete set {0, 1}[5].

$$S : U \bullet \{0, 1\}. \quad \textbf{(5)}$$

Where 0 is used to indicate than *x* does not belong to *S* at all and 1 indicates that *x* belongs completely to *S*.

On the other hand, in fuzzy set theory [6] μ is a membership function that can be valued on the continuum range [0, 1]:

$$:U \bullet [0, 1]. \quad \textbf{(6)}$$

where μ(u) represents the degree in which $u \in U$ belongs to the fuzzy set S. This is a generalization of the classical concept of a set (sharp set). μ is a set of ordered pair for discrete variables and a real function for continuous variables.

Certain families of functions are conventionally used to define the membership function, given their coincidence with the linguistic meaning of the most commonly used labels. Among the more frequent are:
The sigmoid function

$$S(u;\gamma,c) = \frac{1}{1+\exp[-\gamma(u-c)]}. \quad \textbf{(7)}$$

The Z function:

$$Z(u;\gamma,c) = 1-S(u;\gamma,c). \quad \textbf{(8)}$$

And the generalized bell function:

$$\Phi(u;\alpha,\beta,\gamma) = \frac{1}{1+\left|\frac{u-\gamma}{\alpha}\right|^{|2\beta|}}. \quad \textbf{(9)}$$

There are 3 main elements that constitute a fuzzy system: the input-output variables, the fuzzy set rule and the fuzzy inference system that defines the way the IF-THEN rules are used to obtain the output values given the current linguistic input variables values.

A neurofuzzy ANFIS network can be viewed as a fuzzy system conformed by Takagi-Sugeno rules [2], [7], [8]. On a system with two input variables x and y the ANFIS architecture can be represented as:

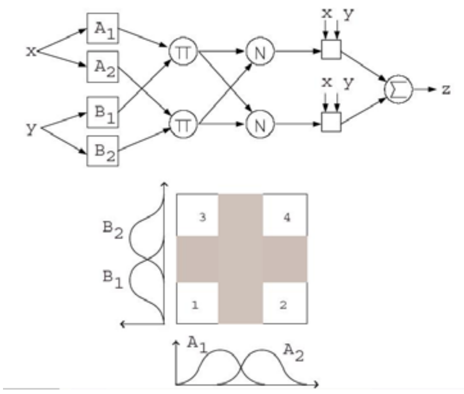

Fig 1: ANFIS architecture.

With a fuzzy rules base with the general structure:

if x is A1 and y is B1 then z = p1x + q1y + r1 **(10)**

if x is A1 and y is B2 then z = p2x + q2y + r2

if x is A2 and y is B1 then z = p3x + q3y + r3

if x is A2 and y is B2 then z = p4x + q4y + r4.

Where the x domain is delimited by fuzzy sets A1 and A2 and the y domain is delimited by fuzzy sets B1 and B2.

Four stages constitute the inference task: the first one is to calculate the membership functions µA1 (x), µA2(x), µB1 (y) y µB2 (y) using 9. The second task is to obtain the inference product wj = µA(x) · µB(y) for each rule. The third task is to obtain a weighted sum of each rule contribution to the outputs. And finally to evaluate the system's solution given by the following expression:

$$\sum_i \overline{w_i} f_i$$

In the rules base (10) the expert's knowledge among with other external considerations are represented in the form of fuzzy rules. It is to be noted that in 10 the antecedent part of the rule includes the same variables as the consequent; also, every variable domain is delimited by at least two fuzzy sets (Takagi-Sugeno form).

Although generally it is possible to use more complex functions in the consequent portion of the rules, usually a simple linear function is utilized to combine the entry variables.

The amount of fuzzy sets assigned to each independent variable, determines the ability of the ANFIS [4] algorithm to approximate non-linear functions and the algorithm's training technique to determine the parameters of the consequent functions [9]. However every time we increase the amount of fuzzy sets on the system, the amount of information required to train the network, and the computational cost of the process is considerably increased.

The fundamental aspect for the application of the ANFIS algorithm lies on the selection of the variables that will form the fuzzy rules, and the number of fuzzy sets assigned to each variable. The process of determination of an initial structure for the system has been investigated and data-mining techniques have been proposed by Jang [9].

3 The ANFIS Hybrid Training Method

Currently there are three main classes of fuzzy inference systems: Mandani, Sugeno and Tsukamoto[4] [11], which are fundamentally different on the structure of the consequent part of their IF-THEN rules, and because of this difference, they also differ from each other on their aggregation and defuzzification methods.

This paper centres on the Sugeno inference system. The ANFIS algorithm is the training routine that better adapts to this type of inference system. The learning rule is hybrid, given that it can apply the steepest descent method and back-propagation for non-linear parameters, and the least mean-squares methods on linear output parameters on the network.

The steepest descent consists on recursively obtaining a gradient vector on which each element is defined by the derived error function relative to a single parameter; this sole procedure can take a long while to converge. This series of iterations to find the gradient vector of an adaptive network structure is also called back-propagation because the gradient is calculated in the direction opposite to the output flow.

The combined use of both methods is effective for a fast identification of the adaptive network parameters. The iterations include a forward step and a backward step.

On the forward step the least mean square method is applied to calculate layer by layer, the output of each node on the net using the input signals vector, until a row corresponding to the matrix A and the vector Y, in the equation $A\Psi = Y$ is obtained. Where A is a matrix of m by n known functions of the input vector, Ψ the parameters to be esteemed vector of size n, and Y the size m output vector. This process is repeated for each pair of training data. The identification of the unknown parameter vector Ψ is obtained once the iterations for each row in A and Y are completed.

It is simple to determine that, when m=n and A isn't singular:

$$\Psi = A^{-1} y . \tag{11}$$

But, generally there are more input-output data pairs than parameters to adjust, thus m > n and a modelation error is generated so:

$$A\Psi + e = y . \tag{12}$$

and e is:

$$e = y - A\psi . \tag{13}$$

Our task consists in finding a value Ψ ($\hat{\Psi}$) that minimizes the sum of the square error defined by:

$$E(\Psi) = e^T e = (y - A\Psi)^T (y - A\Psi) . \tag{14}$$

Since E(Ψ) is in quadratic form, we have a unique minimum in ($\Psi \equiv \hat{\Psi}$) called the least mean squares estimator:

The back step uses back-propagation, where the error signals flow from the output nodes back to the input nodes. The gradient vector is calculated for each training data pair.

By the end of the processing of the training data, the non linear portion parameters are updated and the next forward step begins.

The proposed architecture uses five layers, the initial layer is conformed by the adaptive nodes that receive the input data. These nodes take that input data and apply the membership functions to each fuzzy set (high, low, tall,…etc.) to determine the degree to which they belong to those sets using a generalized bell function:

$$C1_i = \mu A_i(x) | i = \{1,2\} . \tag{15}$$

Where: $C1_i$ is the ith node output on layer 1, x is the node input and A_i is a linguistic label (low, medium, high, etc).
As the function parameters are changed, several membership functions are generated for each fuzzy set. Those parameters are called premise parameters.
The second layer is composed by fixed nodes also called "II nodes". The purpose of the second layer is to calculate the product of all the input signals ($C2_i$ = wi).

The importance of this layer is that each neurons output represents a percentile degree in which the rule is fulfilled.
The third layer calculates the fulfilment relation of the rule with the sum of fulfilment of all the rules. The output of this layer is a normalized this layers output (W_i^*) is called a normalized output.

$$\text{C3}_i = W_i^* = \frac{W_i}{W_1 + W_2} \quad \bigg| \; i = \{1,2\} \,. \tag{16}$$

Each node on the fourth layer is an adaptive node with a function of the type:

$$\text{C4}_i = W_i^* fi = wi(Pix + qiy + ri) \,. \tag{17}$$

Where pi, qi and ri along with the set of parameters of those nodes, are called consequent parameters.

The fifth and final layer has an only node named $\sum$ that has as an output the sum of all its input signals.

$$\text{C5}_i = \sum_i W_i^* f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \,. \tag{18}$$

The forward step starts with the premise parameters and calculates the output of the nodes on the fourth layer and the consequent parameters of that layer are identified by the MCL method.

In the ANFIS architecture the values of the premise parameters are fixed, thus, the total output can be expressed as a linear combination of the consequent parameters.

This way, the hybrid training method guarantees that the consequent parameters are optimum under the fixed premise parameter condition. The hybrid training method converges a lot faster than the original back-propagation method because it reduces the search space. As a rule, the membership functions must be fixed while the process of training is taking place.

4 System Identification

The rules on a fuzzy system are initially determined by a human expert that uses his knowledge about the system for this task. However, when there is no expert available for this purpose, the number of membership function of each variable is assigned empirically. In such case the expert's task is limited to a validation of the initial considerations inferred by the algorithm, as well as the determination of the samples sizes used to train and to validate the model. It is also the expert's task to assess the quality of the identification.

The relative humidity readings on the south region of Jalisco, México, are used by the farmers to determine which products to seed, however these readings are obtained from the report generated by the Guadalajara City Airport, located 90 km away, thus the data has little reliability for the farming zone.

A meteorological station located a half kilometre of Ciudad Guzman (in the south central region of Jalisco) is equipped with sensors to monitor some weather parameters, among them, the temperature, radiation level and relative humidity. Those readings are stored in a database every 15 minutes.

A set of readings containing 34,945 measurements were obtained. Due to the periodicity of the data, 342 daily averages are obtained to form the training matrix (251) and the validation matrix (92). The variables included in the model were: Temperature (t-1), radiation (t-1), radiation (t), relative humidity (t-2), relative humidity (t-1) and finally, the relative humidity (t) as the output variable.

To establish the initial SID a structure of 10 membership functions, with the least possible amount of rules was tried, but with this structure, the identification presented unacceptable errors, due to the big amount of state variables with an influence on the system's behaviour that couldn't be modelled with a low number of membership functions.

The number of membership functions was increased to 20 with the generalized bell structure, and using the training data, a MatLab program using the Fuzzy ToolBox was developed. The initial phase of the program uses the genparam subroutine to automatically generate initial arbitrary values for the membership functions. Then the genfisl routine was used to make a grid fractioning of the input data. Out of this process the initial Sugeno fuzzy inference system is obtained. Once the initial membership functions are obtained, a 200 epochs training using the input-output data is performed through the Anfis subroutine. The results of this process are the modified fuzzy sets and the error plot.

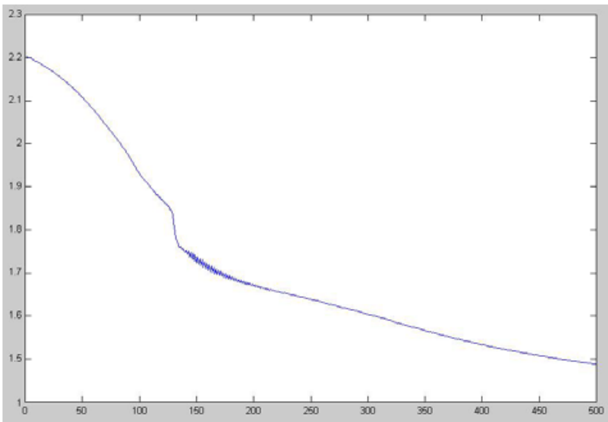

Figure 2

Still, the obtained error does not satisfy our criteria, thus a new training process is initiated using 400 epochs. Figure 2 shows that by the 130th iteration, the error is above 1.8, when the criterion is 1.5, but by the 400th iteration, the obtained value is 1.4. Since the grid partitioning defines that the number of rules is identical to the number of membership functions, we can obtain the 20 rules that identify the system with the showrule subroutine. On figure 3 the plotting of the identification results shows a total coincidence on the first 4000 chosen cases between the real process output (blue line) and the ANFIS identification output (green line).

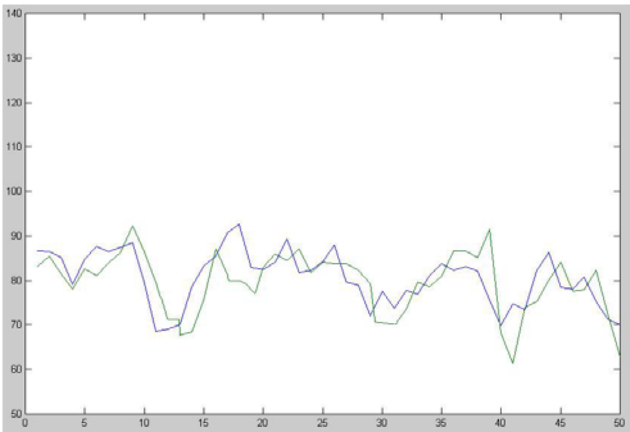

Figure 3

5 Conclusions

In this paper, the use of ANFIS identification with Sugeno fuzzy rules was discussed. Satisfactory results were obtained with the applied methodology to meet the needs of the farmers on the southern region of Jalisco. The results obtained are clearly better than the ones produced by the current system.

The advantage of this identification method lies on the fact that we need only a set of input/output measurements as initial data to identify non-linear systems. In the literature there are other examples of non-linear identification applications of ANFIS where a superior performance than classical methods and non hybrid artificial neural networks is obtained.

When looking to improve the identification structure it is convenient to pay attention to the effective fractioning (clustering) of the input space, this way we can reduce the number of rules and increase the speed of the identification process.

Thanks to their great adaptation possibilities, the generalized bell membership functions were successfully applied to very different input-output parameters.

A setback of this method is that when we encounter highly non linear systems, their identification demands an increased number of membership functions that make its on-line application very difficult due to the increased processing time needed to calculate the resulting parameters. On this case 20 membership functions where used and a time of 17 minutes of processing on a Pentium 3, were required to satisfy the error criteria.

References

1. Babuska, R.: Fuzzy Modeling for Control. 1 ed. International Series in Intelligence Technologies, ed. H. –J. Zimmerman. 1998, Boston: Kluwer Academic Publishers. 260
2. Wang, L. X.: Fuzzy System and Universal approximators. In: Poc. IEEE Int Conf. Fuzzy Systems, San Diego, Ca. Mar (1992)
3. Takagui T., Sugeno M.: Fuzzy Identification of systems and its applications to modeling and control". IEEE Trans. Syst., Man, Cybern, Vol. 15, 116-132 (1985)
4. Jang, J-S R. Anfis: Adaptative-network-based fuzzy Inference system. In: IEEE Trans. Syst., Man, Cybern, Vol. 23, 665-684 (1993)
5. Cox, E.: The fuzzy system Handbook, Second edition AP Professional, Paperback, 2nd Bk&Cd Edition, (1998)
6. Zadeth, L. A.: Fuzzy Sets. Information and Control, Vol. 8, 338-53 (1965)
7. Sugeno M., Kang T.: Structure identification on fuzzy model. Fuzzy sets and systems, Vol 28, 15-33 (1988)
8. Takagi, T., Sugeno, M. Derivation of fuzzy control rules from human operator's control actions. In Proceeding of IFAC Symposium on Fuzzy Information, Knoledge Reservation and Decision Analysis. Pg. 55-60 (1983)
9. Jang, J-S R.Input selection for ANFIS learning. In: Procededing of IEEE Int. conf. On Fuzzy Systems (1996)
10. Jang, J-S., Sun, C. T., Mizutani, E. : Neurofuzzy and Soft Computing. Math Lab Curriculum series. Prentice Hall (1997)