EFFICIENT FUZZY IDENTIFICATION BASED ON INFERENCE ERROR [1]

Juan Contreras Montes [1,3], Roger Misa Llorca [2], Luis Murillo Fernández [3]
Justo Sarabia Agamez [3], Juan Pablo Paz Grau [3]

[1]Escuela Naval Almirante Padilla
Facultad de Ingeniería Naval, Cartagena, Colombia
E-mail: epcontrerasj@ieee.org
[2]Instituto Superior Politécnico José Antonio Echeverría
Facultad de Ingeniería Eléctrica, La Habana, Cuba
E-mail: rmisa@electrica.cujae.edu.cu
[3]Corporación Universitaria Rafael Núñez, Facultad de Ingeniería
Programa de Ingeniería de Sistemas, Cartagena, Colombia
E-mail: {justo.sarabia, luisfernandom, juanpablop}@curn.edu.co

Abstract: We present an efficient methodology to obtain linguistically interpretable fuzzy models using the inference error method. It includes the algorithms to determine classes, rules generation, and partition sum-1 of input variables: shape, number and distribution of fuzzy sets. The load centre of each class represents the output variable's singleton location. The novelty of our proposal lies on the equilibrium between precision and interpretability, and its low computational complexity. The algorithm was applied to some benchmark classics like the Box-Jenkins' gas furnace, the Mackey-Glass series and others, getting better results that achieved by others, comparing the MSE, parameter number and interpretability.

Keywords: Fuzzy logic, Identification algorithms, Mean square error, Membership functions, Fuzzy models.

1. INTRODUCTION

Fuzzy models have got a great success because of the similitude in which they describe real systems the same way a human operator would understand them. The construction of fuzzy models implies the selection and tuning of many parameters like: shape and distribution of input variables membership functions, rule base, logic operators used, shape and distribution of consequents, etc. The great number of parameters to obtain a fuzzy model has hindered the development of a unique modeling technique, especially in the case of fuzzy identification from experimental input – output data.

Many investigators have fall back on hybrid techniques to achieve an acceptable precision but sacrificing the model interpretability especially because they generate a great number of rules, a great number of tuning parameters, and interpolations of more than two sets or in points where the membership degree sum is much greater than 1.

One of the first proposals to automatically design fuzzy systems from data was the table look-up écheme (Wang and Mendel, 1992). Sugeno and Yasukawa (1993) proposed a methodology to identify fuzzy models parameter identification using singleton consequents, but it requires many rules and offers a poor description capacity. Another

[1] This investigation is been founded by Colciencias, Cod 14191417481. Programa ETI

important methods use the descendant gradient, clustering techniques and evolutive algorithms. Except for the latest, all of them require a previous design of the membership functions.

Fuzzy clustering algorithms represent the most adequate techniques to obtain fuzzy models, being the fuzzy C-Means (Bezdek, 1987) and Gustafson-Kessel (Gustafson and Kessel, 1979) methods the most employed. Diverse variations have been applied to these clustering algorithms. Nauck and Kruse, (1995, 1999) proposed neuro-fuzzy clustering algorithms; Espinosa and Vanderwalle (2000) presented a methodology to extract rules from data in a framework of linguistic integrity including fusion algorithms to group sets whose modal values are at a very close distance. Sala (1998, 2001) introduced a novel technique based on the inference error to approximate functions using partition sum-1 with triangular sets; Diez, *et al*. (2004) proposed variations to the clustering algorithms to improve interpretability and find related local structures on the obtained fuzzy models.

Paiva and Dourado (2004) presented a model generated from training a neufo-fuzzy network implemented in two phases: at the first phase, the model structure is obtained using a substractive clustering algorithm, to extract rules from input – output data; at the second phase, the parameter tuning is performed by a neural network using back propagation.

Guillaume and Charnomordic (2004) proposed a strategy to generate fuzzy interpretable partitions from data with a methodology, which they call hierarchical fuzzy partitioning (HFP), where despite of adding data at each iteration, they add fuzzy sets. They also present a fusion algorithm of fuzzy sets based on adequate metrics that guarantee semantic interpretability. Chen and Saif (2006) proposed a novel fuzzy system that uses dynamic rules base, which imply that rules can change with the inputs, making it usable for modeling as long as for control.

The methodology used in this paper to obtain the fuzzy model from input – output data is based on the inference error method of (Sala, 1998) and is presented in three phases: On the former, the inference error method is used to generate an interpretable fuzzy model and, also, to detect possible classes of clusters on data; on the next, by using minimum squares to adjust the consequents and the descendant gradient to adjust the membership functions of the antecedent, the model is optimized; and in the latter, the method is used to construct control adaptive systems.

2. INFERENCE ERROR

On fuzzy logic is common to have rules like “if *u* is A, then *y* is B”, where u and y represents two numeric variables, and A $\subset$ U and B $\subset$ Y, are two input and output fuzzy sets respectively, defined on the universes U and Y. The mentioned rule is equivalent to the inequation

$$u_A(\mathrm{u}) \le u_B(\mathrm{y}) \qquad (1)$$

The inference error e, conceived as the conceptual distance to the conclusion set of the rule, is

$$\in \approx \begin{cases} 0 & \cdots \quad u_A(\mathrm{u}) \le u_B(y) \\ u_A(\mathrm{u}) - u_B(y) & \ldots \quad u_A(\mathrm{u}) > u_B(y) \end{cases} \qquad (2)$$

The proposed method to create a fuzzy system based on inference error that approximates a one input one output function, with null inference error should comply with the condition

$$u_A(\mathrm{u}) = u_B(\mathrm{y}) \qquad (3)$$

In the case of a rule of the kind “if *u* is A, then *y* is B”. If the system has *n* inputs, it should be represented by a rule of the kind “if u_1 is A_1, u_2 is A_2, …, u_m is A_m then *y* is B”, and the generated system should comply with the condition

$$((u_{A1}(x_k) \wedge (u_{A2}(x_k) \wedge \ldots \wedge (u_{Am}(x_k)) = u_B(y_k) \qquad (4)$$

Where $\wedge$ represent the combination operator.

3. FUZZY IDENTIFICATION BY INFERENCE ERROR

3.1 Fuzzy model structure

Searching for equilibrium between numeric approximation and the resulting fuzzy system interpretability the following criteria are considered for the parameter selection procedure.

Membership functions. The input variables universe partition in the learning process will be made with triangular normalized sets using a 0.5 specific overlapping. Triangular membership functions allow the reconstruction of the linguistic value in the same numeric value, after applying a defuzzyficator method (Gaweda and Zurada, 2003); also, the 0.5 overlapping assures that the fuzzy sets supports are different. The fuzzy sets generated for the output variable will be singletons.

Membership functions distribution. The input variables triangular fuzzy sets will be distributed symmetrically on each respective universe, so each of the elements on a variable universe should be a member of, at least, one fuzzy set.

Operators. To combine the rules antecedents the OWA operators will be used.

Inference method. The inference method used will be given by

$$f\left(x^{(i)}\right)=\frac{\sum_{j=1}^{L}\overline{y}^{j}\, m_j\left(x^{(i)}\right)}{\sum_{j=1}^{L} m_j\left(x^{(i)}\right)} \qquad (5)$$

Where:

$$m_j\left(x^{(i)}\right)=u_{A_1^j}\left(x_1^{(i)}\right).u_{A_2^j}\left(x_2^{(i)}\right).\ldots.u_{A_n^j}\left(x_n^{(i)}\right) \qquad (6)$$

Is the j^{th} rule output degree of a Mamdani fuzzy system, $\overline{y}^{j}$ is the singleton value that corresponds to the j rule. The denominator is always 1 when working with partitions sum-1, which is the case of the approximation method proposed in this article.

3.2 Fuzzy identification algorithm

The interpretable fuzzy system generator from data algorithm is based on the minimization of the inference error. It only requires the previous definition of the membership functions shape that will make up the antecedent partition and calculates the final number and distribution of the antecedent membership functions, guaranteeing partition sum-1, and the consequent position, of type singleton.
Given a collection of input and output experimental data $\{x_k, y_k\}$, $k = 1, \ldots, N$, where x_k is the p - dimensional input array $x_k^1, x_k^2, \ldots, x_k^p$ and y_k is a one - dimensional output array.

1. Organize the N pairs set of input – output data $\{(x_i, y_i) \mid i = 1,2,\ldots, N\}$ where $x_i \in \Re^p$ are input arrays and y_i are output scalars.

2. Determine the universe's ranges of each variable according to its maximum and minimum associated data values $\left[x_i^-, x_i^+\right]$, $\left[y^-, y^+\right]$.

3. Distribute the triangular membership functions over each universe. As a general condition, the vortex with membership value one (modal value) lies in the centre of the region covered by the membership function, while the other two vortexes, with membership value zero, lie in the center of the two neighboring regions. To efficiently approximate the inferior and superior sides of the data represented function is necessary that in the triangular partition of the membership function that covers the beginning and the ending of the universe, coincides its vortexes with membership value one with its left and right vortexes, respectively, as shown in figure 1.

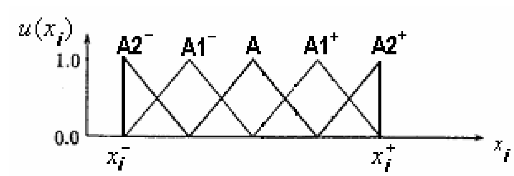

Fig. 1. Triangular partition sum 1

4. Calculate the modal value position at the input variables modal values, according to:

   if $u_{A_k^{(n)}}\left(x_k^{(i)}\right)=1$

   $$ys_k^{(n)} = y[i] \qquad (7)$$

   end

   Where $ys_k^{(n)}$ corresponds to the projection over the output space of the $x^{(i)}$ data evaluation of the k^{th} input variable on the n^{th} set of the corresponding partition. The output value that corresponds to this projection is given by the value of the i^{th} position of the output array y.

5. Rule determination. For this, a fuzzy hierarchical algorithm is used starting with a number of groups equal to the number of fuzzy sets for each input variable multiplied by the number of variables; in other words: n xk. The number of resulting groups will be the number of rules L, and the value of each grouping will be the singleton $\overline{y}^{j}$ value associated to the respective rule.

6. Model validation, or approximation calculation, using the inference method described by (5), where $\overline{y}^j$ is the singleton value that corresponds to rule j. The denominator is always equal to 1 when working with partitions sum-1, which is the case on the proposed approximation method.

7. Tune parameters, relocating the output singletons, using the minimum squares method. Equation (5) can be expressed as

$$f\left(x^{(i)}\right)=\sum_{j=1}^{L}\overline{y}^{j}\,w_{j}\left(x^{(i)}\right) \quad (8)$$

Where

$$w_{j}\left(x^{(i)}\right)=\frac{m_{j}\left(x^{(i)}\right)}{\sum_{j=1}^{L}m_{j}\left(x^{(i)}\right)}=w_{j}^{i} \quad (9)$$

And the output values can be represented as $Y = Wq + E$, that rearranged as a matrix can be expressed like

$$\underbrace{\begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^L \end{bmatrix}}_{Y}=\underbrace{\begin{bmatrix} w_1^1 & w_2^1 & \cdots & w_L^1 \\ w_1^2 & w_2^2 & \cdots & w_L^2 \\ \vdots & \vdots & \ddots & \vdots \\ w_1^n & w_2^n & \cdots & w_L^n \end{bmatrix}}_{W}\underbrace{\begin{bmatrix} \overline{y}^1 \\ \overline{y}^2 \\ \vdots \\ \overline{y}^L \end{bmatrix}}_{q}+\underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}}_{E} \quad (10)$$

Being E the approximation error which should be minimized. Using the quadratic error norm, we get

$$E^2=(Y-Wq)^2=(Y^2-2YWq+(Wq)^2) \quad (11)$$

The solution to this minimum square problem is given by

$$\frac{\partial E^2}{\partial q}=0=-2YW+2W^TWq \quad (12)$$

Where we obtain

$$q=\frac{Y^TW}{W^TW}=(W^TW)^{-1}Y^TW \quad (13)$$

This solution is only valid if (W^TW) is not singular; otherwise the minimum squares estimation should be performed recursively.

8. Finish if the quadratic medium error is less than a previously defined value. In other case, increment the number n of input variable sets in 1 and get back to step 3.

With the described algorithm an interpretable fuzzy model with accurate precision is obtained, and it only requires the adjustment of the singleton consequent parameters, which diminishes the training time. It is possible to achieve a better approximation ("*fine tuning*") if the descendant gradient method is applied at the end of the process to adjust the location of the modal values at the triangular fuzzy sets of the antecedent, keeping the partition sum-1 and, at the same time, system interpretability.

The proposed method allows the finding of a fuzzy interpretable system of rules with great precision, as will be seen forward when we present the results achieved resolving some classical problems. The projection of the modal values in the output space lets the detection of groupings or classes, which can help reduce the number of rules and the number of sets in the input universe.

4. RESULTS

Next, we present the results achieved by applying the proposed method to classic problems.

4.1 Box-Jenkins' Gas Furnace

One of the classic problems in system modeling and identification is the *gas furnace* proposed by Box and Jenkins (1976). The dataset is composed of 296 pairs of input-output data. Input data corresponds to the gas flow rate that is going to be burned and the output data to the carbon dioxide concentration. The objective is to predict the output using past input and output values.

Many authors have work on this problem with different number of input and output past values. Gaweda and Zurada (1998) use, at the regression equation, variables u(k-1), u(k-2), u(k-3), and y(k-2) and y(k-3), while Pavia and Dourado (2004) use at the regression equation just the variables y(t-1) and u(t-4), achieving a RMS error of 0.390 at the identification stage.

With the variables proposed by Pavia and Dourado (2004) we applied the proposed method and obtained the partition showed in figure 2 and the results shown in figure 3. The consequent singletons were adjusted with minimum squares at the training process and then the modal values of the fuzzy sets were adjusted using descendant gradient.

Table 1 presents the fuzzy model rule base and Table 2 compares our results with those obtained with other methods.

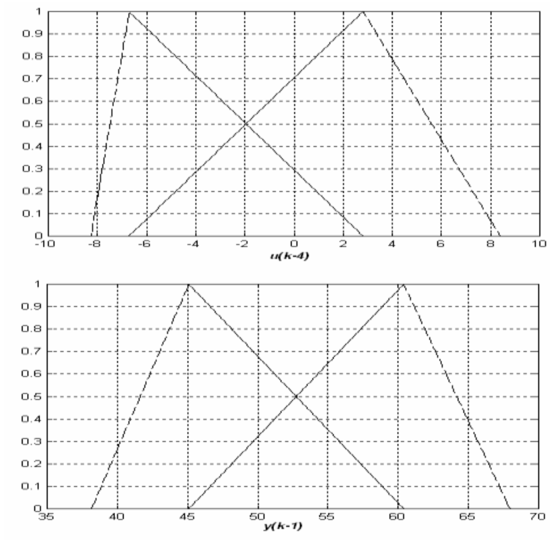

Fig.2. Membership functions for the modelling example of the Box-Jenkins furnace

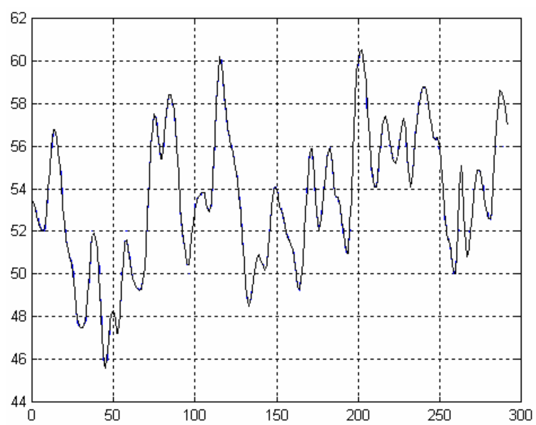

Fig. 3. Performance of the Box-Jenkins gas furnace fuzzy model

The output sets are of type singleton with locations on 44.78 and 65.3.

Table 1. Linguistic description of the fuzzy model

Rule	**u(k-4)**	Ù	**y(k-1)**	⇒ y(k)
1	A		B	44.78
2	B		A	65.30

The obtained fuzzy model has a great performance with a small number of parameters and an adequate interpretability, so a human operator will have a clear understanding of the rules.

Table 2. Comparing results of various methods to the Box-Jenkins gas furnace problem

Model	**MSE**	**Rules**	**Parameters**
Kim et al SI Model (1998)	0.048	2	110
Gaweda and Zurada Model (2003)	0.055	2	38
This Model	0.066	2	6
Wang and Langari (1995) Sugeno Model	0.066	2	N/A
Box-Jenkins (1976) ARMA Model	0.202	N/A	N/A

4.2 Mackey-Glass' chaotic series

Mackey-Glass time series is one of the most employed functions to test identification methods, and is described by the equation

$$\bar{x}(t) = \frac{0.2x(t-t)}{1+x^{10}(t-t)} - 0.1x(t) \qquad (14)$$

It's assumed that x(t) = 0 for t <0, a 0.1 time interval, x(0) = 1.2 and t = 12. x(t-18), x(t-12), x(t-6) and x(t) will be used to predict x(t+6). 500 instances where used to train and test. The obtained fuzzy model has three triangular fuzzy sets for each variable and eight rules. The number of tuning parameters is 20, a very small number compared with other methods. Figure 4 shows the results obtained.

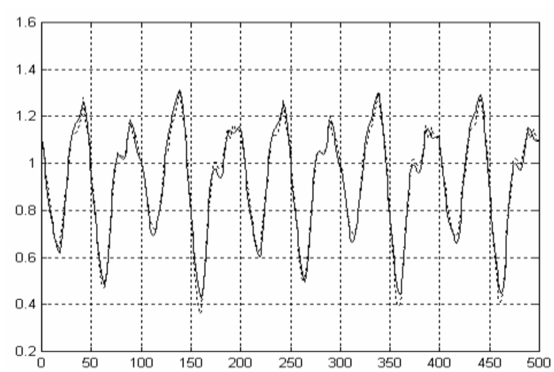

Fig. 4. Mackey-Glass series identification

The medium square error (RMSE) obtained was 0.0223 at training time and 0.0229 at validation time.

5. CONCLUSIONS

A method based on the inference error minimization was presented to identify systems from data with interpretable fuzzy systems with consequent parameters tuning, of singleton type, using square minimums and adjustment of membership functions of the antecedent with descendant gradient.

The fuzzy identification method shown presents an accurate precision without sacrificing interpretability, as occurs with most of the available methods.

The proposed method can easily be extended to adaptive control methods using the inverse control scheme and consequent parameters tuning.

REFERENCES

Bezdek J. C. (1987). *Pattern recognition with Fuzzy Objective Function Algortithms*. Ed. Plenum Press.

Box, G., Jenkins, G., (1976) *Times Series Analysis, Forescasting and Control*, 2nd ed. Holden Day, San Francisco.

Chen, W., and Saif, M. (2005) "A Novel Fuzzy System With Dynamic Rule Base", *IEEE Trans. Fuzzy Systems*, vol. 13, No. 5, pp. 569-582.

Díez J. L., Navarro J. L., Sala A. (2004). Algoritmos de Agrupamiento en la Identificación de Modelos Borrosos. *RIAI: Revista Iberoamericana de Automática e Informática Industrial*.

Emami, M., Turksen, I. Goldenberg A., (1998) "Development of a systematic methodology of fuzzy logic modeling. *Transaction of Fuzzy Systems*, vol. 6, No. 3, pp. 346-360.

Espinosa, J., Vandewalle, J. (2000) "Constructing fuzzy models with linguistic integrity form numerical data-afreli algorithm", *IEEE Trans. Fuzzy Systems*, vol. 8, pp. 591 – 600.

Gaweda, A., Zurada, J. (2003) "Data-Driven Lingustic Modeling Using Relational Fuzzy Rules", *IEEE Trans. Fuzzy System*, vol 11, pp. 121-134.

Guillaume, S., Carnomordic, B., (2004) "Generating an interpretable Family of Fuzzy Partitions Form Data", *IEEE Trans. Fuzzy Systems*, vol. 12, No. 3, pp. 324 – 335.

Guztafson E. E., Kessel W. C. (1979) Fuzzy Clustering with a Fuzzy Covariance Matrix. *IEEE CDC*, San Diego, California, pp. 503 – 516.

Kim, E., Park, M, Kim, S., (1998) "A transformed input-domain approach to fuzzy modeling", *IEEE Trans. Fuzzy Systems*, vol 6, pp. 596-604

Nauck, D., Kruse, R., (1995) "Nefclass - a neuro-fuzzy approach for the classification of data", In *Proceedings of the Symposium on Applied Computing*.

Nauck, D., Kruse, R. (1999) "Neuro-fuzzy systems for function approximation". *Fuzzy Sets and System*. 101(2), pp. 261-271.

Paiva, R. P., Dourado, A. (2004) "Interpretability and learning in neuro-fuzzy systems", *Fuzzy Sets and System*. 147(2004), pp. 17-38.

Pedriycz, W. (1994) "Why Triangular Membership Functions?" *IEEE Trans. Fuzzy Sets and System*, vol. 64, pp. 21-30.

Sala, A. (1998). Validación y Aproximación Funcional en Sistemas de Control Basados en Lógica Borrosa. *Universidad Politécnica de Valencia. Tesis Doctoral*.

Sala, A., Albertos, P. (2001) Inference error minimisation: fuzzy modelling of ambiguous functions. *Fuzzy Sets and Systems*, 121 (2001) pp. 95 – 111.

Sugeno, M., Yasukawa, T. (1993) "A fuzzy logic based approach to qualitative modeling". *Transactions on Fuzzy Systems*, vol. 1, No. 1, pp. 7-31.

Wang, L-X, Mendel, J.M. (1992) "Generating fuzzy rules by learning form examples", *IEEE Transactions on Fuzzy Systems*.

Wang, L, Langari, R. (1995), "Building sugeno type models using fuzzy discretization and orthogonal parameters estimation techniques", *IEEE Transactions Fuzzy Systems*, vol. 3, pp 454-458.