An Approach for Dynamical Adaptive Fuzzy Modeling

M. Cerrada[1], J. Aguilar[2], E. Colina[3], A. Titli[4]

Universidad de Los Andes, Facultad de Ingeniería

[1,3]Departamento de Sistemas de Control. CEMISID. e-mail: cerradam,colina@ing.ula.ve

[2]Departamento de Computación. CEMISID, e-mail: aguilar@ing.ula.ve

Avenida Tulio Febres Cordero.

Telf: 58-74-402983 Fax: 58-74-402846

Mérida 5101. VENEZUELA

[4]Laboratoire d'Analyse et d'Architecture de Systèmes (LAAS-CNRS)

Groupe Automatique Symbolique, e-mail: titli@laas.fr

7, avenue du Colonel Roche-31077.

Phone: 33-5-61336200 Fax: 33-5-61553577

Toulouse, Cedex. FRANCE

Abstract— **In this work, an approach for the development of adaptive fuzzy models is presented. The approach allows to incorporate the system dynamics into the fuzzy membership functions which are defined in terms of a dynamic function with adjustable parameters. These parameters are adapted using a gradient descent based algorithm. Some application examples to illustrate the performance of the dynamical adaptive fuzzy models on system identification are presented.**

I. INTRODUCTION

An important aspect about fuzzy modeling is the search for the design methods to develop accurate representation models of real process using the available knowledge about them. A rather classic method for the development of fuzzy models is the so called direct procedure which does not allow the incorporation of quantitative observations about the system operation in order to determine the structure and parameters of the model. Also, if there is a poor expert knowledge, the fuzzy model obtained from such a background will have an inadequate performance. In order to improve the development of fuzzy models, new design methods like the associate with the adaptive fuzzy systems design, allows the incorporation of the available data [1]. Other approaches, like the ones based upon artificial neural networks, have provided supervised learning algorithms to adapt parameters of fuzzy systems . The resulting fuzzy models have both the advantages of neural networks and fuzzy logic systems: they are universal approximators, they can learn through different methods, and the knowledge about the process may be incorporated into the model parameters.

In this work, a new approach for the development of adaptive fuzzy models is presented. In this approach the system variables dynamics may be incorporated into fuzzy membership functions of the proposed model. As a result, the fuzzy model incorporates dynamical membership functions whose parameters are adjusted via descent gradient learning algorithm.

In the following section, basic concepts about linguistic models and adaptive fuzzy models are revised and an analytic description of these model classes is presented. The third section includes the new dynamical adaptive fuzzy model, followed by illustrative examples about the construction of identification models in section four. Section five is devoted to conclusions and recommendations.

II. LINGUISTIC AND ADAPTIVE FUZZY MODELS

Without loss of generality, a linguistic fuzzy logic model described by a base of M fuzzy rules may be given by the following generic rule:

$$R^{(l)}: \quad IF\, x_1\ is\ F_1^l\ AND ...\ AND\ x_n\ is\ F_n^l \quad THEN\ y\ is\ G^l \tag{1}$$

where $\underline{X} = (x_1\ x_2\ ...\ x_n)^T$ is a vector of input linguistic variables x_i defined on an universe of discourse U_i. The output linguistic variable y is defined on an universe of discourse V. On the other hand, F_i^l and G^l are fuzzy sets on U_i and V, respectively, $(i = 1, ..., n)$, $(l = 1, ..., M)$.

The analytic expression that summarizes the inference mechanism the fuzzy logic system (FLS) described in (1), using the fuzzification method of ordinary sets, the center-average defuzzification method and gaussian membership functions for the fuzzy sets associated to the input variables, is given by the

0-7803-7280-8/02/$10.00 ©2002 IEEE

following equation [1]:

$$y(\underline{X}) = \frac{\sum_{l=1}^{M} \gamma^l \left(\prod_{i=1}^{n} exp\left[-\frac{(x_i - \alpha_i^l)^2}{\beta_i^l} \right] \right)}{\sum_{l=1}^{M} \left(\prod_{i=1}^{n} exp\left[-\frac{(x_i - \alpha_i^l)^2}{\beta_i^l} \right] \right)} \quad (2)$$

where γ^l is the centre of the fuzzy set G^l; α_i^l defines the membership function mean value for the fuzzy set F_i^l associated to the variable x_i in the rule l; β_i^l is the variance with respect α_i^l, for the fuzzy set F_i^l associated to the variable x_i in the rule l.

Note that the FLS represented by equation (2) may be transformed into an Adaptive Fuzzy System (AFS) by properly adjusting the parameters γ^l, α_i^l and β_i^l using a learning algorithm. Reference [1] proposes a gradient descent based supervised learning mechanism for the tuning the before mentioned parameters.

In most practical applications, a FLS like (2), with a gradient descent based learning algorithm for adjusting the parameters α_i^l, β_i^l and γ^l, is not good enough for the construction of an adequate fuzzy model. In particular, based on the generic base of rules described by (1) and the characteristics of gradient descent methods there would be a different membership function of each fuzzy set defined in the base of rules and therefore the linguistic values F_i^l are only used into rule l [2]. As a consequence, there is not any guarantee of overlapping different fuzzy sets and the rules may be weakly activated, or not activated at all, for some input data presented after the training is completed. This fact is particularly true when gaussian and triangular membership functions are used.

Figure 1 illustrates a partition of the variables space when an AFS as the one described by (2) is used. Reference [3] presents an approach that improves the fuzzy rules activation, by defining a base of knowledge with some rules showing the same fuzzy sets for some of the input variables. This approach allows to generate a base of rules similar to the ones obtained when the classical "Mamdani's tables" are used. This way, the number of tuning parameters in the resulting base of rules is decreased. Figure 2 illustrates a fuzzy partition of the variables space as suggested in [3].

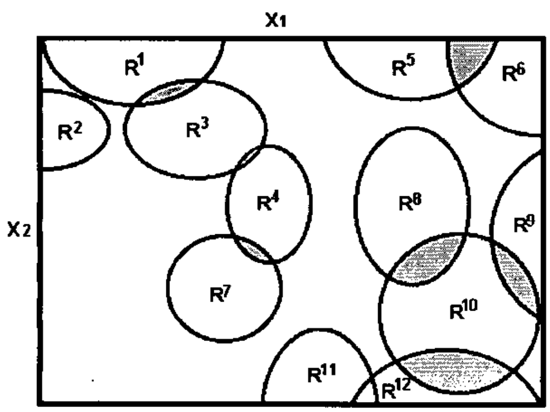

Fig. 1. Fuzzy partition by using a classic AFS

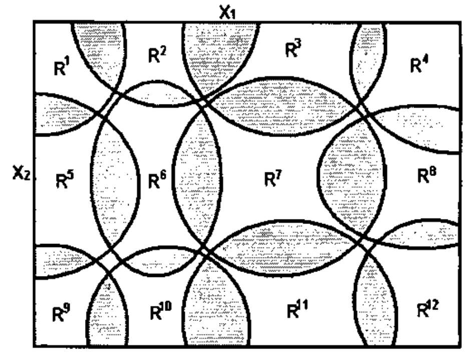

Fig. 2. Fuzzy partition using a Mamdani's table

III. DYNAMICAL ADAPTIVE FUZZY MODELS

The AFS presented in the previous section may be improved by defining dynamical membership functions which incorporates the temporal behaviour of the system into fuzzy models. This way, the resulting Dynamical Adaptive Fuzzy Model (DAFM) can adapt itself to changes in the domain of discourse of the system's variables.

Figure 3 depicts the idea behind of the definition of dynamical membership functions. Note that the dynamical characteristic of the membership functions avoids the before mentioned weakness associated with the activation of fuzzy rules in AFS.

The corresponding analytic expression of a DAFM based on (2), is as follows:

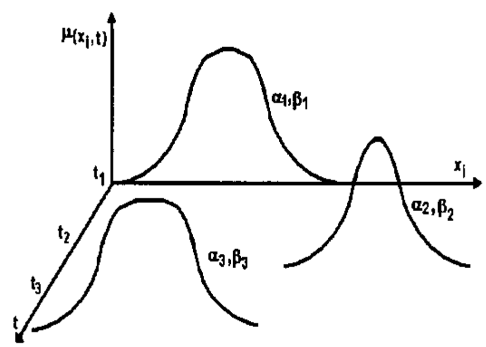

Fig. 3. Dynamical membership function for a given variable x_i

$$y(\underline{X}, t) = \frac{\sum_{l=1}^{M} \gamma^l(\underline{u}^l, t) \left(\prod_{i=1}^{n} exp\left[-\frac{\left(x_i - \alpha_i^l(\underline{v}_i^l, t)\right)^2}{\beta_i^l(\underline{w}_i^l, t)} \right] \right)}{\sum_{l=1}^{M} \left(\prod_{i=1}^{n} exp\left[-\frac{\left(x_i - \alpha_i^l(\underline{v}_i^l, t)\right)^2}{\beta_i^l(\underline{w}_i^l, t)} \right] \right)} \quad (3)$$

where $\underline{X}$ is a vector of input variables x_i; t is the time; $\underline{u}^l$ is a vector of P parameters u_p^l of the function γ^l; $\underline{v}_i^l$ is a vector of Q parameters v_{iq}^l of the function α_i^l; $\underline{w}_i^l$ is a vector of R parameters w_{ir}^l of the function β_i^l; $p = 1, ..., P, q = 1, ..., Q, r = 1, ..., R$.

Once the DAFM is defined, the general structure of the functions $\gamma^l(\underline{u}^l, t)$, $\alpha_i^l(\underline{v}_i^l, t)$ and $\beta_i^l\ (\underline{w}_i^l, t)$ has to be specified and a procedure to obtain the parameters u_p^l, v_{iq}^l and w_{ir}^l should be proposed.

A. General structure of the functions

Functions $\gamma^l(\underline{u}^l, t)$, $\alpha_i^l(\underline{v}_i^l, t)$ and $\beta_i^l\ (\underline{w}_i^l, t)$ should be chosen in such a way they represent the whole domain of discourse range of input and output variables through time.

Let $x_i(t_j)$, $i = 1, ..., n$, be the input variables values to the DAFM at time t_j which generate the output $y(t_j)$. Taking into account the meaning of the functions $\alpha_i^l(\underline{v}_i^l, t_j)$ and $\beta_i^l\ (\underline{w}_i^l, t_j)$ in the gaussian expression given in (3), a general structure for these functions may be stablished as:

$$\alpha_i^l(\underline{v}_i^l, t_j) = f(\underline{v}_i^l, \overline{x}_i(t_j)) \tag{4}$$

$$\beta_i^l(\underline{w}_i^l, t_j) = g(\underline{w}_i^l, \sigma_i^2(t_j)) \tag{5}$$

$$\gamma^l(\underline{u}^l, t_j) = h(\underline{u}^l, \overline{y}(t_j)) \tag{6}$$

where:

$$\overline{x}_i(t_j) = \frac{\sum_{k=1}^{j}(x_i(t_k))}{j} \tag{7}$$

$$\sigma_i^2(t_j) = \frac{\sum_{k=1}^{j}(x_i(t_k) - \overline{x}_i(t_k))^2}{j} \tag{8}$$

$$\overline{y}(t_j) = \frac{\sum_{k=j-\delta}^{j-1}\ (y(t_k))}{\delta},\ \ \delta \in \aleph \tag{9}$$

or, alternatively:

$$\overline{y}(t_j) = \frac{\sum_{k=1}^{j-1}(y(t_k))}{j-1} \tag{10}$$

The equation (7) is the sample mean of the previous observations of the input variables x_i until time t_j. The equation (8) is the average of the sample deviation of the value $x_i(t_k)$ with respect to $\overline{x}_i(t_k)$, until time t_j. The equation (9) is the average of the δ previous observations of the output variable y until time $t_{(j-1)}$, while the equation (10) is the sample mean of the previous observations of the output variable y until $t_{(j-1)}$.

In this work, the general structure of the previous functions are proposed as follows:

$$\alpha_i^l(\underline{v}_i^l, t_j) = v_{i1}^l * \overline{x}_i(t_j) \tag{11}$$

$$\beta_i^l(\underline{w}_i^l, t_j) = w_{i1}^l * (\sigma_i^2(t_j) + w_{i2}^l) \tag{12}$$

$$\gamma^l(\underline{u}^l, t_j) = u_1^l * \overline{y}(t_j) \tag{13}$$

The expression proposed in (11) allows the adjusting of the membership functions mean value of the fuzzy sets F_i^l around the sample mean $\overline{x}_i(t_j)$, through the parameter v_{i1}^l. The expression proposed in (12) allows the adjusting of such a membership functions base around $\sigma_i^2(t_j)$. If $\sigma_i^2(t_j)$ has a very small value, the parameter w_{i2}^l avoid an indeterminate number in the equation (3) when the function $\beta_i^l(\underline{w}_i^l, t_j)$ is computed. The equation (13), allows the adjusting of the centre of fuzzy set G^l around $\overline{y}(t_j)$ through the parameter u_1^l.

B. The parameters adjustment

In this work, the gradient descent based algorithm is used for the parameters tuning of DAFM.

Based on the mean quadratic error E given by the equation (14):

$$E = \frac{1}{2}(y_e(t_j) - y(t_j))^2 \tag{14}$$

where $y(t_j)$ is the output of the system at time t_j and $y_e(t_j)$ is the output estimated by the fuzzy model (3) at time t_j; the adjustment laws using the gradient descent based method, are given by the equations (15), (16), (17):

$$u_p^l(K+1) = u_p^l(K) - \rho_1 \left.\frac{\partial E}{\partial u_p^l}\right|_K \tag{15}$$

$$v_{iq}^l(K+1) = v_{iq}^l(K) - \rho_2 \left.\frac{\partial E}{\partial v_{iq}^l}\right|_K \tag{16}$$

$$w_{ir}^l(K+1) = w_{ir}^l(K) - \rho_3 \left.\frac{\partial E}{\partial w_{ir}^l}\right|_K \tag{17}$$

where K is the current iteration in the training phase and ρ_j is the learning rate ($j = 1, 2, 3$).

Developing the previous expressions, it has that:

$$\left.\frac{\partial E}{\partial u_p^l}\right|_K = \left.\frac{\partial E}{\partial \gamma^l(\underline{u}^l, t_j)} \frac{\partial \gamma^l(\underline{u}^l, t_j)}{\partial u_p^l}\right|_K \tag{18}$$

$$\left.\frac{\partial E}{\partial v_{iq}^l}\right|_K = \left.\frac{\partial E}{\partial \alpha_i^l(\underline{v}_i^l, t_j)} \frac{\partial \alpha_i^l(\underline{v}_i^l, t_j)}{\partial v_{iq}^l}\right|_K \tag{19}$$

$$\left.\frac{\partial E}{\partial w_{ir}^l}\right|_K = \left.\frac{\partial E}{\partial \beta_i^l(\underline{w}_i^l, t_j)} \frac{\partial \beta_i^l(\underline{w}_i^l, t_j)}{\partial w_{ir}^l}\right|_K \tag{20}$$

In order to simplify the notation, it is denoted $\gamma^l(\underline{u}^l, t_j) = \gamma^l$, $\alpha_i^l(\underline{v}_i^l, t_j) = \alpha_i^l$ and $\beta_i^l(\underline{w}_i^l, t_j) = \beta_i^l$. Then,

$$\frac{\partial E}{\partial \gamma^l} = \frac{(y_e(t_j) - y(t_j))}{d(t_j)} h^l(t_j) \tag{21}$$

$$\frac{\partial E}{\partial \alpha_i^l} = 2\frac{(y_e(t_j) - y(t_j))}{d(t_j)}(\gamma^l - y_e(t_j))h^l(t_j)\frac{(x_i - \alpha_i^l)}{\beta_i^l} \tag{22}$$

$$\frac{\partial E}{\partial \beta_i^l} = \frac{(y_e(t_j) - y(t_j))}{d(t_j)}(\gamma^l - y_e(t_j))h^l(t_j)\frac{(x_i - \alpha_i^l)^2}{(\beta_i^l)^2} \tag{23}$$

where:
$d(t_j) = \sum_{l=1}^{M} h^l(t_j)$
$h^l(t_j) = \prod_{i=1}^{n} exp\left[-\frac{(x_i - \alpha_i^l(t_j))^2}{\beta_i^l(t_j)}\right]$

Substituting (18), (19) and (20) respectively into (15), (16) and (17), the adjustment laws of the parameters are obtained. In the particular case of the functions proposed in (11), (12) and (13), it has that:

$$\frac{\partial \alpha_i^l(t_j)}{\partial v_{i1}^l} = \overline{x_i}(t_j) \quad (24)$$

$$\frac{\partial \beta_i^l(t_j)}{\partial w_{i1}^l} = (\sigma_i^2(t_j) + w_{i2}^l) \quad (25)$$

$$\frac{\partial \beta_i^l(t_j)}{\partial w_{i2}^l} = w_{i1}^l \quad (26)$$

$$\frac{\partial \gamma^l(t_j)}{\partial u_1^l} = \overline{y}(t_j) \quad (27)$$

IV. ILLUSTRATIVE EXAMPLES

In the following sections, two examples illustrate the performance of the DAFM in system identification. The performance of the fuzzy model in system identification is evaluated according to the identification error $e(t)$ defined as:

$$e(t) = [y_e(t_j) - y(t_j)] \quad (28)$$

where $y(t_j)$ is the output of the system and $y_e(t_j)$ is the estimated output by the fuzzy model (3) at time t_j.

The first example shows a system with an unknown nonlinear part; in the second the output is not well-known. The DAFM will be used in order to estimate the unknown parts.

In order to propose an identification fuzzy model, the equation (9) is used taking $\delta = 1$, then $\overline{y}(t_j) = y(t_{j-1})$. This way, the centre of fuzzy set G^l depend on the last available value of the output y.

A. Example 1

In this example, the system has been described by the following difference equation:

$$y(k+1) = 0.3y(k) + 0.6y(k-1) + g[u(k)] \quad (29)$$

where $g[u(k)] = 0.6\sin(\pi u(k)) + 0.3\sin(3\pi u(k)) + 0.1\sin(5\pi u(k))$

The estimated function $y_e(k+1)$ is:

$$y_e(k+1) = 0.3y(k) + 0.6y(k-1) + g_e[u(k)] \quad (30)$$

where $g_e[u(k)]$ is estimated by using a DAFM.

The input variable to the fuzzy model is $x_1(k) = u(k)$ and $y(t_{(j-1)}) = g[u(k-1)]$. A set of 1000 training patterns has been obtained from a random input on the interval [-1,1], and 1000 training cycles has been made in the training phase with an arbitrary initial values of the parameters on the interval [0,1]. The best models with respect to the identification error have been obtained with $M = 10, 20, 30$ and $\rho_i = 0.1$

The performance of the previous models has been tested with the input signal $u(k) = \sin(2\pi k/250)$. Figure 4 shows the identification error of $y(k+1)$. The low error has been achieved with $M = 10$. The figure 5 illustrates the performance of such a fuzzy model. Also, the performance of the previous fuzzy model has been tested using the input $u_1(k)$. Figure 6 shows the performance of this fuzzy model.

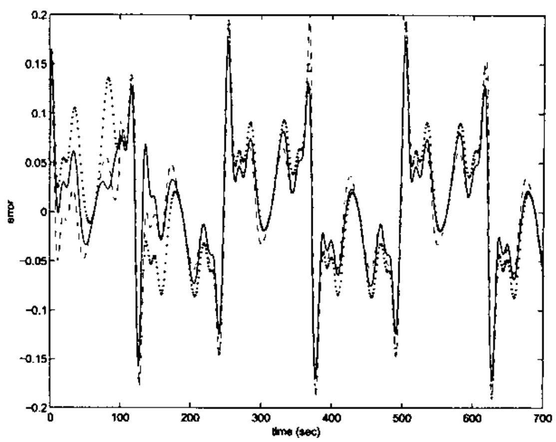

Fig. 4. Identification error using $u(k)$ in the example 1. The first model ($M = 10$) has the solid line, the second model ($M = 20$) has the dashed line and third model ($M = 30$) has the pointed line.

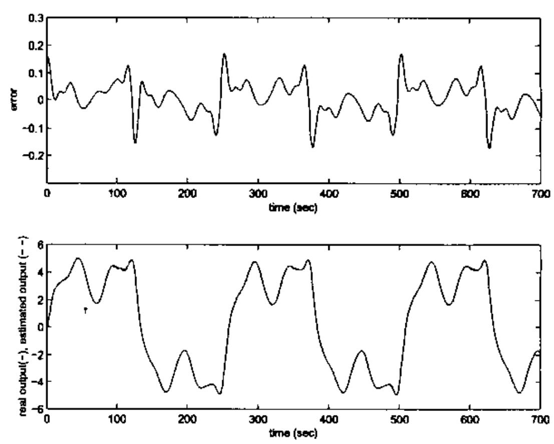

Fig. 5. Real output and estimated output (dashed line) using the fuzzy model with $M = 10$ and the input $u(k)$, in the example 1

$$u_1(k) = \begin{cases} \sin(2\pi k/250) & if \quad otherwise \\ 3 + (0.5(u_a + u_b)) & if \quad 251 < k < 500 \end{cases} \quad (31)$$

where $u_a = \sin(2\pi k/250)$ and $u_b = \sin(2\pi k/25)$.

Based on the identification error, the performance of the proposed fuzzy model performance is adequate. The identification

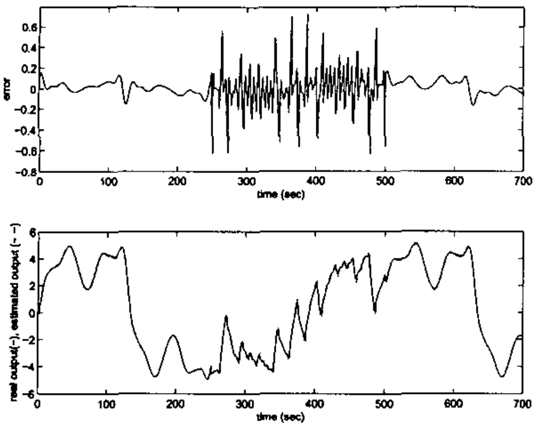

Fig. 6. Real output and estimated output (dashed line) using the fuzzy model with $M = 10$ and the input $u_1(k)$, in the example 1

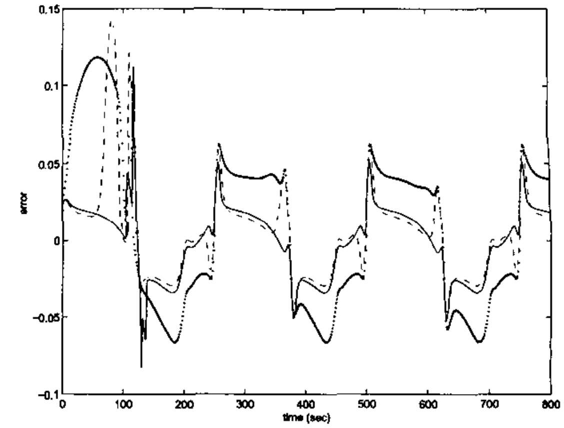

Fig. 7. Identification error using $u(k)$ in the example 2. The first model ($M = 10$) has the solid line, the second model ($M = 20$) has the dashed line and third model ($M = 30$) has the pointed line.

error in the figure 6 is increased when the input signal suddenly changes, however, the identification fuzzy model follows the real output. In [1] an adaptive fuzzy model with $M = 40$, 120 adjustable parameters and 5000 training cycles is proposed. Here, the proposed model uses $M = 10$ and 30 adjustable parameters.

B. Example 2

In this example, the system has been described by the following difference equation:

$$y(k+1) = g[y(k), y(k-1), y(k-2), u(k), u(k-1)] \quad (32)$$

where $g[.] = \frac{y(k)y(k-1)y(k-2)u(k-1)(y(k-2)-1)+u(k)}{1+y(k-2)^2+y(k-1)^2}$
The estimated function $y_e(k+1)$ is:

$$y_e(k+1) = g_e[y(k), y(k-1), y(k-2), u(k), u(k-1)] \quad (33)$$

where $g_e[.]$ is estimated using a DAFM.

The input variables to the fuzzy model are $x_1(k) = y(k)$, $x_2(k) = y(k-1)$, $x_3(k) = y(k-2)$, $x_4(k) = u(k)$ y $x_5(k) = u(k-1)$ and $y(t_{(j-1)}) = g[u(k-1)]$. A set of 1000 training patterns has been obtained using a random input on the interval [-1,1], and 1000 training cycles has been made in the training phase from arbitrary initial values of the parameters on the interval [0,1]. The best models with respect to the identification error are given on the table I. The performance

TABLE I
TRAINING PHASE. EXAMPLE 1

M	ρ_1	ρ_2	ρ_3
10	0.3	0.1	0.3
20	0.3	0.1	0.3
30	0.5	0.3	0.5

of the previous models has been tested with the input signal $u(k) = \sin(2\pi k/250)$. In the figure 7, the identification errors of $y(k+1)$ is shown. The low error has been achieved with $M = 10$. Figure 8 shows its performance using the input $u(k)$. Figure 9 illustrates the performance of this model using

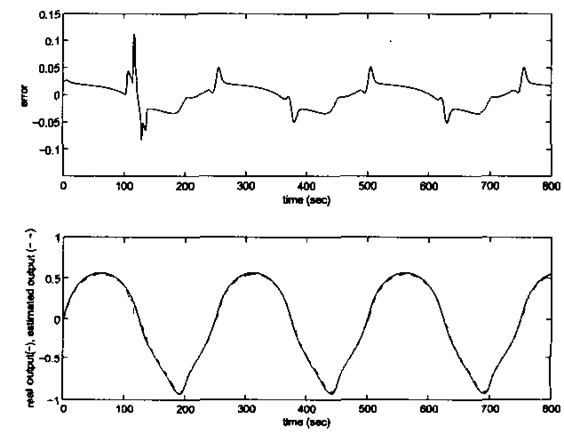

Fig. 8. Real output and estimated output (dashed line) using the fuzzy model with $M = 10$ and the input $u(k)$, in the example 2

the input $u_1(k)$.

$$u_1(k) = \begin{cases} \sin(2\pi k/250) & if \quad otherwise \\ 1.5 + (u_a + u_b) & if \quad 501 < k < 800 \end{cases} \quad (34)$$

where $u_a = 0.8\sin(2\pi k/250)$ and $u_b = 0.2\sin(2\pi k/25)$
According to the identification error, the proposed fuzzy model has an adequate performance. In the figure 9 The identification error is increased just at time when the input signal suddenly changes; however, like the previous example, the identification fuzzy model follows the real output and the identification error remain around cero.

This example has been developed in [1], using a model with $M = 40$, 440 adjustable parameters and 5000 training

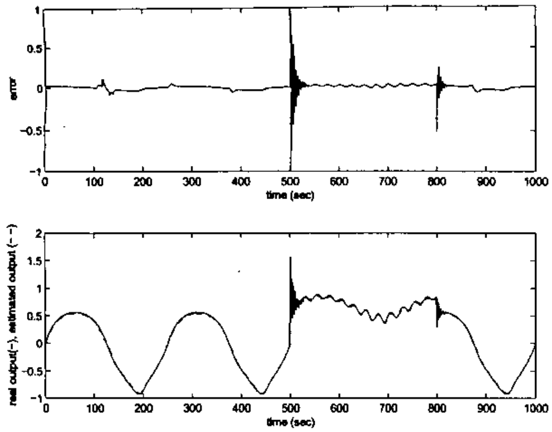

Fig. 9. Real output and estimated output (dashed line) using the fuzzy model with $M = 10$ and the input $u_1(k)$, in the example 2

cycles from an adequate selection of the initial values of the parameters. Here, the proposed model uses $M = 10$ and 110 adjustable parameters from an initial randomly selection of parameters values.

V. CONCLUSION

New approaches in fuzzy modeling that permit to solve practical limitations found in classic adaptive fuzzy modeling, are considered an interesting contribution in the fuzzy logic field.

In this work, an approach for dynamical adaptive fuzzy modeling is proposed. This approach permits incorporate into the fuzzy membership functions the temporal behaviour of the system variables, allowing to the fuzzy model adapt itself to the changes that dynamically can be presented in the domains of discourse. The resulting dynamical adaptive fuzzy model permits to improve the fuzzy rules activation and the overlapping of the fuzzy sets.

The functions that describes the dynamical membership functions are based on the sample mean and sample deviation of the available observation about the variables of the system. The parameters adjustment algorithm is based on the descent gradient supervised learning method, but the design of on-line learning algorithms could be an interesting goal.

The illustrative examples in system identification show that the performance of the proposed fuzzy models based on the identification error is adequate. These models follows the real output using input signals with sudden changes on the time.

REFERENCES

[1] L. Wang, Adaptive Fuzzy Systems and Control. Design and Stability, New Jersey: Prentice Hall, 1994.

[2] Y. Shi and M. Mizumoto, "Some considerations on conventional neuro-fuzzy learning algorithms by gradient descent method", *Fuzzy Sets and Systems*, Vol. 112:1, pp. 51-63, 2000.

[3] Y. Shi and M. Mizumoto, "A new approach of neuro-fuzzy learning algorithm for tuning fuzzy rules", *Fuzzy Sets and Systems*, vol. 112:1, pp. 99-116, 2000.

[4] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", *IEEE Trans. Syst., Man, Cybern*, vol. 15, pp. 116-132, 1985.

[5] M. Cerrada, "Systèmes Flous Adaptatifs Dynamiques pour l'identification des Systèmes Non-Linéaires Variants avec le Temps", *D.E.A. Report*, Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS-CNRS), Toulouse, France, 2000.