A RELATIONAL APPROACH FOR COMPLEX SYSTEM IDENTIFICATION

Ricardo José G.B. Campello
ricardo@dca.fee.unicamp.br

Wagner C. Amaral
wagner@dca.fee.unicamp.br

DCA/FEEC/Unicamp, Campinas, SP
CP 6101, CEP 13083-970

Resumo - Este trabalho aborda a questão de identificação de sistemas complexos com o auxílio de lógica nebulosa. Apresentam-se dois métodos numéricos para estimar os parâmetros de modelos relacionais, que são modelos construídos a partir de equações relacionais nebulosas. O primeiro método é baseado na solução de uma sequência de problemas quadráticos com o objetivo de refinar modelos previamente identificados através de outros algoritmos. A partir deste método, obtém-se também um algoritmo recursivo adequado a aplicações "on-line". Avalia-se o desempenho dos métodos propostos através da modelagem de um sistema dinâmico real.

Palavras-Chave: Modelagem *Fuzzy*, Otimização, Sistemas Não-lineares.

Abstract In this paper the issue of complex system identification with the aid of fuzzy logic techniques is addressed. Models based on fuzzy relational equations, i.e. fuzzy relational models, are presented. Two numerical methods to estimate the parameters of such models are proposed. The first one is an optimization based methodology using sequential procedures of quadratic programming to refine models previously estimated by other methods. From this methodology, a recursive algorithm suitable for on-line identification is derived. The performance of the methods proposed is evaluated by modeling a real dynamic process.

Keywords: Fuzzy modeling, Optimization, Non-linear systems.

1 INTRODUCTION

The concept of fuzzy modeling can be defined as the representation of systems based on fuzzy logic techniques. Its essence comes from the early ideas of Zadeh, 1973 on human behavior representation using fuzzy algorithms. Fuzzy models have become an important mathematical tool for the identification of complex systems, such as the physically unknown and strongly non-linear ones, especially because these models can be constructed as universal approximators (Wang and Mendel, 1992, Kosko, 1992, Sudkamp and Hammell II, 1994, Kosko, 1997) and can deal with linguistic knowledge about the systems to be identified.

The first attempt at system identification using fuzzy models was made by Tong, 1978. In his work Tong used a rule-based model (constituted by a set of fuzzy rules) to model dynamic systems. In this kind of fuzzy model, called linguistic models, the problem of getting an adequate rule set is not a trivial task since the rules involve linguistic terms (fuzzy sets) in their antecedents and consequents. An alternative structure was proposed by Takagi and Sugeno, 1985. The Takagi-Sugeno fuzzy model is constituted by a set of fuzzy rules whose consequents are crisp (non-fuzzy) functions which map the model inputs into the output. The parameters of these functions can be estimated using Least Squares (LS) methods or Kalman filter (Ljung, 1987). In this approach, however, the linguistic interpretability of the rules is lost. Another approach for fuzzy modeling is based on the theory of fuzzy relational equations (Pedrycz, 1993). This kind of model, called fuzzy relational models, can be viewed as a simplification of the linguistic models since a set of fuzzy rules can be written as a relational equation (Yager and Filev, 1994). The main advantage of this simplification is that in linguistic models the fuzzy relation of the relational equation is derived from the aggregation of the rule set whose linguistic terms must be determined, whereas in relational models the fuzzy relation is only a matrix to be estimated.

The use of relational models in system identification and control was first investigated by Czogała and Pedrycz, 1981. Nevertheless, this kind of model became quite useful in practical applications when Pedrycz, 1984 proposed the utilization of the fuzzy discretization technique (see Appendix) for data representation in relational equations. This technique was widely spread in the field of fuzzy relational models since it can provide significant reductions in the dimensions of these models as well as linguistic meaning for them (Pedrycz, 1984, Graham and Newell, 1989, Campello and Amaral, 1999).

In the present paper two new methods to estimate the parameters of fuzzy relational models are proposed. The first one is an optimization based methodology (called refinement or fine-tuning algorithm) using sequential procedures of quadratic programming to refine models previously estimated by other methods (e.g. Pedrycz, 1984, Xu and Lu, 1987, Campello et al., 1998). From this methodology, a recursive algorithm with a very simple updating law is derived. This algorithm is well-suited for on-line identification.

The paper is organized as follows. In the next section the detailed formulation of the fuzzy relational models is presented. Following, the methods for refinement and recursive estimation

[0] Artigo submetido em 08/10/1998
1a. Revisão em 03/03/1999;
Aceito sob recomendação do Ed. Cons. Prof. Dr. Ricardo Tanscheit

of these models are derived in Section 3. Next, in Section 4 a numerical example is provided to evaluate the performance of the proposed methods in the identification of a dynamic system. Finally, the conclusions are addressed in Section 5.

2 DYNAMIC AND STATIC FUZZY RELATIONAL MODELS

Consider the following dynamic system[1]:

$$\begin{aligned} y(k) \;&=\; \mathcal{F}(\, y(k-1), \cdots, y(k-p_y), \\ &\qquad u_1(k-t_1-1), \cdots, u_1(k-t_1-p_{u_1}), \cdots \\ &\qquad u_v(k-t_v-1), \cdots, u_v(k-t_v-p_{u_v})\,) \end{aligned} \qquad (1)$$

where p_y and p_{u_i} for $i = 1, \cdots, v$ are the "orders" of the output y and inputs u_i, respectively, t_i are the time delays, k is the discrete time variable and $\mathcal{F}$ is a non-linear operator. This system can be rewritten as

$$y(k+1) \;=\; \mathcal{F}(\, x_1(k), x_2(k), \cdots, x_n(k)\,) \qquad (2)$$

where $n = p_y + p_{u_1} + p_{u_2} + \cdots + p_{u_v}$ and

$$\begin{aligned} &x_1(k) \,=\, y(k),\; x_2(k) \,=\, y(k-1), \cdots \\ &x_{p_y}(k) \,=\, y(k-p_y+1),\; x_{p_y+1}(k) \,=\, u_1(k-t_1), \cdots \\ &x_n(k) \,=\, u_v(k-t_v-p_{u_v}+1) \end{aligned} \qquad (3)$$

System (2) can be represented (modeled) using the following fuzzy relational equation:

$$Y(k+1) = X_1(k) \bullet X_2(k) \bullet \cdots \bullet X_n(k) \bullet R \qquad (4)$$

where $Y = [Y_1 \cdots Y_{c_0}]$ and $X_i = [X_{i_1} \cdots X_{i_{c_i}}]$ ($i = 1, \cdots, n$) are linguistic (fuzzy) representations of the numerical (non-fuzzy) output y and inputs x_i, respectively, R ($c_1 \times \cdots \times c_n \times c_0$) is the fuzzy relational matrix (fuzzy relation) and "$\bullet$" denotes the fuzzy composition operator.

Equation (4) can be simplified as

$$Y(k+1) \;=\; X(k) \bullet R \qquad (5)$$

where X ($c_1 \times \cdots \times c_n$) is the Cartesian product (Lee, 1990) of the fuzzy inputs,

$$X(k) \;=\; X_1(k) \,\times\, X_2(k) \,\times\, \cdots \,\times\, X_n(k) \qquad (6)$$

which is generically defined as the product of a triangular norm (t-norm) (Pedrycz, 1993, Pedrycz and Gomide, 1998) over the cross product space of the fuzzy inputs.

Equations (5) and (6) can be written pointwisely for specific fuzzy composition operator and t-norm of the Cartesian product, respectively. With respect to the fuzzy composition operator, the present paper deals with the well-known max-t fuzzy composition which is the most commonly used composition operator in fuzzy systems such as fuzzy models and controllers. Then, Equations (5) and (6) are rewritten as

$$Y_j(k+1) \;=\; \bigvee_{\substack{l_i = 1 \\ i = 1, \cdots, n}}^{c_i} X_{l_1,\ldots,l_n}(k) \tau R_{l_1,\ldots,l_n,j}, \quad j = 1, \cdots, c_0 \qquad (7)$$

and

$$X_{l_1,\ldots,l_n}(k) \;=\; X_{1_{l_1}}(k)\, \psi\, X_{2_{l_2}}(k)\, \psi\, \cdots\, \psi\, X_{n_{l_n}}(k) \qquad (8)$$

where $X_{l_1,\ldots,l_n}$ and $R_{l_1,\ldots,l_n,j}$ ($l_i = 1, \cdots, c_i$; $i = 1, \cdots, n$; $j = 1, \cdots, c_0$) are elements of X and R, respectively, "$\vee$" stands for the max operator, τ is the t-norm associated with the max-t composition and ψ is the t-norm associated with the Cartesian product in (6).

Given a set of input-output data pairs the identification problem is to find a matrix R such that Equation (4) is completely or approximately satisfied for these data. According to the fuzzy discretization concept (see Appendix) the input and output data can be represented (fuzzified) using normal and convex reference fuzzy sets, as follows:

$$X_i(k) \;=\; [\, \mathcal{X}_{i_1}(x_i(k)) \;\; \mathcal{X}_{i_2}(x_i(k)) \;\; \cdots \;\; \mathcal{X}_{i_{c_i}}(x_i(k))\,] \qquad (9)$$

$$Y(k+1) \;=\; [\, \mathcal{Y}_1(y(k+1)) \;\; \cdots \;\; \mathcal{Y}_{c_0}(y(k+1))\,] \qquad (10)$$

where $\mathcal{X}_{i_{l_i}}$ is the l_i-th reference fuzzy set in the i-th input interface and $\mathcal{Y}_j$ is the j-th reference fuzzy set in the output interface. The reference fuzzy sets $\mathcal{X}_{i_{l_i}}$ ($l_i = 1, \cdots, c_i$; $i = 1, \cdots, n$) and $\mathcal{Y}_j$ ($j = 1, \cdots, c_0$) are defined over the universes of discourse $\mathbf{X}_i$ and $\mathbf{Y}$ of x_i and y, respectively. The number of fuzzy sets used in the fuzzy model determines its degree of specificity. This means that the approximation capability of the model is proportional to the quantities c_0 and c_i ($i = 1, \cdots, n$). By the other side, the number of parameters of the model to be estimated (given by the size of the relational matrix R) is also proportional to these quantities.

Now, consider the following static system:

$$y \;=\; \mathcal{G}(\, x_1, \cdots, x_n\,) \qquad (11)$$

where $x_1, \cdots, x_n$ and y are numerical input and output variables, respectively, and $\mathcal{G}$ is a non-linear function which maps $x_1, \cdots, x_n$ into y. The representation of this system by means of a fuzzy relational model is straightforward and can be derived just by omitting the time variable k in Equations (4) to (10). In this way, although the present paper focus on the problem of dynamic system identification, the algorithms and results presented in the following sections also hold in the context of static systems.

[1]For simplicity, only multi-input/single-output (MISO) discrete-time dynamic systems are discussed.

3 RELATIONAL MATRIX ESTIMATION

3.1 Optimization criterion

Consider the estimated output of the fuzzy model $\tilde{y}$, the measured output of the real system y and their linguistic representations $\tilde{Y}$ and Y, respectively[2]. The relations between these numerical and linguistic outputs can be written as

$$Y = \mathcal{L}(y) \tag{12}$$
$$\tilde{y} = \mathcal{N}(\tilde{Y}) \tag{13}$$

where $\mathcal{L}$ and $\mathcal{N}$ represent the mappings between numerical and linguistic information in the fuzzification and defuzzification procedures, respectively. In the present paper, the fuzzification $\mathcal{L}$ is based on the fuzzy discretization concept, as discussed in Section 2. It is given by Equations (9) and (10) (for the input and output variables, respectively). The defuzzification used is a slight modification of the center of gravity method, called weighted average (Pedrycz, 1993, pp. 109).

Let the input (fuzzification) and output (defuzzification) interfaces of the fuzzy relational model be constructed as optimal interfaces (Oliveira, 1995), meaning that they satisfy the information equivalence criterion[3] (Pedrycz and Oliveira, 1996) for all numerical values of their universes of discourse, as follows:

$$\forall \;\; a \in \mathbf{A}: \;\; \mathcal{N}(\mathcal{L}(a)) = a \tag{14}$$

where $\mathbf{A}$ is the universe of discourse of a hypothetic optimal interface. Examples of optimal interfaces are those associated with the fuzzification and defuzzification methods considered above, having fuzzy sets with triangular membership functions overlapped at a degree of 0.5 and equally spaced centers. Optimal interfaces with other shapes of membership functions can be constructed using optimization algorithms (Oliveira, 1995).

From the above considerations an optimization criterion for fuzzy relational model identification can be derived based on the following proposition.

Proposition 1

Let the mappings $\mathcal{N}$ and $\mathcal{L}$ in (13) and (12), respectively, be implemented through an optimal interface. Then, the equality $\tilde{Y} = Y$ between the fuzzy outputs of the model and the system results in the equality $\tilde{y} = y$ between their respective non-fuzzy outputs.

Proof: If $\tilde{Y} = Y$, then Equation (13) can be rewritten as

$$\tilde{y} = \mathcal{N}(Y) \tag{15}$$

Substituting (12) into (15) results in

$$\tilde{y} = \mathcal{N}(\mathcal{L}(y)) \tag{16}$$

Then, the equality $\tilde{y} = y$ is obtained using both Equation (16) and the optimal interfaces concept (14). ■

The above result means that an approach for fuzzy model identification can be derived by the minimization, over the relational matrix R, of a criterion of distance between the fuzzy outputs of the system and model, i.e.

$$\min_R \; f(Y, \tilde{Y}) \tag{17}$$

where f is a generic distance criterion. However, the fuzzy output of the model is bounded above by the fuzzy input in the relational equation, i.e., $\tilde{Y}$ (computed by means of Equation (7)) belongs to the interval $[\,0, G'\,]^{c_0}$, where $G' \in [0,1]$ is given by

$$G' \;=\; \max(X) \;=\; \bigvee_{\substack{l_i = 1 \\ i = 1, \cdots, n}}^{c_i} X_{l_1,\ldots,l_n} \tag{18}$$

Since the fuzzy output of the system Y (computed by means of Equation (10)) belongs to the interval $[0,1]^{c_0}$, $\tilde{Y}$ belongs to a subinterval of Y and, consequently, a direct minimization as in (17) may be inefficient.

For the worst case scenario concerning the numerical inputs $x_i \in \mathbf{X}_i$ $(i = 1, \cdots, n)$, G' takes on the lowest possible value, designated G, that is

$$G \;=\; \mathop{Inf}_{\substack{x_i \in \mathbf{X}_i \\ i = 1, \cdots, n}} \; \max(X) \tag{19}$$

Then, Problem (17) can be changed into

$$\min_R \; f(GY, \tilde{Y}) \tag{20}$$

where $GY \in [0, G]^{c_0}$. Since by definition $G' \geq G$ for all $x_i \in \mathbf{X}_i$ $(i = 1, \cdots, n)$, GY belongs to a subinterval of $\tilde{Y}$ and the minimization in (20) can be successfully undertaken.

In this context it is important to notice that, in general, a gain in the values of the membership function of an output fuzzy set, for example A, does not change the defuzzified output $\tilde{a}$, as follows (see Figure (1)):

$$\tilde{a} \;=\; \mathcal{N}(A) \;=\; \mathcal{N}(GA) \tag{21}$$

This is the case for classical defuzzification methods, such as the center of gravity, mean of maxima and weighted average, as shown in the equation below:

$$\tilde{a} \;=\; \frac{\sum_j GA_j \theta_j}{\sum_j GA_j} \;=\; \frac{\sum_j A_j \theta_j}{\sum_j A_j} \tag{22}$$

where θ_j is the modal value of the j-th reference fuzzy set of the output interface. As a consequence, the result of Proposition

[2] The time variable k is omitted for simplicity.

[3] The information equivalence criterion means that a numerical value can be completely recovered after a fuzzification-defuzzification sequence through an interface.

1 can be extended for Problem (20) by means of the following proposition.

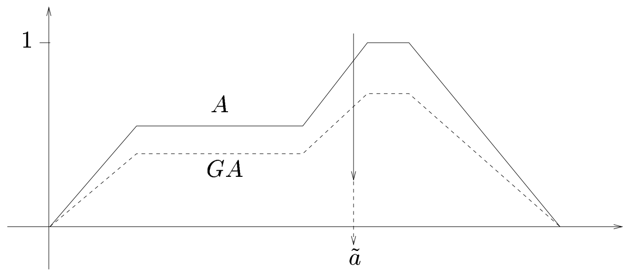

Figure 1: Defuzzification of two proportional fuzzy sets.

Proposition 2

Regarding the conditions of Proposition 1, if the defuzzification mapping $\mathcal{N}$ is such that Equation (21) is satisfied, then the equality $\tilde{Y} = GY$ is optimal since it results in the equality $\tilde{y} = y$ between the non-fuzzy outputs of the system and the model.

Proof: If $\tilde{Y} = GY$, then Equation (13) can be rewritten as

$$\tilde{y} = \mathcal{N}(GY) \tag{23}$$

From (23) and (21),

$$\tilde{y} = \mathcal{N}(Y) \tag{24}$$

and analogously to Proposition 1,

$$\tilde{y} = \mathcal{N}(Y) = \mathcal{N}(\mathcal{L}(y)) = y \tag{25}$$

■

The computation of G is derived from the following theorem.

Theorem 1 *The value G in (19) is given by:*

$$G = g_1 \psi g_2 \psi \cdots \psi g_n \tag{26}$$

$$g_i = Inf \max_{x_i \in \mathbf{X}_i} (\mathcal{X}_{i_1}(x_i), \mathcal{X}_{i_2}(x_i), \cdots, \mathcal{X}_{i_{c_i}}(x_i)) \tag{27}$$

where ψ is the t-norm defined in (8) and g_i is the minimum grade for the union of the reference fuzzy sets in the i-th input interface (see Figure (2)).

Proof: From Equation (8) and due to the monotonicity of the triangular norms (i.e. for $a \leq b$ and $d \leq e$: $a \,\psi\, d \leq b \,\psi\, e$ $\forall a, b, d, e \in [0, 1]$) the following equation can be written:

$$\max(X) = \max(X_1) \;\psi\; \max(X_2) \;\psi\; \cdots \;\psi\; \max(X_n) \tag{28}$$

Analogously,

$$Inf \max_{\substack{x_i \in \mathbf{X}_i \\ i = 1, \cdots, n}} (X) = Inf \max_{x_1 \in \mathbf{X}_1} (X_1) \;\psi\; \cdots \;\psi\; Inf \max_{x_n \in \mathbf{X}_n} (X_n) \tag{29}$$

Defining g_i as

$$g_i \overset{\triangle}{=} Inf \max_{x_i \in \mathbf{X}_i} (X_i) \tag{30}$$

and using Equation (19), Equation (29) becomes

$$G = g_1 \psi g_2 \psi \cdots \psi g_n \tag{31}$$

Moreover, from Equations (9) and (30) the following equation holds:

$$g_i = Inf \max_{x_i \in \mathbf{X}_i} (\mathcal{X}_{i_1}(x_i), \mathcal{X}_{i_2}(x_i), \cdots, \mathcal{X}_{i_{c_i}}(x_i)) \tag{32}$$

■

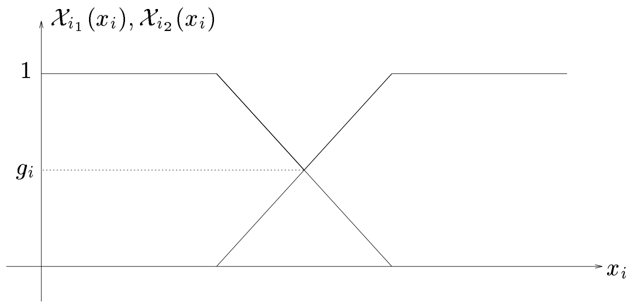

Figure 2: Minimum grade g_i for the union of two reference fuzzy sets.

It is important to notice that if the input interfaces are optimal, then they satisfy the condition of completeness (coverage of the universes of discourse) (Pedrycz, 1993). This property assures that null g_i's can not occur. Consequently, it can avoid a null gain G. However, to avoid numerical problems due to a very small value of G the t-norm ψ should be carefully chosen, especially when the number of inputs is too large (n too large). In these cases, an adequate choice could be, for instance, the min operator. Despite this, a triangular norm commonly used to implement the Cartesian product in fuzzy systems is the algebraic product, which is considered in the example presented in Section 4.

3.2 Refinement algorithm

Let a set of $N + 1$ input/output data pairs be available to be used in the refinement of a fuzzy relational model previously identified by any method (e.g. Pedrycz, 1984, Xu and Lu, 1987, Campello et al., 1998). The distance criterion in Problem (20) is defined as the following set of independent cost functions:

$$J_j = \frac{1}{2} \sum_{k=1}^{N} (GY_j(k+1) - \tilde{Y}_j(k+1))^2 \tag{33}$$

where $j = 1, \cdots, c_0$. Since $\tilde{Y}_j$ is computed using (7), Equation (33) can be rewritten as

$$J_j = \frac{1}{2}\sum_{k=1}^{N}(GY_j(k+1) \; - \bigvee_{\substack{l_i = 1 \\ i = 1, \cdots, n}}^{c_i} X_{l_1,\ldots,l_n}(k)\tau R_{l_1,\ldots,l_n,j})^2 \tag{34}$$

Moreover, since an initial model to be refined (namely an initial relational matrix) is available, the information about which element $X_{(\cdot)}\tau R_{(\cdot),j}$ of the fuzzy composition is maximum at every time instant $k = 1, \cdots, N$ is known (*a priori*). Then, Equation (34) can be simplified as

$$J_j = \frac{1}{2}\sum_{k=1}^{N}(GY_j(k+1) - X_{m^k}(k)\tau R_{m^k,j})^2 \tag{35}$$

where $m^k \in \{1, 2, \ldots, l\}$ ($l = c_1 \cdot c_2 \cdots c_n$) is the combination of $l_1, \ldots, l_n$ such that $X_{m^k}\tau R_{m^k,j}$ is the maximum element of the composition at the time instant k, i.e.

$$m^k = arg \max_{m=1,\cdots,l} X_m(k)\tau R_{m,j} \tag{36}$$

To guarantee that $X_{m^k}\tau R_{m^k,j}$ remains maximum during the optimization procedure, for every $j = 1, \cdots, c_0$ a set of $N(l-1)$ inequalities must be taken into consideration in the optimization problem, as follows:

$$\begin{aligned} & X_{m^k}(k)\tau R_{m^k,j} \geq X_{q^k}(k)\tau R_{q^k,j} \\ & k = 1, \cdots, N \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{37}$$

Due to the monotonicity of the triangular norms, the inequalities in (37) can be replaced by

$$\begin{aligned} & R_{m^k,j} \geq \Gamma_{m^k,q^k,j}(k) \\ & k = 1, \cdots, N \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{38}$$

where $\Gamma_{m^k,q^k,j}$ is the smallest value of $R_{m^k,j}$ for which the inequality in (37) related to the respective k and q^k is satisfied. If the t-norm τ is continuous on its domain ($[0,1]^2$) and satisfies the condition $a' \, \tau \, b \; < \; a \, \tau \, b \; : \; a' < a \; \forall \, a, a', b \in [0,1]$, then the smallest value of $R_{m^k,j}$ for which an inequality in (37) is satisfied is that value for which the lower bound of the inequality (the equality) is satisfied. Then, the constraints $\Gamma_{m^k,q^k,j}$ are given by the solution of the following equations:

$$\begin{aligned} & X_{m^k}(k)\tau\Gamma_{m^k,q^k,j}(k) = X_{q^k}(k)\tau R_{q^k,j} \\ & k = 1, \cdots, N \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{39}$$

Furthermore, it is necessary to assure during the optimization procedure that $R_{m^k,j}$ belongs to $[0,1]$ (membership function interval) for every $k = 1, \cdots, N$. Since any solution $\Gamma_{m^k,q^k,j}$ of Equation (39) belongs to the interval $[0,1]$, the condition $R_{m^k,j} \geq 0$ is assured in (38). Then, it is only necessary to assure that the condition $R_{m^k,j} \leq 1$ is satisfied.

For the sake of the issues presented above, a set of optimization problems to refine the relational matrix R is given by

$$\begin{aligned} & \min_{R_{m^k,j}} \quad J_j = \frac{1}{2}\sum_{k=1}^{N}(GY_j(k+1) - X_{m^k}(k)\tau R_{m^k,j})^2 \\ & \text{s. to} \qquad\qquad \Gamma_{m^k,q^k,j}(k) \leq R_{m^k,j} \leq 1 \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{40}$$

where $j = 1, \cdots, c_0$. If the t-norm τ is the algebraic product then (40) becomes a convex and continuous quadratic problem, that is

$$\begin{aligned} & \min_{R_{m^k,j}} \quad J_j = \frac{1}{2}\sum_{k=1}^{N}(GY_j(k+1) - X_{m^k}(k) R_{m^k,j})^2 \\ & \text{s. to} \qquad\qquad \Gamma_{m^k,q^k,j}(k) \leq R_{m^k,j} \leq 1 \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{41}$$

where $\Gamma_{m^k,q^k,j}$ is derived from (39) as

$$\Gamma_{m^k,q^k,j}(k) = \frac{X_{q^k}(k) R_{q^k,j}}{X_{m^k}(k)} \tag{42}$$

Problem (41) can be solved by means of quadratic programming (Bazaraa and Shetty, 1979) using as an initial feasible condition the fuzzy relational matrix of the model to be refined. It is worthwhile to remark that in this problem only the elements of the relational matrix which maximize the fuzzy composition are optimized. Also, the optimization is done inside the bounds (constraints) determined by the structure of the model to be refined together with the available data. Hence, the refinement will achieve better results for a model such that the majority of its elements (which are significant to represent the system) maximize the fuzzy composition at least at one time instant, and also for a consistent data set.

It is important to notice that the values of the elements $R_{q^k,j}$ in (42) are in general given by the initial relational matrix to be refined. However, if $m^{k_1} \neq m^{k_2}$ for $k_1 \neq k_2$, then there exist q^{k_1} and q^{k_2} such that $q^{k_1} = m^{k_2}$ and $q^{k_2} = m^{k_1}$, i.e., specific elements $R_{m^{k_1},j}$ and $R_{q^{k_2},j}$ (also $R_{m^{k_2},j}$ and $R_{q^{k_1},j}$) are the same element of the relational matrix. Then, the elements $R_{m^k,j}$ to be optimized in (41) can also be involved in constraints $\Gamma_{m^k,q^k,j}$. In these cases, these constraints are not single constants to be computed using Equation (42); They are function of the elements to be optimized and, consequently, the respective inequalities $\Gamma_{m^k,q^k,j} \leq R_{m^k,j}$ in (41) should be manipulated to be numerically implemented. For example, the constraint $\Gamma_{m^{k_1},m^{k_2},j}$ is given by

$$\Gamma_{m^{k_1},m^{k_2},j}(k) = \frac{X_{m^{k_2}}(k) R_{m^{k_2},j}}{X_{m^{k_1}}(k)} \tag{43}$$

and the inequality $\Gamma_{m^{k_1},m^{k_2},j} \leq R_{m^{k_1},j}$ should be rewritten as

$$X_{m^{k_2}}(k)R_{m^{k_2},j} - X_{m^{k_1}}(k)R_{m^{k_1},j} \leq 0 \quad (44)$$

i.e., with two optimization variables.

Finally, if the refinement procedure does not provide satisfactory results according to some criteria for model evaluation, then the structure of the fuzzy relational matrix can be changed. Roughly speaking, since the relational matrix of the model to be refined is by definition non-optimal, the information about the elements which maximize the composition for $k = 1, \cdots, N$ may be incorrect. In this case, a methodology can be used after the solution of Problem (41) to substitute the elements which maximize the composition in such a way that the cost functions J_j ($j = 1, \cdots, c_0$) are decreased. This methodology is explained using the following example and subsequently a generic algorithm is presented.

Example: Consider a specific element $R_{\hat{m},j}$ of the column j of the relational matrix which maximizes the fuzzy composition, for example, at $k = k_1$,k_2,k_3. Then, according to Equation (36),

$$\hat{m} = m^{k_1} = m^{k_2} = m^{k_3} \quad (45)$$

For every $k = k_1, k_2, k_3$, there is a set of $(l-1)$ constraints $\Gamma_{m^k,q^k,j}$ in (41), but only the greatest one can restricts the value of $R_{\hat{m},j}$. After the solution of (41) all constraints $\Gamma_{m^k,q^k,j}$ are known and can be computed by Equation (42) using R derived from (41). For each k, the greatest constraint is given by:

$$\begin{aligned} &\bar{\Gamma}_{m^k,j}(k) = \max_{q^k} \Gamma_{m^k,q^k,j}(k), \\ &q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \quad (46)$$

Suppose that $\bar{\Gamma}_{m^{k_1},j} < \bar{\Gamma}_{m^{k_2},j} < \bar{\Gamma}_{m^{k_3},j}$ in this example. Suppose also that $R_{\hat{m},j}$ was bounded below in (41). This means that $R_{\hat{m},j} = \bar{\Gamma}_{m^{k_3},j}$ after the solution of (41). From this equality and Equations (46) and (42),

$$\begin{aligned} R_{\hat{m},j} \;&=\; \max_{q^{k_3}} \Gamma_{m^{k_3},q^{k_3},j}(k_3) \;= \\ &\max_{q^{k_3}} \left(\frac{X_{q^{k_3}}(k_3) R_{q^{k_3},j}}{X_{m^{k_3}}(k_3)} \right) \;\triangleq\; \frac{X_{\hat{q}^{k_3}}(k_3) R_{\hat{q}^{k_3},j}}{X_{m^{k_3}}(k_3)} \end{aligned} \quad (47)$$

where

$$\begin{aligned} &\hat{q}^{k_3} = arg \max_{q^{k_3}} \Gamma_{m^{k_3},q^{k_3},j}(k_3) \\ &q^{k_3} = 1, 2, \cdots, m^{k_3} - 1, m^{k_3} + 1, \cdots, l \end{aligned} \quad (48)$$

Equation (47) can be rewritten as

$$X_{\hat{m}}(k_3) R_{\hat{m},j} \;=\; \frac{X_{\hat{m}}(k_3) X_{\hat{q}^{k_3}}(k_3) R_{\hat{q}^{k_3},j}}{X_{m^{k_3}}(k_3)} \quad (49)$$

From Equation (45) the equality $X_{\hat{m}}(k_3) = X_{m^{k_3}}(k_3)$ holds and Equation (49) becomes

$$X_{\hat{m}}(k_3) R_{\hat{m},j} = X_{\hat{q}^{k_3}}(k_3) R_{\hat{q}^{k_3},j} \quad (50)$$

By analyzing Equation (50) it can be noted that decreasing $R_{\hat{m},j}$ the element $R_{\hat{q}^{k_3},j}$ of the relational matrix automatically starts to maximize the composition at $k = k_3$, but the maximum value ($X_{\hat{q}^{k_3}} R_{\hat{q}^{k_3},j}$) of the composition remains constant. Hence, the component of the cost function J_j related to $k = k_3$ does not change when $R_{\hat{m},j}$ is decreased. Moreover, since $X_{\hat{m}} R_{\hat{m},j}$ decreases when $R_{\hat{m},j}$ is decreased, $R_{\hat{m},j}$ does not start to maximize the composition at $k \neq k_1, k_2, k_3$. Hence, the components of the cost function J_j related to $k \neq k_1, k_2, k_3$ do not change either. Consequently, only the components of J_j related to $k = k_1$ and $k = k_2$ changes. Therefore, J_j can be minimized (over $R_{\hat{m},j}$) by minimizing only its components which are related to $k = k_1$ and $k = k_2$ inside the range so that $R_{\hat{m},j}$ decreases without stopping to maximize the composition at $k = k_1$ and $k = k_2$, as follows[4]:

$$\begin{aligned} &\min_{R_{m^k,j}} \quad J_j' = \frac{1}{2} \sum_{k=k_1,k_2} (GY_j(k+1) - X_{m^k}(k) R_{m^k,j})^2 \\ &\text{s. to} \qquad\qquad \bar{\Gamma}_{m^{k_2},j}(k_2) \leq R_{m^k,j} \leq \bar{\Gamma}_{m^{k_3},j}(k_3) \end{aligned} \quad (51)$$

The optimal solution $R^*_{\hat{m},j}$ ($= R^*_{m^k,j}$ for $k = k_1, k_2, k_3$) is obtained from

$$\frac{\partial J_j'}{\partial R_{m^k,j}} = \sum_{k=k_1,k_2} (X^2_{m^k}(k) R_{m^k,j} - G X_{m^k}(k) Y_j(k+1)) = 0 \quad (52)$$

Then, $R^*_{\hat{m},j}$ is given by

$$R^*_{\hat{m},j} = \begin{cases} \bar{\Gamma}_{m^{k_2},j}(k_2), & \text{if } \dfrac{G \sum_{k=k_1,k_2} X_{m^k}(k) Y_j(k+1)}{\sum_{k=k_1,k_2} X^2_{m^k}(k)} \leq \bar{\Gamma}_{m^{k_2},j}(k_2) \\ \bar{\Gamma}_{m^{k_3},j}(k_3), & \text{if } \dfrac{G \sum_{k=k_1,k_2} X_{m^k}(k) Y_j(k+1)}{\sum_{k=k_1,k_2} X^2_{m^k}(k)} \geq \bar{\Gamma}_{m^{k_3},j}(k_3) \\ \dfrac{G \sum_{k=k_1,k_2} X_{m^k}(k) Y_j(k+1)}{\sum_{k=k_1,k_2} X^2_{m^k}(k)}, & \text{otherwise} \end{cases} \quad (53)$$

Afterwards, if $R^*_{\hat{m},j} = \bar{\Gamma}_{m^{k_2},j}$ (i.e. if $R^*_{\hat{m},j}$ is bounded below) then the present context is equivalent to the initial context of the example and a similar analysis can be done only for $k = k_1, k_2$. In this way,

$$R^*_{\hat{m},j} = \begin{cases} \bar{\Gamma}_{m^{k_1},j}(k_1), & \text{if } \frac{GY_j(k_1+1)}{X_{m^{k_1}}(k_1)} \leq \bar{\Gamma}_{m^{k_1},j}(k_1) \\ \bar{\Gamma}_{m^{k_2},j}(k_2), & \text{if } \frac{GY_j(k_1+1)}{X_{m^{k_1}}(k_1)} \geq \bar{\Gamma}_{m^{k_2},j}(k_2) \\ \frac{GY_j(k_1+1)}{X_{m^{k_1}}(k_1)}, & \text{otherwise} \end{cases} \quad (54)$$

In this example it can be seen that after the solution of the quadratic problem (41) specific elements of the relational matrix

[4] The inequality $R_{m^k,j} \geq \bar{\Gamma}_{m^{k_1},j}$ is not taken into consideration in Problem (51) because it is redundant since $\bar{\Gamma}_{m^{k_1},j}$ is lower than $\bar{\Gamma}_{m^{k_2},j}$.

(i.e. elements which were bounded below) can be optimized in such a way that they stop to maximize the composition at some time instants, reducing the values of cost functions J_j. The algorithm is summarized below.

Algorithm for changing the relational matrix structure

Given a fuzzy relational matrix derived from (41), its structure can be changed using the following basic algorithm:

a) For $j = 1$ to c_0 execute all operations until Step **j**.

b) For $i = 1$ to l execute all operations until Step **j**.

c) If the element $R_{i,j}$ maximizes the fuzzy composition at least at one $k \in \{1, \cdots, N\}$ (i.e. if there exist $k \in \{1, \cdots, N\}$ such that $i = m^k$, where m^k is given by (36)) then execute all operations until Step **j**, else do nothing.

d) Find the set $\mho = \{k_1, \cdots, k_\rho\}$ of time instants k for which $R_{i,j}$ maximizes the composition, i.e., the instants $k = k_1, \cdots, k_\rho$ for which $i = m^k$ ($i = m^{k_1} = \cdots = m^{k_\rho}$).

e) For each element of $\mho$ compute the greatest constraint $\Gamma_{m^k,q^k,j}$, i.e., for $k = k_1, \cdots, k_\rho$ compute

$$\begin{aligned} \bar{\Gamma}(k) &= \max_{q^k} \Gamma_{m^k,q^k,j}(k), \\ q^k &= 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{55}$$

where the constraints $\Gamma_{m^k,q^k,j}$ are given by Equation (42).

f) Construct a ordered set $\{\bar{\Gamma}_1, \cdots, \bar{\Gamma}_\rho\}$ with $\bar{\Gamma}(k)$ such that $\bar{\Gamma}_1 \geq \bar{\Gamma}_2 \geq \cdots \geq \bar{\Gamma}_\rho$, where

$$\bar{\Gamma}_1 = \max_{k \in \mho} \bar{\Gamma}(k) \tag{56}$$

$$\bar{\Gamma}_\rho = \min_{k \in \mho} \bar{\Gamma}(k) \tag{57}$$

g) For $\beta = 1$ to $\rho - 1$ execute all operations until Step **j**.

h) If $R_{i,j} = \bar{\Gamma}_\beta$ execute all operations until Step **j**, else do nothing.

i) Find the solution $R^*_{m^k,j}$ of the following optimization problem:

$$\begin{aligned} \min_{R_{m^k,j}} \quad & J_j^{'} = \frac{1}{2} \sum_{k \in \mho} (GY_j(k+1) - X_{m^k}(k) R_{m^k,j})^2 \\ \text{s. to} \quad & \bar{\Gamma}_{\beta+1} \leq R_{m^k,j} \leq \bar{\Gamma}_\beta \end{aligned} \tag{58}$$

j) Compute $R_{i,j} = R^*_{m^k,j}$.

If the the values of the cost functions J_j ($j = 1, \cdots, c_0$) are reduced after the execution of the algorithm presented above, then Problem (41) can be solved again using the new fuzzy relational matrix as an initial condition. Otherwise, the refinement has achieved its best possible solution. However, if there are elements of the final optimized matrix which never maximizes the composition for the available data set, then these elements can be made equal zero since they do not influence the model output. In this case, some constraints of (41) are relaxed and the refinement can go on.

3.3 Recursive algorithm

An algorithm for recursive identification of fuzzy relational models can be derived from the issues presented in the last section as follows. Starting from an initial estimate of the fuzzy relational matrix and taking into consideration only a single input-output data pair ($X(k)$ and $Y(k+1)$), i.e. $N = 1$, Problem (40) is rewritten as

$$\begin{aligned} \min_{R_{m^k,j}} \quad & J_j = \frac{1}{2}(GY_j(k+1) - X_{m^k}(k) \tau R_{m^k,j})^2 \\ \text{s. to} \quad & \Gamma_{m^k,q^k,j}(k) \leq R_{m^k,j} \leq 1 \\ & q^k = 1, 2, \cdots, m^k - 1, m^k + 1, \cdots, l \end{aligned} \tag{59}$$

where $j = 1, \cdots, c_0$ and $R_{m^k,j}$ is the element of the column j of the relational matrix which maximizes the fuzzy composition at k. Unlike Problem (40), since only one instant k is taken into consideration in (59) the use of $l - 1$ constraints is redundant because the greatest one ($\bar{\Gamma}_{m^k,j}$) is known. Thus, Problem (59) can be rewritten as

$$\begin{aligned} \min_{R_{m^k,j}} \quad & J_j = \frac{1}{2}(GY_j(k+1) - X_{m^k}(k) \tau R_{m^k,j})^2 \\ \text{s. to} \quad & \bar{\Gamma}_{m^k,j}(k) \leq R_{m^k,j} \leq 1 \end{aligned} \tag{60}$$

where $\bar{\Gamma}_{m^k,j}$ (given by Equation (46)) is the smallest value of $R_{m^k,j}$ for which $X_{m^k} \tau R_{m^k,j}$ remains the maximum element of the composition.

Using the penalty function method (Bazaraa and Shetty, 1979), the constrained optimization problem (60) is replaced by an unconstrained one, as follows:

$$\min_{R_{m^k,j}} \quad J_j^{'} = \frac{1}{2}(GY_j(k+1) - X_{m^k}(k) \tau R_{m^k,j})^2 + J_{p_j} \tag{61}$$

with

$$\begin{aligned} J_{p_j} = \frac{\Omega}{2} \Big(& [\, \max(0\, , R_{m^k,j} - 1) \,]^2 + \\ & + [\, \max(0\, , \bar{\Gamma}_{m^k,j}(k) - R_{m^k,j}) \,]^2 \Big) \end{aligned} \tag{62}$$

where $\frac{\Omega}{2}$ is the gain of the penalty function J_{p_j} which must be large enough in relation to the original cost function J_j in (60). Since J_j in (60) belongs to the interval $[0, \frac{1}{2}]$, the condition $\Omega > 1$ should be satisfied.

The fuzzy relational matrix can be updated using the gradient method as

$$R^*_{m^k,j} = R_{m^k,j} - \eta \frac{\partial J'_j}{\partial R_{m^k,j}}, \qquad j = 1, \cdots, c_0 \quad (63)$$

where $R^*_{m^k,j}$ is the new value of $R_{m^k,j}$ and η is the learning rate of the method. The gradient is given by

$$\frac{\partial J'_j}{\partial R_{m^k,j}} =$$
$$(X_{m^k}(k)\tau R_{m^k,j} - GY_j(k+1))\frac{\partial(X_{m^k}(k)\tau R_{m^k,j})}{\partial R_{m^k,j}} +$$
$$+ \Omega \max\left(0\,,\, R_{m^k,j} - 1\right) - \Omega \max\left(0\,,\, \bar{\Gamma}_{m^k,j}(k) - R_{m^k,j}\right) \quad (64)$$

Equations (63) and (64) are a generalization of the adaptation law proposed in Campello and Amaral, 1998, where the t-norm τ used is the product and the gradient is given by

$$\frac{\partial J'_j}{\partial R_{m^k,j}} =$$
$$X^2_{m^k}(k)R_{m^k,j} - GX_{m^k}(k)Y_j(k+1) +$$
$$+ \Omega \max\left(0\,,\, R_{m^k,j} - 1\right) - \Omega \max\left(0\,,\, \bar{\Gamma}_{m^k,j}(k) - R_{m^k,j}\right) \quad (65)$$

It is important to note that the algorithm updates the relational matrix recursively and takes the constraints given by an estimate of this matrix into consideration. This estimate is given by the relational matrix itself at the previous time instant, but its information remains reliable at the present instant only if there are no abrupt changes in the system. Thus, the algorithm is well-suited for the identification of slowly time-variant systems, i.e., systems which have no abrupt changes in their dynamics. For the identification of generic time-varying systems a more complex adaptive algorithm can be utilized, as that one developed in previous work (Campello et al., 1997, Campello et al., 1998).

4 NUMERICAL EXAMPLE

The gas furnace of Box and Jenkins, 1970 is used to evaluate the performance of the proposed methods in the identification of a dynamic system. The data set consists of 296 input/output data pairs observed from the process. The input (u) is methane gas feed rate and the output (y) is carbon dioxide concentration in a mixture of gases. The relational model used for the identification of this process follows the structure investigated in Pedrycz, 1984 and Lee et al., 1994, i.e.

$$Y(k+1) = Y(k) \bullet U(k-2) \bullet R \quad (66)$$

where "$\bullet$" denotes the max-product composition (t-norm τ is the product). In this example, ψ is the product t-norm and the interfaces have $c_0 = c_1 = c_2 = 5$ reference fuzzy sets with Gaussian membership functions optimized by the Σ-PAFIO algorithm (Oliveira, 1995). The defuzzification method used is the weighted average (see Equation (22)). To get round the well-known difficulty of the method to defuzzify extreme values, the output interface is expanded (10%) over its universe of discourse.

A random initial fuzzy relational matrix is used to evaluate the performance of the recursive algorithm presented in Section 3.3 when a model to be refined is not available. This matrix is considered to be the iteration zero ($\varphi = 0$) of the identification procedure, since there is no identification associated with it, and is set up as the initial condition for the recursive algorithm. The choice of the recursive algorithm parameters is as follows. The gain of the penalty function method should be greater than 0.5 (see Section 3.3). The choice of this gain is not critical and in this example it is set equal to 2. The value of the learning rate of the gradient method basically changes the convergence rate of the algorithm. A high value (0.9) is assigned to it in this example since the algorithm starts the identification procedure from a random relational matrix (i.e. without an initial model). In this algorithm the data set is recursively used four times, i.e. $\varphi = 1, 2, 3$ and 4, where each iteration φ means the utilization of the entire data set. The algorithm is stopped at $\varphi = 4$ because beyond this iteration it can not provide significant improvement of the model anymore. More specifically, the mean squared error (MSE) between the non-fuzzy outputs of the process and model can not be reduced with rate greater than a previously established threshold of 5% between two iterations φ. After the execution of the recursive algorithm, the resulting model is refined using the refinement algorithm presented in Section 3.2 until a satisfactory fine-tune is achieved ($\varphi = 5, \cdots, 9$).

The evolution of the MSE between the non-fuzzy outputs of the process (y) and model ($\tilde{y}$) throughout the iterations φ is illustrated in Figure (3). In this figure the convergent behavior of both the proposed algorithms can be observed.

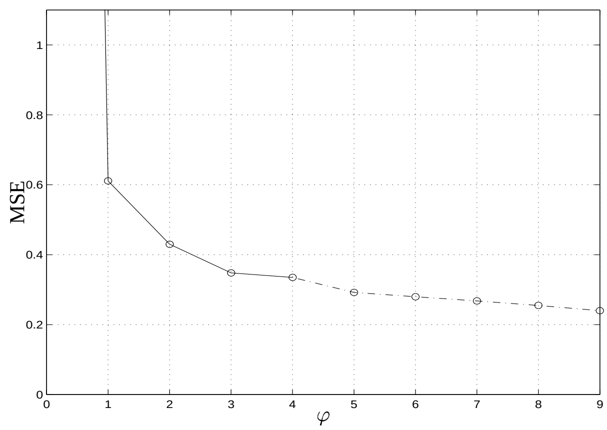

Figure 3: Evolution of the mean squared error between y and $\tilde{y}$: Recursive algorithm (solid line) and refinement algorithm (dashed line).

Figure (4) displays the output of the system together with the output of the model for one-step-ahead prediction and synthetic data (also called recursive or open-loop simulation). The synthetic data are generated to evaluate the generalization of the model, since in this kind of simulation there is feedback of the prediction errors. Figure (4) illustrates the accuracy of the model obtained using the proposed algorithms, showing that an efficient identification can be performed even starting from a random relational matrix. Despite this, in the case of hard user requirements these results can be significantly improved using a more complex model with two samples of the input signal instead just one (as in Sugeno and Yasukawa, 1993), that is $Y(k+1) = Y(k) \bullet U(k-2) \bullet U(k-3) \bullet R$. In this case, however, the dimensions of the model increase, especially the size of the relational matrix.

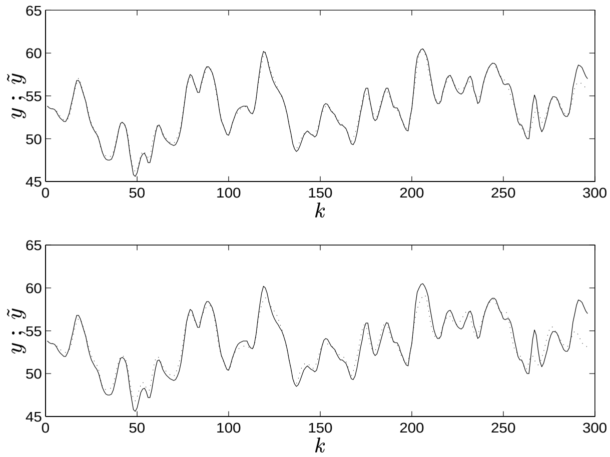

Figure 4: System output (solid line) and model output (dotted line): One-step-ahead prediction (above) and synthetic data (below).

5 CONCLUSIONS

Two numerical methods for identification of fuzzy relational models have been proposed. The first one is an optimization-based methodology which solves a set of quadratic problems to refine rough estimative of relational matrices. The second one is a recursive algorithm which is derived as a simplification of the first method and is well-suited for on-line applications. The proposed methods were evaluated using a real dynamic system. The recursive algorithm provided a model with reasonably small prediction errors, even starting the identification from a random initial relational matrix. The refinement algorithm provided a fine-tune for the model identified using the recursive algorithm, showing that it can improve relational models previously identified by other methods. These results illustrate that efficient models can be derived to be used in several application fields such as, for example, system monitoring, time series forecasting and (adaptive/predictive) model-based control. In the latter, specifically, better models can lead to more accurate control strategies and, consequently, better performances for the closed-loop systems.

References

Bazaraa, M. S. and Shetty, C. M. (1979). *Nonlinear Programming Theory and Algorithms*. John Wiley & Sons.

Box, G. and Jenkins, G. (1970). *Time Series Analysis, Forecasting and Control*. Holden Day.

Campello, R. J. G. B. and Amaral, W. C. (1998). Refinement and identification of fuzzy relational models. In *Proc. 7th IEEE Internat. Conference on Fuzzy Systems*, pages 651–656, Anchorage/USA.

Campello, R. J. G. B. and Amaral, W. C. (1999). Extracting linguistic knowledge from fuzzy relational models. In *Proc. 8th IFSA World Congress*, Taipei/Taiwan (to be published).

Campello, R. J. G. B., Nazzetta, R. M., and Amaral, W. C. (1997). A new methodology for fuzzy model identification. In *Proc. 7th IFSA World Congress*, pages 366–370, Prague/Czech Republic.

Campello, R. J. G. B., Nazzetta, R. M., and Amaral, W. C. (1998). A highly adaptive algorithm for fuzzy modelling of systems. *Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems*, 6:35–50.

Czogała, E. and Pedrycz, W. (1981). On identification in fuzzy systems and its applications in control problems. *Fuzzy Sets and Systems*, 6:73–83.

Graham, B. P. and Newell, R. B. (1989). Fuzzy adaptive control of a first-order process. *Fuzzy Sets and Systems*, 31:47–65.

Kosko, B. (1992). *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall.

Kosko, B. (1997). *Fuzzy Engineering*. Prentice Hall.

Lee, C. C. (1990). Fuzzy logic in control systems: Fuzzy logic controller - Part 1. *IEEE Trans. Systems, Man and Cybernetics*, 20:404–418.

Lee, Y. C., Hwang, C., and Shih, Y. P. (1994). A combined approach to fuzzy model identification. *IEEE Trans. Systems, Man and Cybernetics*, 24:736–743.

Ljung, L. (1987). *System Identification: Theory for the user*. Prentice Hall.

Oliveira, J. V. (1995). A design methodology for fuzzy system interfaces. *IEEE Trans. Fuzzy Systems*, 3:404–414.

Pedrycz, W. (1984). An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems*, 13:153–167.

Pedrycz, W. (1993). *Fuzzy Control and Fuzzy Systems*. Research Studies Press/John Wiley & Sons, 2nd edition.

Pedrycz, W. (1995). *Fuzzy Sets Engineering*. CRC Press.

Pedrycz, W. and Gomide, F. (1998). *An Introduction to Fuzzy sets. Analysis and Design*. MIT Press.

Pedrycz, W. and Oliveira, J. V. (1996). An algorithmic framework for development and optimization of fuzzy models. *Fuzzy Sets and Systems*, 80:37–55.

Sudkamp, T. and Hammell II, R. J. (1994). Interpolation, completion, and learning fuzzy rules. *IEEE Trans. Systems, Man and Cybernetics*, 24:332–342.

Sugeno, M. and Yasukawa, T. (1993). A fuzzy-logic-based approach to qualitative modeling. *IEEE Trans. Fuzzy Systems*, 1:7–31.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Systems, Man and Cybernetics*, SMC-15:116–132.

Tong, R. M. (1978). Synthesis of fuzzy models for industrial processes-some recent results. *Int. J. General Systems*, 4:143–162.

Wang, L. and Langari, R. (1996). Complex systems modeling via fuzzy logic. *IEEE Trans. Systems, Man and Cybernetics*, 26:100–106.

Wang, L. and Mendel, J. M. (1992). Fuzzy basis functions, universal approximation and orthogonal least squares learning. *IEEE Trans. Neural Networks*, 3:807–814.

Willaeys, D. and Malvache, N. (1981). The use of fuzzy sets for the treatment of fuzzy information by computer. *Fuzzy Sets and Systems*, 5:323–327.

Xu, C. W. and Lu, Y. Z. (1987). Fuzzy model identification and self-learning for dynamic systems. *IEEE Trans. Systems, Man and Cybernetics*, SMC-17:683–689.

Yager, R. R. and Filev, D. P. (1994). *Essentials of Fuzzy Modeling and Control*. John Wiley & Sons.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man and Cybernetics*, SMC-3:28–44.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28.

APPENDIX: FUZZY DISCRETIZATION

The fuzzy discretization technique for data representation in fuzzy systems was proposed by Willaeys and Malvache, 1981 and have been widely utilized in fuzzy model identification (for instance see (Pedrycz, 1984) and (Wang and Langari, 1996)). This technique is outlined in this appendix as follows.

Let $\mathcal{X}_1, \cdots, \mathcal{X}_\eta$ be fuzzy sets, called reference fuzzy sets, defined in the universe of discourse $\mathbf{X}$ of the non-fuzzy variable x, i.e.

$$\mathcal{X}_i : \mathbf{X} \to [0,1], \qquad i = 1, \cdots, \eta \tag{67}$$

such that they satisfy the frame of cognition concept[5], especially the condition of completeness:

$$\forall x \in \mathbf{X} \quad \exists i : 1 \le i \le \eta : \quad \mathcal{X}_i(x) > 0 \tag{68}$$

Then, according to the fuzzy discretization concept any fuzzy set X in $\mathbf{X}$ can be represented by a possibility vector as follows:

$$P_x = [\, P_{x_1} \ \ \cdots \ \ P_{x_\eta} \,] \tag{69}$$

where P_{x_i} represents a possibility measure (Zadeh, 1978) of X with respect to the i-th reference fuzzy set $\mathcal{X}_i$, that is

$$\begin{aligned} P_{x_i} \ &= \ \text{Poss}(\, X | \mathcal{X}_i \,) \ \stackrel{\triangle}{=} \ \sup_{x \in \mathbf{X}} [\, X(x) \ \ \text{t} \ \ \mathcal{X}_i(x) \,]\,, \\ i &= 1, \cdots, \eta \end{aligned} \tag{70}$$

with “t” being a triangular norm.

The formulation presented above means that a non-fuzzy variable $x \in \mathbf{X}$ can be fuzzified generating a fuzzy set X, and further, it can be represented by means of reference fuzzy sets $\mathcal{X}_1, \cdots, \mathcal{X}_\eta$ using equations (69) and (70). If the fuzzification is implemented using the (non-fuzzy) singleton, then for a given value x':

$$X(x) \ = \ \begin{cases} 1, & \text{if} \quad x = x' \\ 0, & \text{otherwise} \end{cases} \tag{71}$$

[5]The frame of cognition (Pedrycz, 1995, Pedrycz and Gomide, 1998) is a set of conditions which determines that a collection of fuzzy sets associated with a linguistic variable is semantically interpretable, having clear linguistic meaning.

Due to the boundary conditions of the triangular norms, i.e., $0 \ \text{t} \ a = 0$ and $1 \ \text{t} \ a = a$ ($\forall\, a \in [0,1]$), equations (70) and (71) yield

$$P_{x_i} \ = \ \mathcal{X}_i(x') \tag{72}$$

and consequently

$$P_x \ = \ [\, \mathcal{X}_1(x') \ \ \cdots \ \ \mathcal{X}_\eta(x') \,] \tag{73}$$

where $\mathcal{X}_i(x')$ is the grade of membership of the numerical value x' in relation to the i-th reference fuzzy set $\mathcal{X}_i$.

Since $\mathcal{X}_i(x') \in [0,1]$ ($i = 1, \cdots, \eta$), the possibility vector (73) is, by definition, a fuzzy set (with η discretization elements). Then, since generally $\eta \le 9$ (because a linguistic variable is commonly associated with 7 ± 2 fuzzy sets (Pedrycz, 1995)) it is possible to represent linguistically a non-fuzzy variable x by means of a fuzzy set P_x using a small number of parameters. As a consequence, the dimensions of fuzzy models can be significantly reduced by the use of the fuzzy discretization technique.