Identification of fuzzy models

Easter TAN, Gilles MOUROT, Didier MAQUIN, José RAGOT

Centre de Recherche en Automatique de Nancy - CNRS URA 821
B.P. 3 - 2, avenue de la Forêt de Haye
France - 54 516 Vandoeuvre-Les-Nancy Cedex
Tél.: (33) 83 59 59 59 Fax: (33) 83 59 56 44
Email: jragot@ensem.u-nancy.fr

ABSTRACT
The objective of this work is to describe a numerical technique to identify parameters of a fuzzy model. When this model and the membership functions of the input variables are continuously differentiable, we show that the estimation of parameters can be performed with a two-levels hierarchical algorithm, comprising the estimation of the model parameters and the estimation of the membership functions parameters. The proposed algorithm is then applied to an example describing a non-linear system.

KEYWORDS
Fuzzy model, identification, parameter estimation.

1. INTRODUCTION

Building of fuzzy models can be applied to fuzzy control and to modeling of complex systems. Moreover, the implementation of a fuzzy logic controller involves the design of linguistic rules to compute the command to be applied on a process according to the measurement of the position error. In simple situations, these rules can be derived from those of a conventional PID controller ; on the other hand, for non-linear systems or for time-varying systems, the design of these rules needs an appropriate modeling. For the systems where the mathematical structures are unknown, fuzzy modeling can constitute an interesting approach since it uses multiple smoothly connected local input-output relations. The problem of fuzzy modeling has received little attention until now. Indeed, the principles of system modeling have been presented by [Tong, 1979]; a method of parameter estimation can be found in [Takagi and Sugeno, 1985] and the problem of structure identification of fuzzy models is addressed in [Sugeno and Kang, 1988 ; Hihi and Richard, 1993].

This paper doesn't deal with the structure identification problem but focuses on the problem of parameter identification of a class of fuzzy model. Section 2 describes this class of fuzzy model. Section 3 formulates the problems and shows the proposed solution. Section 4 describes an example.

2. FUZZY MODEL

2.1. SOME FUZZY CONCEPTS

Let us define some useful notations. X is a reference set which elements are noted x; $\mu_A(x)$ represents the membership function of a fuzzy subset A, $x \in X$. In the following, we use gaussian membership functions which are characterized by two parameters giving the greatest grade 1 and the least grade 0. To deal with several fuzzy sets, for instance A and B, we define operations on sets : the value of proposition "x is A and y is B" is characterized by its membership degre V:

$$V\ (x \text{ is } A \text{ and } y \text{ is } B) = T(\mu_A(x)\ ,\ \mu_B(y)) \qquad (1)$$

where the operator T is a triangular norm. In particular, t-norm "min" is the logical AND. It is possible to use others norms, such as algebrical product, but it is a priori difficult to justify the choice of a particular norm.

In this paper, we only deal with static multi-input/multi-output process ; command variables vector is $x = \{ x_1 , ... , x_i , ... , x_n \}$ and output variables vector is $y = \{ y_1 , ... , y_i , ... , y_m \}$. Each variable x_i is fuzzified ; this operation consists of three different steps. First, the choice of the universe of discourse which can be determined knowing the values x_{imin} and x_{imax} corresponding to the variation interval of x_i. The number of linguistic terms n_i must then be fixed for each variable. This information may be given by an expert of the system or by a pattern recognition technique. Each level is then associated with a linguistic qualifier. Finally, the third step consists in choosing the structure of the membership function for each linguistic term. Several types of membership functions can be used and it is a common and convenient practice to use triangular or trapezoidal functions. In our case, we chose a gaussian function which has the advantage of being continuous on all the variation interval (figure 1). In the following, these functions are denoted $\mu_H(x,q)$ where q represents the parameters (nodal value and width) of this function.

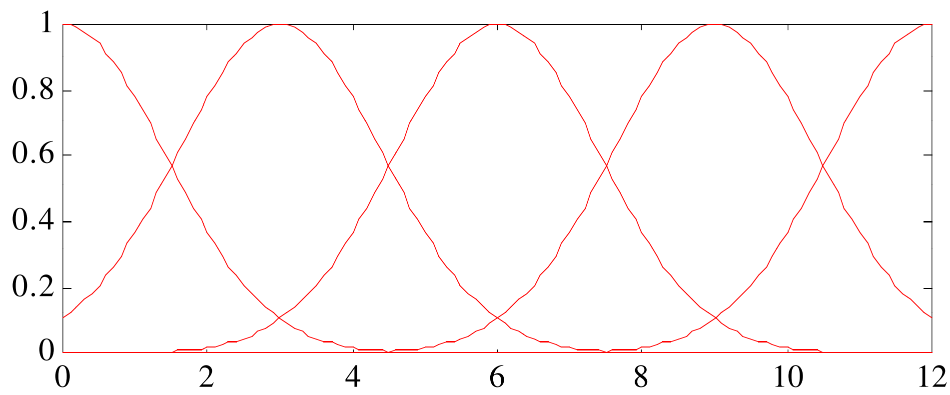

Figure 1: The gaussian membership functions.

2.2. FUZZY MODEL

We consider a fuzzy model consisting of a set of linguistic rules. For example, the following model is defined with three input variables (x_1, x_2 and x_3), single output variable y, and two rules that are of the format:

R_1 : if x_1 is $small_1$ and if x_2 is big_2 then $y = x_1 + x_2 + 2\,x_3$

R_2 : if x_1 is $medium_1$ and if x_2 is $small_2$ then $y = 3\,x_1 + 2\,x_2$

We shall notice that the qualifiers "small", "medium" and "big" are different for each variable ; for example $small_1$ for x_1 and $small_2$ for x_2 do not have necessarilly the same meaning.

A model is determined by a set of rules and in general only part of the variables appears in each rule. An inference table can be defined for each variable. For example, table 1 presents the inference table for x_i : it comprises five rules and shows that x_i can receive four qualifiers A_i and that a rule doesn't use twice the same variable. It is also possible to define the list $L_i = \{ 1, 2, 1, 4, 3 \}$ which indicated the qualifiers of variable x_i in the different rules R_j.

	A_{i1}	A_{i2}	A_{i3}	A_{i4}
R_1	X			
R_2		X		
R_3	X			
R_4				X
R_5			X	

Table 1: The inference table for x_i.

When a variable j is used let us say in rules 1, 3, and 4 with respective qualifiers 1, 3, and 2, but not in rules 2 and 5, the notation $L_j = \{ 1, \varnothing, 3, 2, \varnothing \}$ is then utilized.

More generally, the following set of rules can be defined:

$R = \{ R_1, ..., R_i, ..., R_M \}$

R_i : if $(x_1 \text{ is } A_{L1(i),1}, x_2 \text{ is } A_{L2(i),2}, ..., x_v \text{ is } A_{Lv(i),v})$

then $(y_1 \text{ is } B_{L1(i),1}, y_2 \text{ is } B_{L2(i),2}, ..., y_w \text{ is } B_{Lw(i),w})$

$v = 1, ..., n$ $w = 1, ..., m$ $i = 1, ..., M$

A_{ij} the qualifiers of the variable x

B_{ij} the qualifiers of the variable y

y_j the consequence variables, $j = 1, ..., m$

x_j the premise variables, j = 1, ..., n
A_i the fuzzy subsets
g_i the function that implies the value of y_i

Each rule R_i can be decomposed such as:

$$R_i = \{ R_{i1}, ..., R_{ij}, ..., R_{iv} \} \tag{2}$$
$$R_{ij} : \text{if} \quad (x_1 \text{ is } A_{L1(i),1}, x_2 \text{ is } A_{L2(i),2}, ..., x_v \text{ is } A_{Lv(i),v})$$
$$\text{then} \quad y = y_{ij} \text{ is } B_{L1(i)}$$

which corresponds to the general n-inputs m-outputs model made with m n-inputs single output models. For this reason, this paper will solely described the study of mono-output system and in the following m = 1 and y_{ij} will be reduced to y_i (i = 1, ..., M).

Semi-fuzzy model

Such a semi-fuzzy model can be obtained when the output variables of the process are represented by a model with non-fuzzy output values. In this particular case, the rules can be noted:

$$R_i : \quad \text{if} \quad (x_1 \text{ is } A_{L1(i),1}, x_2 \text{ is } A_{L2(i),2}, ..., x_v \text{ is } A_{Lv(i),v}) \tag{3}$$
$$\text{then} \quad y_i = f_i(x_1, x_2, ..., x_v, p_i)$$

where p_i represents the parameters vector of the model.

For a given observation of variable x, one can calculate a particular value of y_i from each rule R_i. The evaluation of the M rules yields M values of y_i ; it is therefore necessary to estimate a single value of y from these M values and their grades of membership. It is reasonnable to use the center of gravity of y_i values as defuzzification algorithm. The grade of membership of a rule is defined from the membership functions of variable x_i such as:

$$V_i(y_i) = g_i(\mu_{Ai1}(x_1, q_{Ai1}), \mu_{Ai2}(x_2, q_{Ai2}), ..., \mu_{Aiv}(x_v, q_{Aiv})) \tag{4}$$

This formulation can be simplified by using membership functions which are independant of the evaluated rule. It is then possible to write:

$$V_i(y_i) = g_i(\mu_{A1}(x_1, q_{A1}), \mu_{A2}(x_2, q_{A2}), ..., \mu_{Av}(x_v, q_{Av})) \tag{5a}$$

The case of a composition in the form of a product is particularly simple to use:

$$V_i(y_i) = \mu_{A1}(x_1, q_{A1})\, \mu_{A2}(x_2, q_{A2}) \cdots \mu_{Av}(x_v, q_{Av}) \tag{5b}$$

Using a center of gravity relationship, the value of consequence variable (output or command) is defined by:

$$y = \frac{\sum_{i=1}^{n} V_i(y_i)\, f_i(x_1, x_2, ..., x_v, p_i)}{\sum_{i=1}^{n} V_i(y_i)} \tag{6}$$

To illustrate these notations and the above definitions, let us consider a simple three rules model (M = 3), two premise variables (n = 2), three qualifiers for the first variable and two qualifiers for the other variable:

R_1 : if x_1 is $small_1$ and if x_2 is big_2 then $y_1 = x_1 + x_2$ (7a)

R_2 : if x_1 is $medium_1$ and if x_2 is $small_2$ then $y_2 = 3\,x_1 + 2\,x_1\,x_2$ (7b)

R_3 : if x_1 is big_1 and if x_2 is $small_2$ then $y_3 = 2\,x_1 + 2\log(x_2\,x_1)$ (7c)

The membership functions may be defined such as:

x_1 is $small_1$: $\mu_{small_1}(x_1) = \exp[-(\frac{x}{2})^2]$ (8a)

x_1 is $medium_1$: $\mu_{medium_1}(x_1) = \exp[-(\frac{x-5}{3})^2]$ (8b)

x_1 is big_1 : $\mu_{big_1}(x_1) = \exp[-(\frac{x-10}{2})^2]$ (8c)

x_2 is $small_2$: $\mu_{small_2}(x_2) = \exp[-(\frac{x+5}{2})^2]$ (8d)

x_2 is big_2 : $\mu_{big_2}(x_2) = \exp[-(\frac{x-15}{2})^2]$ (8e)

From a couple of given measurement x_1 and x_2, the calculation of the output y is carried out by evaluating:

the consequence rules from definition (7).

the membership functions of x_1 and x_2 from definition (8).

the grade of membership of each rule from definition (5b).

the output from the center of gravity equation given by definition (6).

3. ESTIMATION OF FUZZY MODEL PARAMETERS

The fuzzy model described by equation (6) uses several parameters which have to be numerically determined. Some of these parameters are related to the structure of the model (number of qualifiers, variation interval of variables, number of rules, number of parameters describing membership functions), others define the values of parameters describing the models and the membership functions. In the following, we emphasized on the latter parameters ; the selection of appropriate structures for the models and for the membership functions is not investigated analytically, but we will mention a few results showing the sensitivity of the estimations to the complexity of these structures.

The principle of identification is simple: it consists in determining the parameters from the N measurement of variables x and y such that the equation (6) is satisfied. The technique used ajust the parameters of the models and the membership functions by minimizing the following least squares criterion:

$$\Phi = \frac{1}{2} \sum_{k=1}^{N} \left(y(k) - \frac{\sum_{i=1}^{n} V_i(y_i)\, f_i(x_1(k), x_2(k), ..., x_v(k), p_i)}{\sum_{i=1}^{n} V_i(y_i)} \right)^2 \qquad (9)$$

The presence of the parameters p_i and q_{Ai} in the functions f and g implies that the problem is strongly non-linear and therefore we suggest a resolution using a two-levels hierarchical algorithm.

Estimation of consequence parameters (first level)

We assume that the parameters q_{Ai} of the premises known. By defining the normalized membership coefficients β_i:

$$\beta_i = \frac{V_i(y_i)}{\sum_{i=1}^{n} V_i(y_i)} \qquad (10)$$

The model (6) can be written:

$$y(k) = \sum_{i=1}^{n} \beta_i\, f_i(x_1(k), x_2(k), ..., x_v(k), p_i) \qquad k = 1, ..., N \qquad (11)$$

In the case of models linear in respect to parameters, equation (11) can be simplified such as:

$$y(k) = \sum_{i=1}^{n} \beta_i\, (p_{i0} + p_{i1}\, x_1(k) + p_{i2}\, x_2(k) + \ ... + p_{iv}\, x_v(k)) \qquad (12)$$

With:

$$p_i = (p_{i0} \;\; p_{i1} \;\; \cdots \;\; p_{iv})^T \qquad (13a)$$

$$p = (p_1 \;\; p_2 \;\; \cdots \;\; p_n)^T \qquad (13b)$$

$$\varphi(k) = (1 \; x_1(k) \; ... \; x_n(k))^T \qquad (13c)$$

and using definition (13), the model (12) is rewritten:

$$y(k) = \varphi^T(x(k))\, p \qquad (14)$$

The estimation of p can be carried out by a classical least squares procedure. We now have a set of statistical tools enabling us to appreciate the validity of the model structure (accuracy of parameters, descriptive variable power, ...) and the coherence of the pseudo-measurement φ.

Estimation of premise parameters (second level)
The consequence parameters are assumed to be determined by the first level. Equation (6) can be rewritten:

$$\sum_{i=1}^{n} V_i(y_i)\, \big(y(k) - f_i(x_1(k), x_2(k), ..., x_v(k), p_i)\big) = 0 \qquad (15)$$

In this expression, we remind that V_i depends on the membership function of x for the different models (equation 5a). If the algebrical product is used to evaluate the membership function (equation 5b), the left member of equation (15) is continuously derivable to the parameters which described the membership functions. It is then possible to satisfy equation (15) with a least squares criterion for all the measurement ($k = 1, ..., N$) ; the gradient of the objective function can be simply explicited and the determination of the parameters can be performed using a Newton type technique.

In summary, the method is composed of an initialisation and three different steps:

Initialisation: choose the number of rules and the initial values of the parameters of the membership function
Step 1: estimation of model parameters
Step 2: estimation of parameters of membership function
Step 3: if convergence, end of the procedure
otherwise go to the step 1

4. EXAMPLES
Figure 2a presents the case of a two variables relationship. The 31 experimental values are represented by open circles ; the curve correspond to the results obtained with a model identified according to the above procedure by choosing three rules and three qualifiers for the variable x. As the variation interval of x is known, the membership functions of the qualifiers "small" and "big" are centred to the extremities of this interval. Therefore ten parameters are to be identified, four for the membership function and six for the consequence variables. Practically, for this example studied, the convergence of the estimation is achieved with twenty iterations.

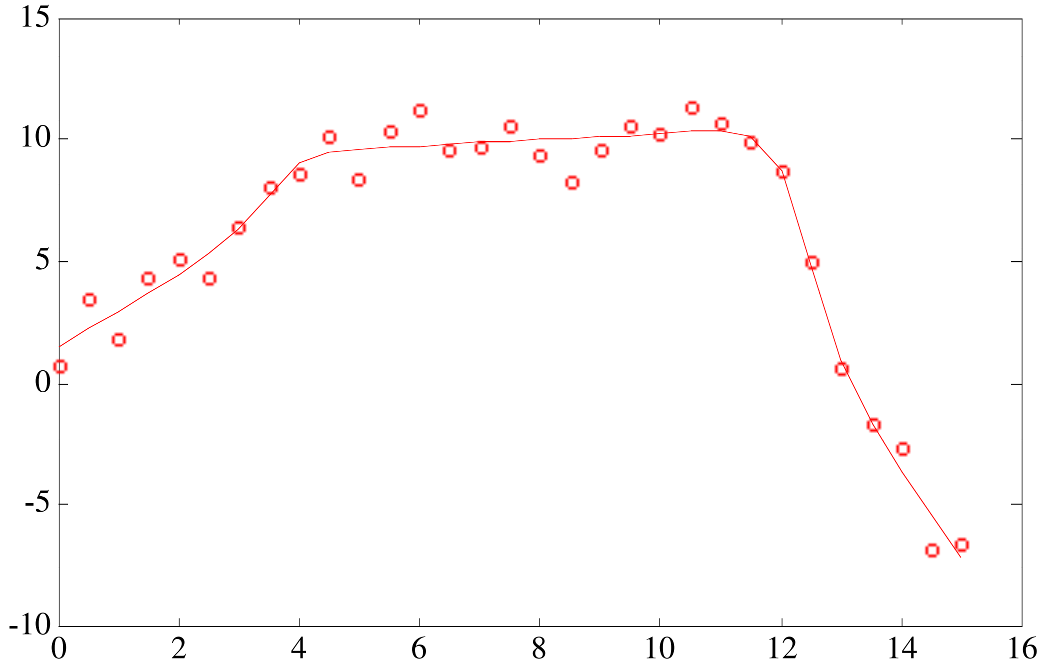

Figure 2a: The experimental and model values.

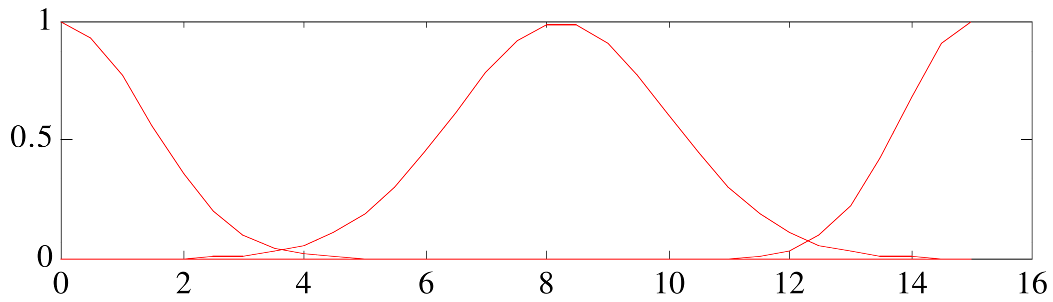

Figure 2b: The identified membership functions.

Once identified, the membership functions (figure 2b) can be written:

x is small : $\mu(x) = \exp\left[-\left(\frac{x}{1.974}\right)^2\right]$

x is medium : $\mu(x) = \exp\left[-\left(\frac{x - 8.244}{2.516}\right)^2\right]$

x is big : $\mu(x) = \exp\left[-\left(\frac{x - 15}{1.635}\right)^2\right]$

Finally, the model is defined by the three following fuzzy rules:

R_1	if	x is small,	then	$y = 1.449 + 1.489\ x$
R_2	if	x is medium,	then	$y = 8.891 + 0.126\ x$
R_3	if	x is big,	then	$y = 44.44 - 3.144\ x$

Let us now consider the case of a three variables system. Figure 3 indicates the distribution of the measurements (left side) and the estimations (right side) given by the model after identification of its parameters (the membership functions are gaussian and the model uses three rules). The top graph is a three dimensional view and the bottom graph shows level curves relatived to the premise variable y.

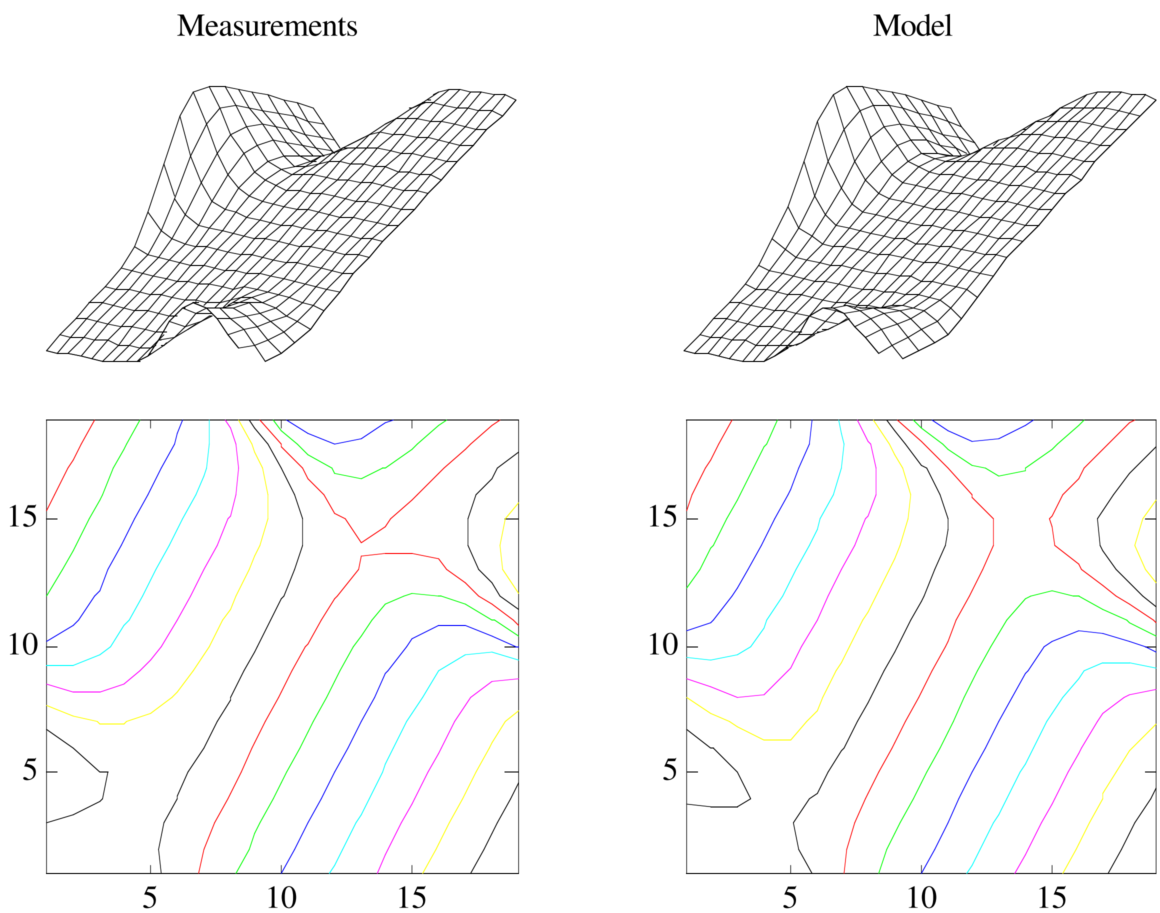

Figure 3: The three variables system.

5. CONCLUSION

The description of a system by a set of fuzzy relationships uses the same principles as classical algebraic systems and includes : a phase of modelisation consists of defining the structure of the model (number of rules, type of rules, number of variables, number of qualifiers, type of membership functions) and a phase of identification (estimation of the parameters of the rules and of the membership functions). The model is also validated by the same kind of tests (residual criteria and parameters consistency). However, in most cases, the identification of the fuzzy rules is a problem of non-linear optimization due to the structure of rules, the structure of membership functions, and also to the coupling between the rule parameters and the membership functions.

REFERENCES

Dubois, G. and E. Levrat. Résolution d'un système d'équations de relation floues, composition T-conorme, T norme générale.

Gorzalczany, M.B., J.B. Kiszka and M.S. Stachowicz (1982). Some problem of studying adequacy of fuzzy models. In : *Fuzzy Set and Possibility Theory*, R.R. Yager Ed., Pergamon, New York.

Hihi, J. and A. Richard (1993). Identification of fuzzy rules based models. *Proceedings of 2nd International Symposium IMACS/IFAC on Mathematical and Intelligent Models in System Simulation*, April 12-16, Brussels, Belgium, 232-236.

Sugeno, M. and G.T. Kang (1988). Structure identification of fuzzy model. *Fuzzy Sets and Systems*, **28**, 15-33.

Sugeno, M. and K. Tanaka (1991). Successive identification of a fuzzy model and its application to prediction of a complex system. *Fuzzy Sets and Systems*, **42**, 315-334.

Sugeno, M. and T. Takagi (1983). Multi-dimensional reasoning. *Fuzzy Sets and Systems*, **9** ,**2.**

Takagi, T. and M. Sugeno (1985). Fuzzy identification of systems and its application to modelling and control. *IEEE Transactions on Systems, Man and Cybernetics*, **15**, 116-132.

Tong, R.M. (1979). The construction and evaluation of fuzzy models. In : *Advances to Fuzzy Set Theory and Applications*, M.M. Gupta Ed., Plenum, New York.