

# A simple method for identification of singleton fuzzy models

C.-L. CHEN<sup>\*†</sup>, S.-H. HSU<sup>†</sup>, C.-T. HSIEH<sup>†</sup> and T.-C. WANG<sup>‡</sup>

<sup>†</sup>Department of Chemical Engineering, National Taiwan University, Taipei 106, Taiwan, Republic of China

<sup>‡</sup>Center for Environmental, Safety and Health, Industrial Technology Research Institute,  
Chutung, Hsinchu 310, Taiwan, Republic of China

(Received 15 November 2000; revised 15 July 2004; accepted 19 August 2005)

This article presents a simple method for constructing a singleton fuzzy model from a given set of input/output data. The method consists of three computational steps: the initial phase, the growth phase, and the optional refining phase. The universe of discourse and two linguistic terms for each input variable and a rule base are established during the initial phase. Additional linguistic terms and rules are then appended sequentially during the growth phase to modify the model structure and to elevate the performance. During the optional refining phase the overall modelling performance can be further improved by adjusting the singleton outputs of the rule set in the sense of least squares. The proposed identification method can simultaneously provide an appropriate model structure and parameters without any time-consuming optimisation. Several numerical examples demonstrate the effectiveness of the proposed identification method.

**Keywords:** Fuzzy sets; Fuzzy model; Identification

## 1. Introduction

System identification focuses mainly on establishing mathematical models to represent system input–output relationships, and selecting a specific model for the class of models. Although numerical data are abundant in most real world problems, the linguistic models are the ones that are widely used for field practitioners. This is because decisions are frequently based on qualitative judgement. Zadeh's fuzzy sets theory was thus proposed to enable individuals to describe and formulate the linguistic mental models apparent in daily life behaviour (Zadeh 1965).

Developing a model from numerical data by fuzzy technology has several advantages over the other competing modelling technologies such as neural networks: (1) The procedures are easily implemented; (2) applying cumbersome optimisation calculation is unnecessary; (3) the structure of the model can be

varied during construction; and (4) the resulting model has more physical meanings than others.

Two types of fuzzy models are widely used, the Mamdani fuzzy model and the Takagi–Sugeno–Kang (TSK) fuzzy model (Takagi and Sugeno 1985, Sugeno and Kang 1986). Practitioners find the latter appealing since the consequent parts of the IF–THEN rules in a TSK fuzzy model are usually simple linear functions of input variables. Thus, the model output value is a weighted average of those linear output equations (Yager and Filev 1994). The singleton fuzzy model is the simplest possible form of the TSK model. Though a TSK fuzzy model with its many parts achieves the same accuracy as models with singleton consequent outputs, the zero-order singleton fuzzy model is still worth using due to its ease of interpretation. The singleton fuzzy model can be viewed as a neural fuzzy system, and the well-known back-propagation learning algorithm can then be applied for the learning of model parameters. In such a case, *a priori* knowledge can be effectively incorporated in the initial parametric values (Yager and Filev 1994). A learning by back-propagation

\*Corresponding author. Email: CCL@ntu.eud.tw

algorithm for neural fuzzy models, which is an application of gradient descent method with Gaussian-type membership functions for antecedent fuzzy sets, is discussed by Wang and Mendel (1992). However, the convergence of such learning algorithm is questionable. Some investigations (Higgins and Goodman 1994, Nakoula *et al.* 1996), while having the convenience of the TSK model, present a constructing algorithm without using optimisations. The modelling effort can thus be reduced dramatically. However, the discussions are restricted in single-input/single-output cases.

In this article, we present a simple yet effective method for generating a multiple-input/single-output singleton fuzzy model from observed input/output data. The singleton fuzzy model studied herein consists of triangular fuzzy partitions for inputs and singletons for the output. An explicit input/output mathematical relation of the singleton fuzzy model is also derived. For a set of observed input/output data pairs, the structure and the parameters of an appropriate singleton fuzzy model with required accuracy are determined by a series of algebraic computations. Notably, the proposed method does not require any time-consuming optimisation procedure. Several numerical examples demonstrate the effectiveness of the proposed model and the identification method.

The remainder of this paper is organised as follows. In section 2, the explicit relation between input and output is proposed. The procedure of constructing the singleton-output type model is explained in section 3. Several numerical examples are given step by step in section 4 to demonstrate the effectiveness of the proposed model and the identification method. The conclusion is in section 5.

## 2. The singleton fuzzy model

The following equation contains a typical singleton fuzzy model with  $p$  inputs and single output:

$$R(i_1, i_2, \dots, i_p) : \text{if } x_1 \text{ is } X_1^{(i_1)} \text{ and } x_2 \text{ is } X_2^{(i_2)} \text{ and } \dots \text{ and } x_p \text{ is } X_p^{(i_p)} \text{ then } y = y^{(i_1, i_2, \dots, i_p)} \quad (1)$$

where  $x_j$  and  $y$  are the  $j$ th input variable and the output variable, respectively;  $y^{(i_1, i_2, \dots, i_p)}$  is the singleton output value of the rule  $R(i_1, i_2, \dots, i_p)$ ;  $X_j^{(i_j)}$  is the  $i_j$ th linguistic term of the input variable  $x_j$ ;  $V_j \equiv \{X_j^{(1)}, X_j^{(2)}, \dots, X_j^{(L_j)}\}$  is the term set of  $x_j$ ; and  $L_j$  is the number of linguistic terms of the input variable  $x_j$ . Each linguistic term is described by a triangular membership function  $X_j^{(i)}(x)$ ,

such as figure 1 and equation (2).

$$X_j^{(i)}(x) = \begin{cases} \frac{x - x_j^{(i-1)}}{x_j^{(i)} - x_j^{(i-1)}} & \text{if } x_j^{(i-1)} \leq x \leq x_j^{(i)}, \\ \frac{x_j^{(i+1)} - x}{x_j^{(i+1)} - x_j^{(i)}} & \text{if } x_j^{(i)} \leq x \leq x_j^{(i+1)}, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

According to the membership function definition in equation (2), two properties are simultaneously satisfied (Ying 1994, Chen *et al.* 1998).

**Property 1:** For any scalar input  $x \in U_x$ , no more than two membership values are nonzero.

**Property 2:** The sum of nonzero membership values is unity.

For an input pattern  $(x_1, x_2, \dots, x_p)$ , the degree of fulfilment of the rule indicated by the  $(i_1, i_2, \dots, i_p)$ ,  $\phi(i_1, i_2, \dots, i_p)$ , is determined by the fuzzy intersection ( $t$ -norm) of the relevant input membership values.

$$\phi(i_1, i_2, \dots, i_p) = \prod (X_1^{(i_1)}(x_1), X_2^{(i_2)}(x_2), \dots, X_p^{(i_p)}(x_p)) \quad (3)$$

By generalising to a  $p$ -input-single-output system, the corresponding model response is a weighted average of singleton outputs of the  $L_1 \times L_2 \times \dots \times L_p$  rules:

$$y = \frac{\sum_{i_1=1}^{L_1} \sum_{i_2=1}^{L_2} \dots \sum_{i_p=1}^{L_p} \phi(i_1, i_2, \dots, i_p) y^{(i_1, i_2, \dots, i_p)}}{\sum_{i_1=1}^{L_1} \sum_{i_2=1}^{L_2} \dots \sum_{i_p=1}^{L_p} \phi(i_1, i_2, \dots, i_p)} \quad (4)$$

$$= \sum_{i_1=1}^{L_1} \sum_{i_2=1}^{L_2} \dots \sum_{i_p=1}^{L_p} \omega(i_1, i_2, \dots, i_p) y^{(i_1, i_2, \dots, i_p)}$$

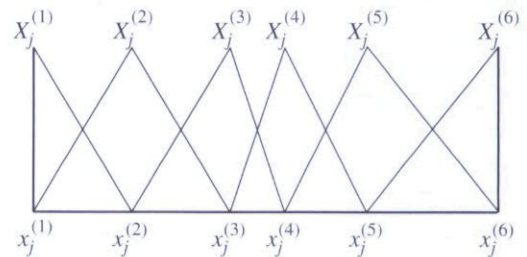


Figure 1. Illustration of linguistic terms  $X_j^{(i)}$ s and triangular membership functions  $X_j^{(i)}(x_j)$ s for variable  $x_j$ , where  $x_j^{(i)}$ s are principle values ( $L_j=6$ ).

is that,

$$\omega(i_1, i_2, \dots, i_p) = \frac{\phi(i_1, i_2, \dots, i_p)}{\sum_{i_1=1}^{L_1} \sum_{i_2=1}^{L_2} \dots \sum_{i_p=1}^{L_p} \phi(i_1, i_2, \dots, i_p)} \quad (5)$$

In this article, the algebraic product  $t$ -norm is chosen:

$$\phi(i_1, i_2, \dots, i_p) = \prod_{j=1}^p X_j^{(i_j)}(x_j) \quad (6)$$

The advantage of selecting the triangular membership function as figure 1 and the algebraic product  $t$ -norm is that the sum of all  $\phi(i_1, i_2, \dots, i_p)$  is unity for any input patterns, i.e.

$$\sum_{i_1=1}^{L_1} \sum_{i_2=1}^{L_2} \dots \sum_{i_p=1}^{L_p} \phi(i_1, i_2, \dots, i_p) = 1 \quad (7)$$

Thus,

$$\omega(i_1, i_2, \dots, i_p) = \phi(i_1, i_2, \dots, i_p) \quad (8)$$

Since Property 1 denotes at most two membership values which are not zero for each element of the input data, no more than  $2^p$  elements are non-zero in the set  $\Phi$ .

$$\Phi = \{\phi(i_1, i_2, \dots, i_p) | i_j = 1, 2, \dots, L_j; j = 1, 2, \dots, p\}$$

Therefore, only  $2^p$  rules at most contribute to the final output for any input pattern.

Without a loss of generality, for a two-input-single-output system ( $p=2$ ) and the algebraic product  $t$ -norm, the output is the weighted average of the four rules (Chen *et al.* 1998). Assume that  $x_1 \in [x_1^{(i_1)}, x_1^{(i_1+1)}], x_2 \in [x_2^{(i_2)}, x_2^{(i_2+1)}]$

$$\begin{aligned} R(i_1, i_2) : & \text{ if } x_1 \text{ is } X_1^{(i_1)} \text{ and } x_2 \text{ is } X_2^{(i_2)} \\ & \text{ then } y = y^{(i_1, i_2)} \\ R(i_1, i_2 + 1) : & \text{ if } x_1 \text{ is } X_1^{(i_1)} \text{ and } x_2 \text{ is } X_2^{(i_2+1)} \\ & \text{ then } y = y^{(i_1, i_2+1)} \\ R(i_1 + 1, i_2) : & \text{ if } x_1 \text{ is } X_1^{(i_1+1)} \text{ and } x_2 \text{ is } X_2^{(i_2)} \\ & \text{ then } y = y^{(i_1+1, i_2)} \\ R(i_1 + 1, i_2 + 1) : & \text{ if } x_1 \text{ is } X_1^{(i_1+1)} \text{ and } x_2 \text{ is } X_2^{(i_2+1)} \\ & \text{ then } y = y^{(i_1+1, i_2+1)} \end{aligned} \quad (9)$$

Define

$$\begin{aligned} \mathcal{X}_1 &\equiv X_1^{(i_1)}(x_1) = \frac{x_1^{(i_1+1)} - x_1}{x_1^{(i_1+1)} - x_1^{(i_1)}} \quad \text{and} \\ \mathcal{X}_2 &\equiv X_2^{(i_2)}(x_2) = \frac{x_2^{(i_2+1)} - x_2}{x_2^{(i_2+1)} - x_2^{(i_2)}} \end{aligned} \quad (10)$$

The explicit equation for the model response can be found,

$$\begin{aligned} y &= \mathcal{X}_1 \mathcal{X}_2 y^{(i_1, i_2)} + \mathcal{X}_1 (1 - \mathcal{X}_2) y^{(i_1, i_2+1)} \\ &\quad + (1 - \mathcal{X}_1) \mathcal{X}_2 y^{(i_1+1, i_2)} + (1 - \mathcal{X}_1)(1 - \mathcal{X}_2) y^{(i_1+1, i_2+1)} \\ &= \mathcal{X}_1 \mathcal{X}_2 y^{(i_1, i_2)} + \mathcal{X}_1 (1 - \mathcal{X}_2) y^{(i_1, i_2+1)} + (1 - \mathcal{X}_1) \mathcal{X}_2 y^{(i_1+1, i_2)} \\ &\quad + (1 - \mathcal{X}_1)(1 - \mathcal{X}_2) y^{(i_1+1, i_2+1)} \end{aligned} \quad (11)$$

This kind expression of a fuzzy rule is commonly referred to as indexed representation. Another useful representation of fuzzy rules is sequential representation:

$$\begin{aligned} R(i) : & \text{ if } x_1 \text{ is } X_1^{(i)} \text{ and } x_2 \text{ is } X_2^{(i)} \\ & \text{ and } \dots \text{ and } x_p \text{ is } X_p^{(i)} \text{ then } y = y^{(i)}, \quad i = 1, 2, \dots, n \end{aligned} \quad (12)$$

Where  $x_j^{(i)} \in V_j$  is the linguistic term in which the  $j$ th input variable  $x_j$  belongs to and  $y^{(i)}$  is the singleton output value of the  $i$ th rule. For an  $n$ -rule fuzzy model, the output is represented as follows:

$$y = \sum_{k=1}^n \omega_k y^{(k)}, \quad \omega_k = \frac{\phi_k}{\sum_{k=1}^n \phi_k} \quad (13)$$

where  $\phi_k$  is the degree of fulfilment of the  $k$ th rule.

### 3. Rule generating procedure

The unknown parameters in a fuzzy model include the number of rules and the number of linguistic terms for each variable (the structural parameters), the locations of principal values  $x_j^{(i)}$ s, the singleton output of each rule  $y^{(i)}$ , i.e. the input/output membership functions (the tuning parameters) (Sugeno and Yasukawa 1993). Previous literature separates the identification of the two types of parameters (Sugeno and Kang 1986). In addition the structural identification method frequently depends on the engineering judgment or optimization (Abonyi *et al.* 1999).

Owing to the above circumstances, a simple yet effective identification method is presented to determine both

model structure and input/output membership functions in an integrated procedure. The proposed method consists of three computational steps: the initial phase, the growth phase, and the optional refining phase. Figure 2 illustrates the procedure. Details of the three identification steps are provided in the following.

### 3.1. Initial phase

The simplest fuzzy model is that two linguistic terms (e.g., {Large, Small} or {High, Low}) are assigned to each input variable. The universes of discourse for the input variables are constructed by these linguistic terms and all ( $m$ ) data points are enclosed by the universe of discourse.

This step focuses on building the initial model with two linguistic terms for each input variable.  $2^p$  rules for  $p$ -input-single-output system will be generated.

1. Find the universe of discourse  $U_j$  for the  $j$ th input variable,  $x_j$ .

$$U_j = \{x_j | x_j \in [x_{j, \min}, x_{j, \max}]\} \quad j = 1, 2, \dots, p$$

where  $x_{j, \min} = \min\{x_{j, 1}, x_{j, 2}, \dots, x_{j, m}\}$

$$x_{j, \max} = \max\{x_{j, 1}, x_{j, 2}, \dots, x_{j, m}\}$$

2. Construct two linguistic terms,  $X_j^{(1)}$  and  $X_j^{(2)}$ , with principal values at  $x_{j, \min}$  and  $x_{j, \max}$ , respectively. Figure 3 depicts the membership functions for  $X_j^{(1)}$  and  $X_j^{(2)}$ . Then all data points are located in the universe of discourse  $U$ , where  $U = U_1 \times U_2 \times \dots \times U_p$ . The  $p$  input variables, each with two linguistic terms, can establish  $2^p$  rules with the following antecedent part, where the consequence part for each rule can be determined as shown in the next step.

<p>If <math>x_1</math> is <math>X_1^{(1)}</math> and <math>x_2</math> is <math>X_2^{(1)}</math> and ... and</p> <p>If <math>x_1</math> is <math>X_1^{(1)}</math> and <math>x_2</math> is <math>X_2^{(1)}</math> and ... and</p> <p style="text-align: center;"><math>\vdots</math></p> <p>If <math>x_1</math> is <math>X_1^{(2)}</math> and <math>x_2</math> is <math>X_2^{(2)}</math> and ... and</p>	<p><math>x_p</math> is <math>X_p^{(1)}</math> then <math>y = y^{(1, 1, \dots, 1)} = y^{(1)}</math></p> <p><math>x_p</math> is <math>X_p^{(2)}</math> then <math>y = y^{(1, 1, \dots, 2)} = y^{(2)}</math></p> <p style="text-align: center;"><math>\vdots</math></p> <p><math>x_p</math> is <math>X_p^{(2)}</math> then <math>y = y^{(2, 2, \dots, 2)} = y^{(2^p)}</math></p>
--	---

3. Determine the consequence part for each of the  $2^p$  rules by the nearest data point or the refining phase (see section 3.3). Although simpler, the nearest data point method may not be accurate enough. For instance,  $y^{(1, 1, \dots, 1)}$  (or  $y^{(1)}$ ) is the rule output to be obtained. If one input pattern  $(x_{1, l}, x_{2, l}, \dots, x_{p, l})$  in the training data set is the nearest point to  $(x_{1, \min}, x_{2, \min}, \dots, x_{p, \min})$ , i.e.  $\sum_{j=1}^p (x_{j, l} - x_{j, \min})^2$  is minimum, the output measurement  $y_l$  corresponding to

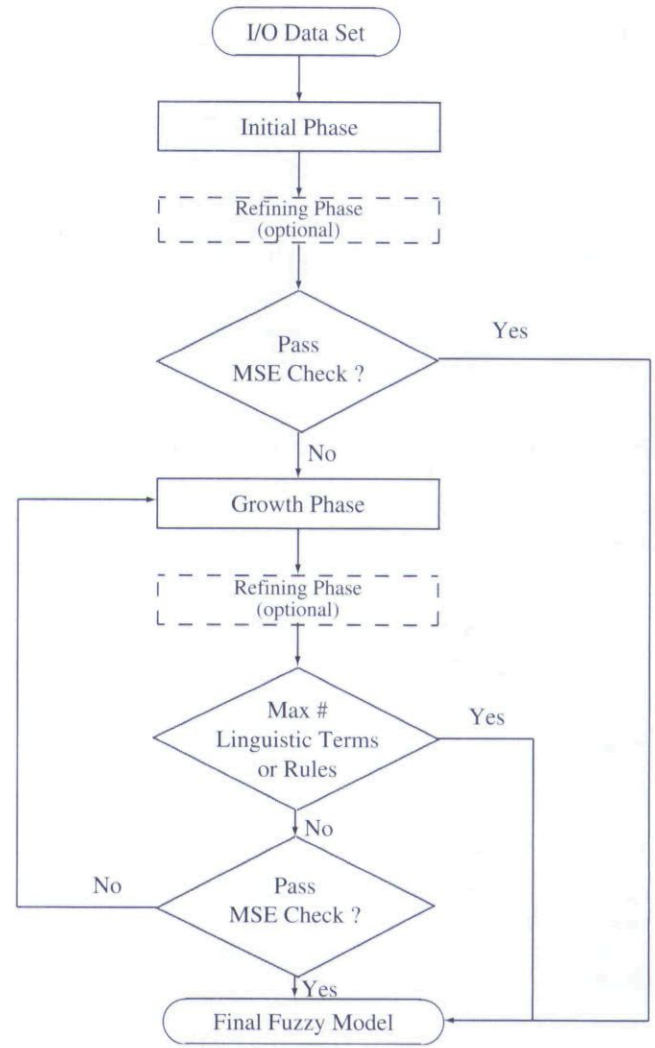


Figure 2. Rule generating procedure.

$(x_{1, l}, x_{2, l}, \dots, x_{p, l})$  is set to be the singleton output of the rule. That is,

$$\begin{aligned} \text{If } x_1 \text{ is } X_1^{(1)} \text{ and } x_2 \text{ is } X_2^{(1)} \text{ and } \dots \text{ and } x_p \text{ is } X_p^{(1)} \\ \text{then } y = y^{(1, 1, \dots, 1)} = y^{(1)} = y_l \end{aligned} \quad (14)$$

For the same reason, all the outputs of rules  $y^{(i)}$ ,  $i = 2, 3, \dots, 2^p$  are also determined.

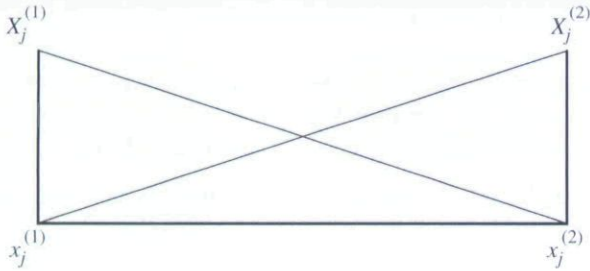


Figure 3. Membership functions for each input variable in the initial phase.

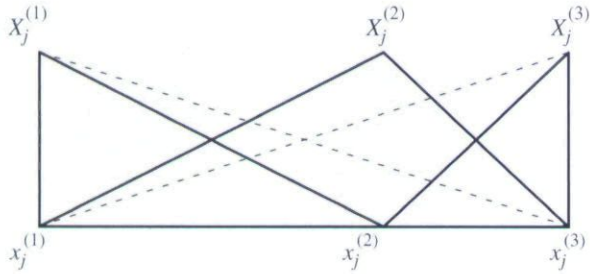


Figure 4. Membership functions for each input variable in the growth phase. ---: Membership functions of two linguistic terms, —: membership functions of three linguistic terms.

### 3.2. Growth phase

Additional linguistic terms and inference rules should be appended on the existing fuzzy model if the latter cannot afford the required accuracy for modelling sampled data. In this subsection, we propose a method to construct more rules to enhance the performance of the model. The underlying concept of this method is, instead of inserting new terms to all inputs, adding a linguistic term to the most effective input variable on the point with the maximum modelling error. Some inference rules are thus constructed due to this new fuzzy term and then inserted into the existing fuzzy model to eliminate the error of that point. The proposed rule-inserting algorithm can effectively avoid suffering from the curse of dimensionality. The algorithm can also be modified to insert new terms to the most important inputs only. The procedure is as follows.

1. Evaluate the model input for all sampling data points. The point with maximum error between the measurement and model is found, denoted by  $(x_{1,y}, x_{2,y}, \dots, x_{p,y}; y_y)$ .
2. Add a linguistic term for  $x_1$  with a principal value  $x_{1,y}$ . The membership function of each linguistic term for  $x_1$  is changed, such as figure 4. However, the number of linguistic terms for the other input variables still remain unchanged. Restated, the number of rules increases from  $L_1 \times L_2 \times \dots \times L_n$  to  $(L_1 + 1) \times L_2 \times \dots \times L_n$ . Determine the consequence part by the previous method in section 3.1.

3. Evaluate the sum of squared errors  $SSE_1$ .
4. Repeat steps 2 and 3 for  $x_2, x_3, \dots, x_p$ , respectively. By doing so, the sum of squared errors  $SSE_2, SSE_3, \dots, SSE_p$  are obtained.
5. Find the minimum of  $\{SSE_1, SSE_2, \dots, SSE_p\}$ ,  $SSE_j$ , and the new linguistic term will be added in the  $x_j$  direction. Therefore,  $L_j$  adds one and other  $L_{k \neq j}$ s are kept.

### 3.3. Refining phase

In addition to the nearest point method, the least squares method can more accurately adjust the singleton output values, while keeping the same input membership functions from the growth phase, to help the model reach optimal performance. Use the representation equation (12) and equation (13) for the  $k$ th sampling point,

$$\hat{y}_k = \sum_{i=1}^n \omega_i y^{(i)} = [\omega_{k1} \quad \omega_{k2} \quad \dots \quad \omega_{kn}] \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} \quad (15)$$

Collect all  $m$  sampling data,

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix} = \begin{bmatrix} \omega_{11} & \omega_{12} & \dots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \dots & \omega_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{m1} & \omega_{m2} & \dots & \omega_{mn} \end{bmatrix} \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n)} \end{bmatrix} \quad (16)$$

or,

$$\hat{\mathbf{Y}} = \Phi \mathbf{A} \quad (17)$$

Thus, the least squares solution of  $\mathbf{A}$  which allows the performance index  $J = \sum_{k=1}^m (y_k - \hat{y}_k)^2$  to reach its minimum is the optimal solution of the principal values. Let the observed  $\mathbf{Y} \equiv [y_1 \quad y_2 \quad \dots \quad y_m]^T$ , then (Hsia 1979)

$$\mathbf{A} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{Y} \quad (18)$$

### 3.4. Stopping criteria

It is a trade-off problem between model complexity and accuracy. Two criteria for stopping the model building procedure are considered in the following. Both criteria can be used to stop the model identification procedure

to avert as well as avoid the complex model structure and obtain an accurate model.

1. *Maximum number of input linguistic terms.* A realistic model should not consist of too many linguistic terms for a specific input. The maximum number of linguistic terms of each variable can be predefined. This criterion can prevent the fuzzy model becoming a large number of rules. Restated, the criterion makes the model hold an acceptable complexity. However, the accuracy of the obtained model is not ensured.
2. *Mean of squared errors.* This index is a common used criterion to specify the model accuracy.

$$\text{MSE} = \frac{1}{m} \sum_{k=1}^m (\hat{y}_k - y_k)^2 \quad (19)$$

Although the MSE index can make the model accurate enough, the model may become extremely complex.

#### 4. Numerical examples

Several numerical examples are given step by step in the following to demonstrate the effectiveness of the proposed model and the identification method.

**Example 1:** Consider a nonlinear function with two inputs and single output, refer to figure 5.

$$y = e^{-x_1} + e^{-x_2}, \quad x_1, x_2 \in [0, 5] \quad (20)$$

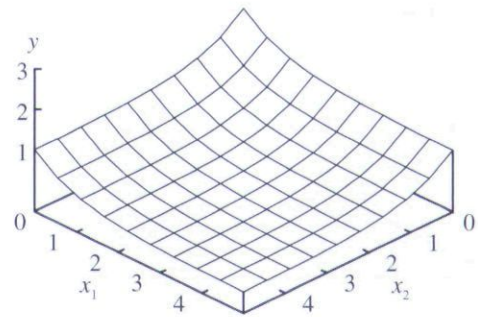


Figure 5. Output surface of Example 1.

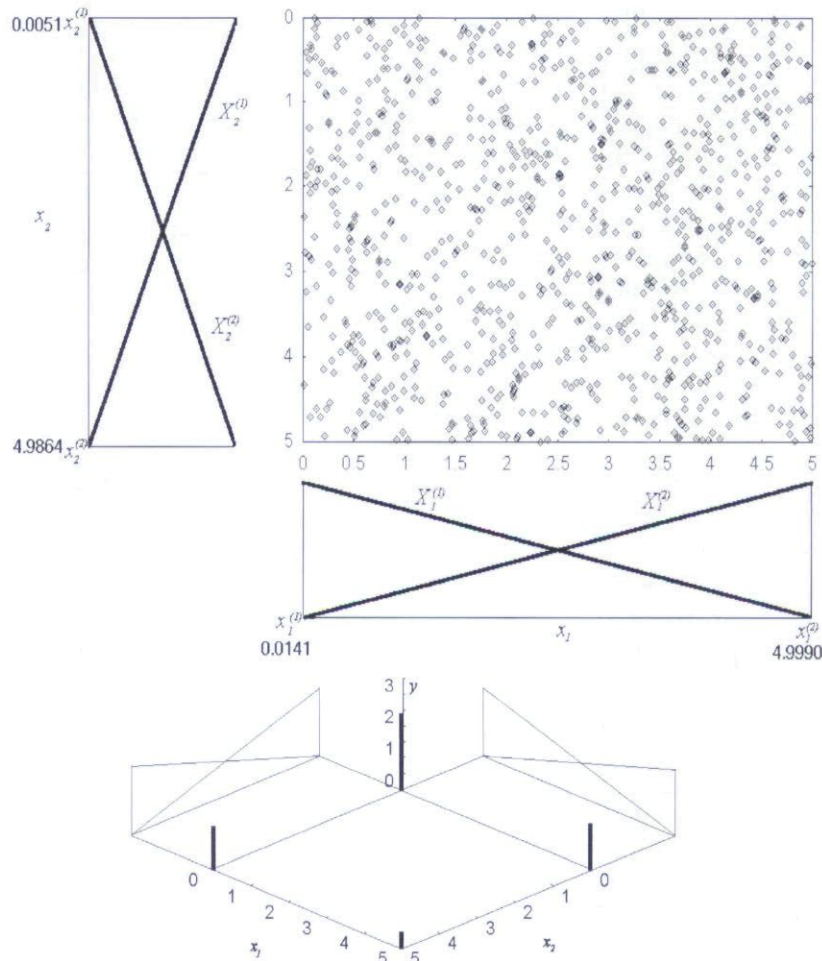


Figure 6. Input membership functions and singleton fuzzy rule outputs for the initial phase of Example 1.

Table 1. Principal values for  $x_1$ ,  $x_2$  and singleton outputs  $y$  of Example 1: (a) initial phase; (b) and (c) growth phase.

(a)		$x_1$		
$y$		0.0141	4.9990	
$x_2$	0.0051	1.884	0.926	
	4.9864	0.812	0.014	
MSE		0.294		

(b)		$x_1$		
$y$		0.0141	1.6067	4.9990
$x_2$	0.0051	1.884	1.196	0.926
	1.8198	1.142	0.363	0.199
	4.9864	0.812	0.175	0.014
MSE		0.0124		

(c)		$x_1$			
$y$		0.0141	0.8136	1.6067	4.9990
$x_2$	0.0051	1.884	1.474	1.064	0.926
	0.8727	1.304	0.861	0.563	0.399
	1.8198	1.142	0.644	0.363	0.199
	4.9864	0.812	0.406	0.175	0.014
MSE		0.00458			

1,000 inputs/outputs of data are randomly distributed on the considered domain.

1. *Initial phase.* The upper and lower bounds for the two inputs,  $x_{1,\min}$ ,  $x_{1,\max}$ ,  $x_{2,\min}$ , and  $x_{2,\max}$  are initially found. Two linguistic terms are defined for  $x_1$  and  $x_2$ . The singleton outputs for the rules are obtained, such as figure 6 and table 1(a). Notably, the output surface of the resulting fuzzy model is the weighted average of the four rule's consequent outputs, i.e.

$$y = 1.884\mathcal{X}_1\mathcal{X}_2 + 0.812\mathcal{X}_1(1 - \mathcal{X}_2) + 0.926(1 - \mathcal{X}_1)\mathcal{X}_2 + 0.014(1 - \mathcal{X}_1)(1 - \mathcal{X}_2)$$

where

$$\mathcal{X}_1 \equiv X_1^{(1)}(x_1) = \frac{x_1^{(2)} - x_1}{x_1^{(2)} - x_1^{(1)}} = \frac{4.9990 - x_1}{4.9990 - 0.0141}$$

$$\mathcal{X}_2 \equiv X_2^{(1)}(x_2) = \frac{x_2^{(2)} - x_2}{x_2^{(2)} - x_2^{(1)}} = \frac{4.9864 - x_2}{4.9864 - 0.0051}$$

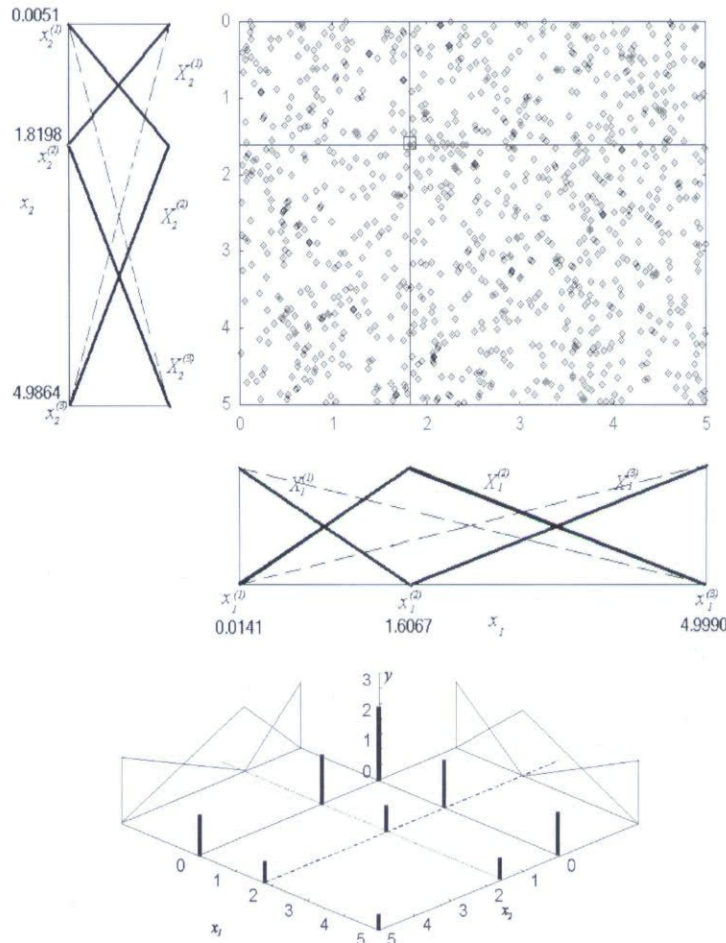


Figure 7. Input membership functions and singleton fuzzy rule outputs for the growth phase of Example 1(I).

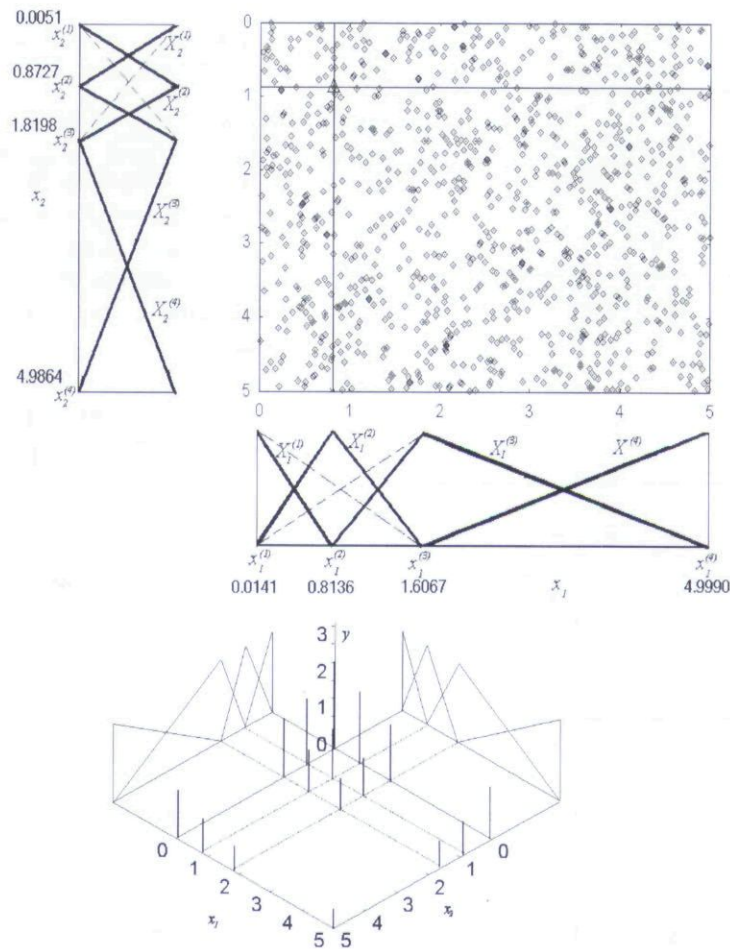


Figure 8. Input membership functions and singleton fuzzy rule outputs for the growth phase for Example 1(II).

2. *Growth phase.* The maximum discrepancy between the model outputs and the observations is found. Additional linguistic terms are appended for both variables to eliminate the maximum error. Figure 7 and figure 8 display some growth processes. Table 1 lists some resulting models.
3. *Refining phase.* If the singleton outputs are determined by the refining phase, the mean of squared errors will be smaller than the rules generated by the nearest point method, as illustrated in table 2.

**Example 2:** The Mackey-Glass chaotic time series (Nakoula *et al.* 1996, Wang and Mendel 1992) Mackey-Glass delay differential equation is used as a benchmark problem in the neural network and fuzzy modelling communities:

$$\begin{aligned} \dot{x}(t) &= \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t), \quad \tau \leq 17 \\ x(0) &= 1.2 \\ x(t) &= 0 \quad \forall t < 0 \end{aligned} \quad (21)$$

The proposed model is a 4-input-single-output system. The inputs are  $x_1 \equiv x(t-18)$ ,  $x_2 \equiv x(t-12)$ ,  $x_3 \equiv x(t-6)$ ,  $x_4 \equiv x(t)$ . The output is  $y \equiv x(t+6)$  (Wang and Mendel 1992). Figure 9 depicts one set of time series with 1800 samples. In this modelling problem, the sampling data in  $t \in [1000, 1500]$  are used for identification, and data in  $t \in [1500, 1800]$  are used for testing the accuracy of the resulting model. Figures 10 and 11 summarise the results.

## 5. Conclusions

This article presents a simple algorithm for identifying a singleton fuzzy model. The singleton fuzzy model is advantageous in that the input-output relationship is much easier than the linguistic fuzzy model, and can be derived explicitly. For a set of observed patterns, the simplest possible fuzzy model is constructed in three steps: the initial phase, the growth phase, and the optional refining phase.

Table 2. Refined singleton outputs  $y$  of Example 1: (a) initial rules; (b) and (c) growth phase.

(a)		$x_1$		
$y$		0.0141	4.9990	
$x_2$	0.0051	1.179	0.390	
	4.9864	0.355	−0.319	
MSE		0.0315		

(b)		$x_1$		
$y$		0.0141	1.6067	4.9990
$x_2$	0.0051	1.722	0.975	0.858
	1.8198	0.978	0.236	0.120
	4.9864	0.769	0.085	−0.034
MSE		0.00242		

(c)		$x_1$			
$y$		0.0141	0.8136	1.6067	4.9990
$x_2$	0.0051	1.911	1.357	1.070	0.917
	0.8727	1.323	0.835	0.518	0.381
	1.8198	1.109	0.560	0.280	0.125
	4.9864	0.898	0.395	0.096	−0.055
MSE		0.00074			

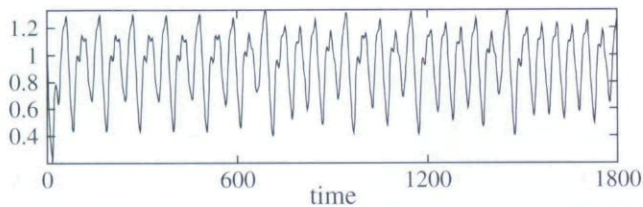


Figure 9. Mackey-Glass chaotic time series.

The universe of discourse of each input variable is determined in the initial phase. Using the minimum and maximum of the input variables to obtain the universe of discourse can ensure that all the training data points are in the universe. More linguistic terms and rules are appended in the growth phase to modify the model structure and enhance the performance. In addition, an alternative method based on the least squares method to evaluate the singleton outputs is proposed in the optional refining phase. This method provides a better model performance.

Compared to other fuzzy modelling methods, such as learning by back-propagation, the proposed method does not require time-consuming optimisation techniques to identify the model structure and to estimate the model parameters. Although the principal values

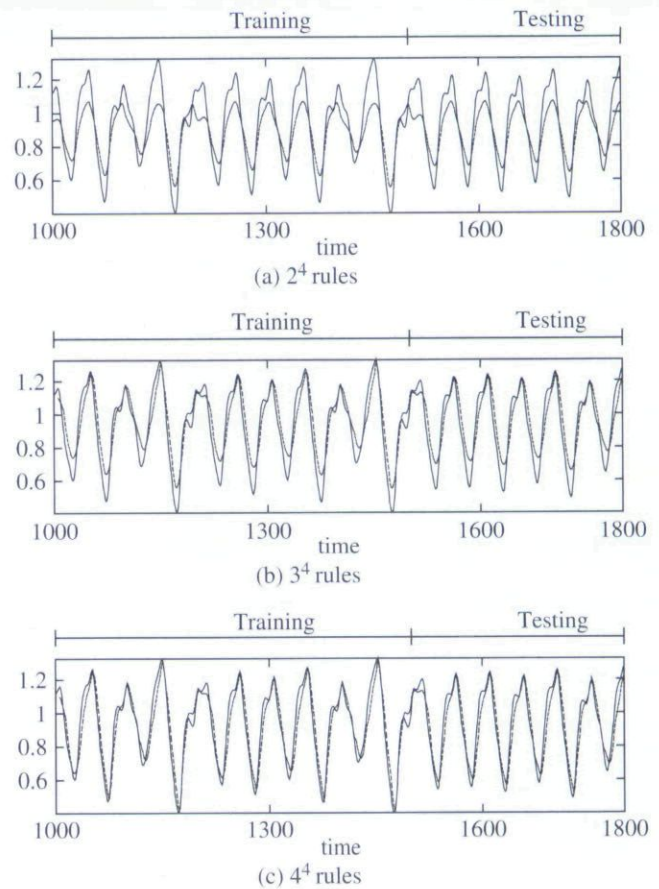


Figure 10. Modelling Mackey-Glass time series.

for input variables are not optimal, the least squares method gives optimal performance under this model structure, and computational time is reduced. Two criteria are suggested to stop the identification process. The maximum number of input linguistic terms makes the rule base simpler, and the minimum sum of squared errors guarantees an accurate model.

The proposed method is highly promising for engineering applications since one can proceed with the model identification without any searching procedure and thus no numerical converging problem. Two numerical examples demonstrate the feasibility of the proposed method in terms of building the singleton fuzzy model from data.

#### Acknowledgement

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC87-2214-E-002-014.

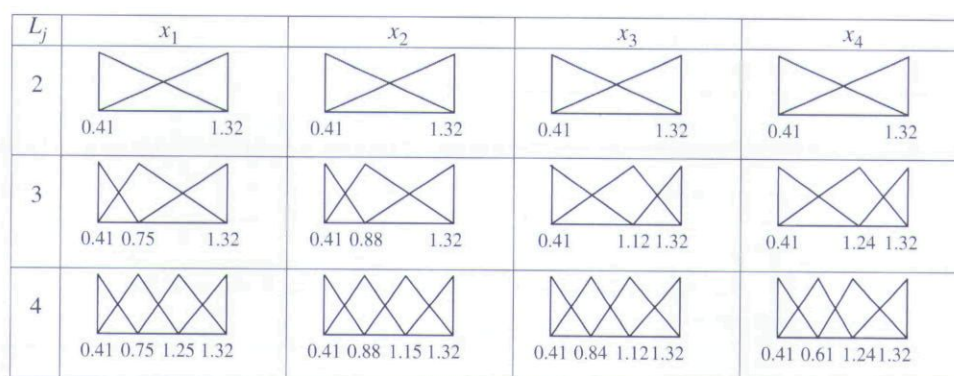


Figure 11. Input membership functions of modelling Mackey-Glass time series.

## References

- J. Abonyi, L. Nagy and F. Szeifert, "Adaptive fuzzy inference system and its application in modelling and model based control", *Trans. IChemE*, 77, pp. 281–290, 1999.
- C.L. Chen, S.N. Wang, C.T. Hsieh and F.Y. Chang, "Theoretical analysis of crisp-type fuzzy logic controllers using various  $t$ -norm-sum-gravity inference methods", *IEEE Trans. on Fuzzy Systems*, 6, pp. 122–136, 1998.
- C.M. Higgins and R.M. Goodman, "Fuzzy rule-based networks for control", *IEEE Trans. on Fuzzy Systems*, 2, pp. 82–88, 1994.
- T.C. Hsia, *System Identification: Least-squares Methods*, Lanham, MD, USA: Lexington Books, 1979, pp. 17–22.
- S. Nakoula, S. Galichet and L. Foulloy, "Simultaneous learning of rules and linguistic terms", in *5th IEEE International Conference on Fuzzy Systems*, New Orleans, 1996, pp. 1743–1749.
- M. Sugeno and G.T. Kang, "Fuzzy modelling and control of multi-layer incinerator", *Fuzzy Sets and Systems*, 18, pp. 329–346, 1986.
- M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modelling", *IEEE Trans. on Fuzzy Systems*, 1, pp. 7–31, 1993.
- T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modelling and control", *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-15, pp. 116–132, 1985.
- L.X. Wang and J.M. Mendel, "Generating fuzzy rules by learning from examples", *IEEE Trans. on Systems, Man, and Cybernetics*, 22, pp. 1414–1427, 1992.
- R.R. Yager and D.P. Filev, *Essentials of Fuzzy Modeling and Control*, New York: John Wiley & Sons, Inc., 1994.
- H. Ying, "A nonlinear fuzzy controller with linear control rules is the sum of a global two-dimensional multilevel relay and a local nonlinear proportional-integral controller", *Automatica*, 29, pp. 1185–1195, 1994.
- L.A. Zadeh, "Fuzzy sets", *Information and Control*, 8, pp. 338–353, 1965.



**Cheng-Liang Chen** received the Ph.D. degree in chemical engineering from National Taiwan University in 1987. He is currently a Professor with the Department of Chemical Engineering, National Taiwan University. His research interests include neural/fuzzy modeling and control, identification and control systems design, process integration and optimization, production scheduling and supply chain management.

**Shuo-Huan Hsu** received the Master degree in chemical engineering from National Taiwan University in 1998. Currently, he is a Ph.D. candidate at the Department of Chemical Engineering, Purdue University, USA. His research interests include system identification, genetic algorithms, system biology.

**Chung-Tyan Hsieh** received the Ph.D. degree in chemical engineering from National Taiwan University in 1997. His research interests include expert systems, fuzzy modeling and control. In addition to the research work, Dr. Hsieh has developed a Chinese TeX system which is one of the main typesetting systems used in the Republic of China.

**Tzu-Chi Wang** received the Ph.D. degree in chemical engineering from National Taiwan University in 2001. Currently, he is a research fellow at the Center for Environmental, Safety and Health, the Industrial Technology Research Institute, Taiwan. His research interests include control systems design, process monitoring and safety engineering.

Copyright of International Journal of Systems Science is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.