2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

T-S Fuzzy Affine Linear Modeling Algorithm by Possibilistic *c*-Regression Models Clustering Algorithm

Chung-Chun Kung and Hong-Chi Ku

***Abstract*—This paper presents a Takagi-Sugeno (T-S) fuzzy affine linear modeling algorithm by the possibilistic *c*-regression models (PCRM) clustering algorithm. We apply the PCRM to partition the given input-output data into hyper-plane-shaped clusters (regression models). We choose the suitable number of cluster by the cluster validity criterion and then to construct the T-S fuzzy affine linear model. A simulation example is provided to demonstrate the effectiveness of the T-S fuzzy affine linear modeling algorithm.**

Index Terms- Takagi-Sugeno (T-S) fuzzy model; affine linear; possibilistic c-regression models (PCRM); cluster validity criterion

I. Introduction

Fuzzy models have been an active research area for several years. Building a fuzzy model for a real system which is usually a nonlinear system will encounter some problems, such as the kind of structure that could be accepted to describe this real system, the number of IF-THEN rules, and the reliable parameter of the antecedent and consequent parts. Takagi-Sugeno (T-S) fuzzy model [1, 2] is a popular fuzzy model because its consequent part is functional type and it has good capability of describing a nonlinear system. It can accurately approximate the given nonlinear systems with fewer rules than other types of fuzzy models.

In particular, Bezdek et al. proposed the fuzzy clustering algorithm for hyper-plane-shaped, that is, fuzzy *c*-regression models (FCRM) clustering algorithm [3] which assumes that the data are draw from c different regression models. Additionally, Kim et al. [4] successfully applied the FCRM clustering algorithm to construct fuzzy model. However, the clustering result of the FCRM clustering algorithm is sensitive to the noisy data. The column sum constraint of the FCRM clustering algorithm could result in each cluster is sensitive to the noisy data. To overcome this problem, Krishnapuram and Keller [5] proposed the possibilistic *c*-means (PCM) clustering algorithm which relaxes the column sum constraint and each cluster reduces the effects of the noisy data effectively. Recently, we successfully applied the characteristic of the PCM clustering algorithm to the FCRM clustering algorithm named the possibilistic *c*-regression models (PCRM) clustering algorithm [6] and it could reduce the effects of the noisy data effectively.

On the other hand, the number of clusters, c , is fixed and must be prior assigned. In order to solve this problem, there were many literatures of cluster validity criterion proposed to help the cluster algorithm to choose an appropriate number of clusters such as Bezdek's partition coefficient [7] and partition entropy [8], Xie-Beni's index [9], and cluster validity criterion for FCRM clustering algorithm [10- 16]. We adopt the cluster validity criterion for FCRM clustering algorithm to our fuzzy modeling algorithm because the PCRM clustering algorithm is similar to the FCRM clustering algorithm.

T-S fuzzy affine linear model also have been used to describe nonlinear systems [3, 10- 16]. In this paper, our objective is to use the T-S fuzzy affine linear models to describe a nonlinear system with high accuracy and as few IF-THEN rules as possible. First, we collect input and output data from original nonlinear system, and partition these data into some different clusters via identification and PCRM clustering algorithm. Then, we apply a cluster validity criterion for the PCRM clustering algorithm to choose an appropriate number of clusters. Finally, we construct fuzzy affine linear model for original nonlinear system via modeling algorithm.

The rest of the paper is organized as follows. Section II introduces the T-S fuzzy model. Section III proposes the T-S fuzzy affine linear modeling algorithm includes data partition, determination of the number of the clusters and fuzzy rule construction. In Section IV, we illustrate the effectiveness of the T-S fuzzy affine linear modeling algorithm with several examples. Section V gives the summary and conclusions.

II. Takagi-Sugeno Fuzzy Model

A fuzzy rule-based model suitable for describing a large class of nonlinear systems was introduced by Takagi and Sugeno [1, 2] as follows:

$$\begin{aligned} R^i : &\text{IF } u_1 \text{ is } \mathbf{A}_1^i \text{ and } \cdots \text{ and } u_n \text{ is } \mathbf{A}_n^i \\ &\text{THEN } y^i = f^i(\mathbf{u}, \boldsymbol{\beta}^i) \end{aligned} \tag{1}$$

where $\mathbf{u} = [u_1, \cdots, u_n]^T \in \Re^n$, $\boldsymbol{\beta}^i \in \Re^{n+1}$ is the parameter of the ith regression model (cluster) for $i = 1, 2, \ldots, r$, R^i denotes the ith IF-THEN rule and c is the number of rules in the rule base. $u_q, q = 1, 2, \ldots, n$, are individual input variables, and $\mathbf{A}_q^i$ are associated individual antecedent fuzzy sets of each input variable. $y^i \in \Re$ is the output of each rule.

For any input vector $\mathbf{u}$, if the singleton fuzzifier, the product fuzzy inference and the centre average defuzzifier are

Chung-Chun Kung is with the Department of Electrical Engineering, Tatung University, No.40, Sec. 3, Zhongshan N. Rd., Taipei City 104, Taiwan. (e-mail: cckung@ttu.edu.tw)

Hong-Chi Ku is with the Department of Electrical Engineering, Tatung University, No.40, Sec. 3, Zhongshan N. Rd., Taipei City 104, Taiwan (e-mail: maple10_08@hotmail.com)

978-1-4799-2072-3/14/$31.00 ©2014 IEEE

applied, the output of the fuzzy model $\hat{y}$ is inferred as follows [17, 18]:

$$\hat{y} = \frac{\sum_{i=1}^{c} w^i(\mathbf{u}) y^i}{\sum_{i=1}^{c} w^i(\mathbf{u})} \tag{2}$$

where

$$\begin{aligned} w^i(\mathbf{u}) &= \mathbf{A}_1^i(u_1) \times \mathbf{A}_2^i(u_2) \times \cdots \times \mathbf{A}_n^i(u_n) \\ &= \prod_{q=1}^{n} \mathbf{A}_q^i(u_q), \end{aligned} \tag{3}$$

$w^i(\mathbf{u})$ denote the degree of fulfillment of the antecedent, that is, the level of firing of the ith rule.

III. Fuzzy Modeling Algorithm

The proposed T-S fuzzy affine linear modeling algorithm is composed of three parts: input-output data partition by the PCRM clustering algorithm, determination of the number of clusters and fuzzy rule construction.

A. Data partition

To overcome the problem of the FCRM clustering algorithm [6], we successfully applied the characteristics of the possibilistic c-means (PCM) [5] clustering algorithm to the FCRM clustering algorithm and proposed a new clustering algorithm named possibilistic c-regression models (PCRM).

Suppose that $\mathbf{Z} = \{\mathbf{z}_1, \cdots, \mathbf{z}_L\}$ is an unlabeled data set, where $\mathbf{z}_h = (\mathbf{u}_h, y_h) \in \Re^n \times \Re$ for all $h \in \{1, \cdots, L\}$, and let $\mathbf{P}$ denote a $c \times L$ possibilistic c-partition matrix generated by PCRM clustering algorithm as follows [5, 6, 19]:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1L} \\ p_{21} & p_{22} & \cdots & p_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ p_{c1} & p_{c2} & \cdots & p_{cL} \end{bmatrix} \tag{4}$$

where p_{ih} of $\mathbf{P}$ satisfies the following conditions [5, 6, 19]:

$$p_{ih} \in [0,1] \text{ for all } i \in \{1, \cdots, c\} \text{ and } h \in \{1, \cdots, L\}, \tag{5}$$

$$0 < \sum_{h=1}^{L} p_{ih} < L \text{ for all } h \in \{1, \cdots, L\}, \text{ and} \tag{6}$$

$$\max_i p_{ih} > 0 \text{ for all } h \in \{1, \cdots, L\}, \tag{7}$$

where c is the number of clusters and p_{ih} is the degree of possibility (typicality) of $\mathbf{z}_h$ belonging to the ith cluster.

The regression model adopted in this paper is as follows [3, 6, 10- 16]:

$$\begin{aligned} y_h &= \beta_0^i + \beta_1^i u_{h1} + \cdots + \beta_n^i u_{hn} + e_h \\ &= [\mathbf{u}_h^T \ 1]\boldsymbol{\beta}^i + e_h, \end{aligned} \tag{8}$$

where $\mathbf{u}_h = [u_{h1}, \cdots, u_{hn}]^T \in \Re^n$, $\boldsymbol{\beta}^i = [\beta_1^i, \cdots, \beta_n^i, \beta_0^i] \in \Re^{n+1}$ is the vector of parameters of the ith regression model (cluster) for all $i \in \{1, \cdots, c\}$ and $\beta_0^i \in \Re$ denotes the offset or shift term. Then the local estimation error is defined by [3, 6, 10- 16]:

$$D_{ih}(\boldsymbol{\beta}^i) = |e_h|. \tag{9}$$

In such case, the regression model parameters $\boldsymbol{\beta}^i$ can be estimated by using the weighted least square (WLS) algorithm [6, 15, 16, 20, 21] as follows:

$$\boldsymbol{\beta}^i = [\boldsymbol{\Gamma}^T \mathbf{W}^i \boldsymbol{\Gamma}]^{-1} \boldsymbol{\Gamma}^T \mathbf{W}^i \boldsymbol{\Pi} \tag{10}$$

where

$$\boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{u}_1^T, 1 \\ \mathbf{u}_2^T, 1 \\ \vdots \\ \mathbf{u}_L^T, 1 \end{bmatrix} \in \Re^{L \times (n+1)}, \boldsymbol{\Pi} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_L \end{bmatrix} \in \Re^L \tag{11}$$

and

$$\mathbf{W}^i = diag[(p_{i1})^m, \cdots, (p_{iL})^m] \in \Re^{L \times L}. \tag{12}$$

The objective function for the PCRM clustering algorithm is defined by [5, 6, 19]

$$J(\mathbf{P}, \boldsymbol{\Theta}) = \sum_{h=1}^{L} \sum_{i=1}^{c} (p_{ih})^m (D_{ih}(\boldsymbol{\beta}^i))^2 + \sum_{i=1}^{c} \rho^i \sum_{h=1}^{L} (1 - p_{ih})^m,$$

$$\rho^i = \gamma \frac{\sum_{h=1}^{L} \sum_{i=1}^{c} (p_{ih})^m (D_{ih}(\boldsymbol{\beta}^i))^2}{\sum_{h=1}^{L} (p_{ih})^m} \text{ for all } i \in \{1, \cdots, c\}, \tag{13}$$

where $\boldsymbol{\Theta} = (\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, ..., \boldsymbol{\beta}^c)$, ρ is a constant and $m \in [2, \infty)$ is the weighting exponent. For the PCRM clustering algorithm, the objective is to find $(\mathbf{P}, \boldsymbol{\Theta})$ such that (13) is minimized.

PCRM Clustering Algorithm:

Step 1: Given c ($1 < c < L$), $m \in [2, \infty)$, the termination threshold $\varepsilon > 0$, set an initial possibilistic c-partition matrix

$$\mathbf{P}^{(0)} = \begin{bmatrix} p_{11}^{(0)} & p_{12}^{(0)} & \cdots & p_{1L}^{(0)} \\ p_{21}^{(0)} & p_{22}^{(0)} & \cdots & p_{2L}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{c1}^{(0)} & p_{c2}^{(0)} & \cdots & p_{cL}^{(0)} \end{bmatrix} \tag{14}$$

satisfying (5)-(7), and set iteration index $k = 0$.

Step 2: By (7), calculate the parameter vectors $\boldsymbol{\beta}^{i(k)}$ for all $i \in \{1, \cdots, c\}$.

Step 3: Update $\mathbf{P}^{(k)}$ to $\mathbf{P}^{(k+1)}$ by [10, 11]

$$p_{ih} = \left(1 + \left[\frac{(D_{ih}(\boldsymbol{\beta}^{i(k)}))^2}{\rho^i} \right]^{\frac{1}{m-1}} \right)^{-1} \tag{15}$$

for all $i \in \{1, \cdots, c\}$ and $h \in \{1, \cdots, L\}$.

Step 4: If $\left\| \mathbf{P}^{(k)} - \mathbf{P}^{(k+1)} \right\| < \varepsilon$, stop; otherwise, set $k = k + 1$ and return to ***Step 2***.

B. Determination of the number of clusters

In the PCRM clustering algorithm, the number of clusters c is fixed and assigned by the user. In practice, the appropriate number of clusters is usually decided with the aid of a reliable index called the cluster validity criterion.

Because the PCRM clustering algorithm is similar to the FCRM clustering algorithm, we adopt the same concept of the compactness-to-separation ratio as a cluster validity criterion for the PCRM clustering algorithm. Referring to the affine linear regression models defined in (16)

$$y = f^i(\mathbf{u}, \boldsymbol{\beta}^i) = \beta_0^i + \beta_1^i u_1 + \cdots + \beta_n^i u_n. \tag{16}$$

These models can be regarded as the shift of linear hyper planes, $y = \beta_1^i u_1 + \cdots + \beta_n^i u_n$, by scale β_0^i in y-axis. By removing β_0^i, we could rewrite each linear hyper-plane, $y = \beta_1^i u_1 + \cdots + \beta_n^i u_n$, as follows:

$$\boldsymbol{\eta}^T \mathbf{n}_i = 0 \tag{17}$$

where $\boldsymbol{\eta} = [u_1, \cdots, u_n, y]^T$ and $\mathbf{n}_i = [\beta_1^i, \cdots, \beta_n^i, -1]^T \in \Re^{n+1}$. In (17), $\mathbf{n}_i$ denotes the normal vector of the ith linear hyper-plane. The corresponding unit normal vector of each linear hyper-plane in (17) can be defined as

$$\mathbf{v}_i = \frac{\mathbf{n}_i}{\|\mathbf{n}_i\|} \tag{18}$$

where $\|\bullet\|$ denotes the Euclidean norm. For any given two hyper-planes, $y = f^i(\mathbf{u}, \boldsymbol{\beta}^i)$ and $y = f^j(\mathbf{u}, \boldsymbol{\beta}^j)$, its linear hyper-planes are $\boldsymbol{\eta}^T \mathbf{n}_i = 0$ and $\boldsymbol{\eta}^T \mathbf{n}_j = 0$, respectively. If $\left| < \mathbf{v}_i, \mathbf{v}_j > \right| = 1$, then these two linear hyper-planes are coincident. If $\left| < \mathbf{v}_i, \mathbf{v}_j > \right| = 0$, then these two linear hyper-planes are orthogonal [22]. Furthermore, we define the distance of the shift term between two hyper-planes as:

$$\delta_{ij} = \frac{\left| \beta_0^i - \beta_0^j \right|}{\max\limits_{i \neq j} \left| \beta_0^i - \beta_0^j \right|}, \text{ for } i, j = 1, 2, \ldots, c \tag{19}$$

where δ_{ij} is normalized, that is, $\delta_{ij} \in [0, 1]$.

Accordingly, we define a separation validity function, V_{sep}, for the affine linear regression models by

$$V_{sep} = \min_{i \neq j} \frac{\delta_{ij} + \lambda_2}{\left| < \mathbf{v}_i, \mathbf{v}_j > \right| + \lambda_1}, \tag{20}$$

where λ_1, λ_2 are rather small real positive constants that prevents the function from being zero or being divided by zero. Obvious, V_{sep} (20) fits the concept of separation measure criterion in [9]: the more diverse the hyper-planes, the larger the separation validity function.

To reflect the compactness of clusters, the compactness validity function for PCRM, V_{com}, is defined as

$$V_{com} = \frac{\sum_{h=1}^{L} \sum_{i=1}^{c} (p_{ih})^m \left| y_h - [\mathbf{u}^T \ 1] \boldsymbol{\beta}^i \right|}{L} \tag{21}$$

which also fits the objective function of the PCRM clustering algorithm.

The cluster validity criterion suitable for PCRM clustering algorithm can be defined by the compactness-to-separation ratio as follows

$$V = \max_{i \neq j} \frac{\sum_{h=1}^{L} \sum_{i=1}^{c} (p_{ih})^m \left| y_h - [\mathbf{u}^T \ 1] \boldsymbol{\beta}^i \right|}{N \dfrac{\delta_{ij} + \lambda_2}{\left| < \mathbf{n}_i, \mathbf{n}_j > \right| + \lambda_1}} \tag{22}$$

where the numerator reflects the compactness of hyper-plane-shaped clusters, and the denominator indicates the separation of hyper-plane-shaped clusters. The optimal number of clusters c is chosen when V reaches its minimum. In practice, the appropriate number c is chosen at which the first local minimum of V has occurred; moreover, when the cluster validity index decreases monotonically, we can choose c at which a significant change in its curvature has occurred [9, 23].

C. Fuzzy rule construction

Once the appropriate number of clusters c is chosen with the aid of the cluster validity criterion (22), we can therefore construct the T-S fuzzy affine linear model from the cluster representatives and the fuzzy partitions obtained by the PCRM clustering algorithm. The fuzzy rules we want to construct for the T-S fuzzy affine linear model are expressed as follows:

$$\begin{aligned} R^i &: \text{IF } u_1 \text{ is } \mathbf{A}_1^i \text{ and } \cdots \text{ and } u_n \text{ is } \mathbf{A}_n^i \\ &\text{THEN } y^i = \beta_0^i + \beta_1^i u_1 + \cdots + \beta_n^i u_n \\ &\qquad = [\mathbf{u}^T \; 1]\boldsymbol{\beta}^i, \quad \text{for } i = 1, 2, \ldots, c. \end{aligned} \tag{23}$$

where the consequent parameters $\boldsymbol{\beta}^i = [\beta_1^i, \cdots, \beta_n^i, \beta_0^i]$ can be directly calculate by (10), but the antecedent fuzzy sets $\mathbf{A}_q^i$ need some additional manipulations. The antecedent fuzzy sets are usually achieved by projecting (the axis-orthogonal projection method [23]) the membership degrees in the fuzzy partitions matrix $\mathbf{P}$ onto the axes of individual antecedent variable u_q to obtain a point-wise defined antecedent fuzzy set $\mathbf{A}_q^i$ and then to approximate it by a normal bell-shaped membership function [18, 23- 25]. The uniform structure of bell-shaped function is advantageous and convenient for identification, analysis and optimization; besides, data of all ranges are supported by this kind of membership function.

Hence, each antecedent fuzzy set $\mathbf{A}_q^i$ in (23) is calculated from the sampled input data $\mathbf{u}_h = [u_{h1}, \cdots, u_{hn}]^T$ and the fuzzy partition matrix $\mathbf{P} = [p_{ih}]$ as follows [4, 26]:

$$\mathbf{A}_q^i(u) = \exp\left\{-\frac{1}{2}\left(\frac{u - \alpha_q^i}{\sigma_q^i}\right)^2\right\} \tag{24}$$

where

$$\alpha_q^i = \frac{\sum_{h=1}^{L} p_{ih} u_{hq}}{\sum_{h=1}^{L} p_{ih}}, \text{ and} \tag{25}$$

$$\sigma_q^i = \left(\frac{\sum_{h=1}^{L} p_{ih}(u_{hq} - \alpha_q^i)^2}{\sum_{h=1}^{L} p_{ih}}\right)^{\frac{1}{2}} \tag{26}$$

denote the mean and standard deviation of the bell-shaped membership function, respectively.

T-S Fuzzy Affine Linear Modeling Algorithm:

Step 1: Given experimental data $(\mathbf{u}_h, y_h)$, for $h = 1, \cdots, L$. Set exponential weighting $m \in [2, \infty)$. Specify the cluster representatives as affine linear models, and define the corresponding measurement error D_{ih} in PCRM clustering algorithm applied to bilinear models. Set $c = c_{\min}$ and pick a termination threshold $\varepsilon > 0$.

Step 2: Set an initial partition $\mathbf{P}^{(0)}$ satisfying (5)- (7). Set iteration index $r = 0$.

Step 3: Calculate c model parameters by (10) that minimize the objective function in (13).

Step 4: Update $\mathbf{P}^{(r)}$ to $\mathbf{P}^{(r+1)}$ by (15).

Step 5: If $\left\|\mathbf{P}^{(r)} - \mathbf{P}^{(r+1)}\right\| \le \varepsilon$, go to ***Step 6***; otherwise, set $r = r + 1$ and return to ***Step 3***.

Step 6: Set $c = c + 1$ and repeat ***Step 2*** to ***Step 6*** until $c = c_{\max}$.

Step 7: Use the cluster validity criterion (22) to find the appropriate number of clusters.

Step 8: According to the appropriate number of clusters, construct the antecedent of fuzzy model from the partition. Apply (23)- (26) to generate the antecedent and consequent parameters of the affine linear model.

IV. NUMERICAL EXAMPLES

To illustrate the T-S fuzzy affine linear modeling algorithm, we consider the following example to demonstrate the effectiveness of the modeling algorithm.

Example: Consider the discrete-time nonlinear system described by the following second-order difference equation [18]:

$$y(k+1) = \frac{y(k)y(k-1)(y(k)+2.5)}{1 + y^2(k) + y^2(k-1)} + u(k) + e(k). \tag{27}$$

where $e(k)$ is an independent white Gaussian random noise having zero mean and standard deviation 0.04.

The T-S fuzzy affine linear model is described as follows:

$$\begin{aligned} R^i &: \text{IF } y(k) \text{ is } \mathbf{A}_1^i \text{ and } y(k-1) \text{ is } \mathbf{A}_2^i \\ &\text{THEN } y^i(k+1) = [y(k) \;\; y(k-1) \;\; u(k) \;\; 1]\,\boldsymbol{\beta}^i, \end{aligned} \tag{28}$$

where $i = 1, 2, \ldots, c$.

We use the input signal with uniformly distributed white random signal as the training input signal [1], i.e. $u(k)$ is a uniformly distributed random signal in the range $[-1, 1]$ for $1 \le k \le 200$. The number of training data L is 200. The termination threshold in the PCRM clustering algorithm is chosen $\varepsilon = 0.00001$. This training input signal is depicted in Fig. 1 with $\lambda_1 = \lambda_2 = 0.001$. The plots in Fig. 2 show that the appropriate number of clusters is $c = 4$. The parameters of the antecedent and the consequent parts are listed in Table 1 and 2, respectively.

To analyze the performance of the obtained T-S fuzzy affine linear model, we use the sinusoidal signal $u(k) = \sin(2\pi k / 25)$ for $1 \le k \le 100$ to validate. The number of testing data L is 100.

Define the mean square error (MSE) as follows:

$$\text{MSE} = \frac{\sum_{k=1}^{L} \left(y(k+1) - \hat{y}(k+1) \right)^2}{L} \tag{29}$$

where $y(k+1)$ and $\hat{y}(k+1)$ are the outputs of the discrete-time nonlinear system (27) and the T-S fuzzy affine linear model (28), respectively.

The result MSE of our fuzzy modeling is 0.0689 for $L = 100$. But we use this fuzzy modeling algorithm with the FCRM clustering method, the result MSE is 0.1027. Fig. 3(a)-(c) illustrate the testing input signal and the output response, respectively.

V. CONCLUSION

In this paper, we have established the T-S fuzzy affine linear modeling algorithm by PCRM clustering. The simulation result shows that the modeling algorithm could construct the T-S fuzzy affine linear model that will approximate the nonlinear system with high accuracy and fuzzy modeling by PCRM clustering method better than by FCRM clustering method.

ACKNOWLEDGMENT

This work was supported by the National Science Council of Taiwan, R.O.C., under Grant NSC 102-2221-E-036-033.

REFERENCES

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

[2] M. Sugeno and G.T. Kang, “Structure identification of fuzzy model,” *Fuzzy Sets Syst.*, vol. 28, no. 1, pp. 15–33, 1988.

[3] R. J. Hathaway and J. C. Bezdek, “Switching regression models and fuzzy clustering,” *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 3, pp. 195-204, Aug. 1993.

[4] E. Kim, M. Park, S. Ji, and M. Park, “A new approach to fuzzy modeling”, *IEEE Trans. Fuzzy Syst.*, 1997, 5, (3), pp. 328–337.

[5] R. Krishnapuram and J. Keller, “A possibilistic approach to clustering,” *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 2, pp. 98–110, Apr. 1993.

[6] C. C. Kung, H. C. Ku, and J. Y. Su, “Possibilistic c-regression models clustering algorithm,” *Proc. of 2013 International Conference on System Science and Engineering* (ICSSE 2013), pp. 297–302, July 2013.

[7] J. C. Bezdek, “Cluster validity with fuzzy set,” *J. Cybern.*, vol. 3, no. 3, pp. 58–72, 1974.

[8] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

[9] X.L. Xie and G.A. Beni, “A validity measure for fuzzy clustering,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, no. 8, pp. 841–846, 1991.

[10] C. C. Kung, Y. F. Nieh, and J. Y. Su, “Fuzzy dynamic model identification by fuzzy c-regressoin models clustering,” *Proc. of 2010 IEEE International Conference on Systems Man and Cybernetics*, pp. 3989-3994, Oct. 2010.

[11] C. C. Kung. J. Y. Su, and Y. F. Nieh, “A novel cluster validity criterion for fuzzy C-regression models,” *Proc. of 2009 IEEE International Conference on Fuzzy Systems*, pp. 1885-1890, Aug. 2009.

[12] C. C. Kung and J. C. Hung, “A novel cluster validity criterion for fuzzy c-regression model clustering algorithm,” *Proc. of 2003 IEEE International Conference on Fuzzy Systems*, May 2003.

[13] C. C. Kung and J. Y. Su, “T-S Fuzzy Model Identification and the Fuzzy Model Based Controller Design,” *Proc. of 2007 IEEE International Conference on Systems, Man and Cybernetics*, pp. 859-864, Oct. 2007.

[14] M. Soltani, A. Chaari, and F. BenHmida, “Takagi-Sugeno fuzzy model parameters identification based on fuzzy c-regression model clustering algorithm and particle swarm optimization,” *Proc. of 2012 IEEE Mediterranean Electrotechnical Conference*, 2012.

[15] C. C. Kung and J. Y. Su, “Affine Takagi-Sugeno fuzzy modeling algorithm by fuzzy c-regression models clustering with a novel cluster validity criterion,” *IET Control Theory Appl.*, vol. 1, no. 5, pp. 1255-1265, Sep. 2007.

[16] C. C. Kung and J. Y. Su, “A Study of Cluster Validity Criteria for the Fuzzy c-Regression Models Clustering Algorithm,” *Proc. of 2007 IEEE International Conference on Syst. Man Cybern.*, pp. 853-858, Oct. 2007.

[17] Wang, L.X., *Adaptive fuzzy systems and control: design and stability analysis*, Prentice-Hall, NJ, 1994.

[18] Wang, L.X., *A course in fuzzy systems and control*, Prentice-Hall, NJ, 1997.

[19] J. Nie and D. Linkens, *Fuzzy-Neural Control: Principles, Algorithms and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[20] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, 1983.

[21] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, Ed. 2, Prentice-Hall, 1989.

[22] S.H. Friedberg, A.J. Insel, and Spence, L.E., *Linear algebra*, Prentice-Hall, NJ, 1989.

[23] R. Babuska, *Fuzzy modeling for control*, Kluwer Academic Publishers, Boston, 1998.

[24] R. Kruse, J. Gebhardt, and F. Klawonn, *Foundations of fuzzy systems*, John Wiley and Sons, Chichester, 1994.

[25] G.J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*, Prentice-Hall, NJ, 1995.

[26] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy cluster analysis, methods for classification, data analysis and image recognition*, John Wiley & Sons, NJ, 1999

TABLE I. List of antecedent parameters

	α_1^i	σ_1^i	α_2^i	σ_2^i
$i=1$	-0.7607	0.3322	-0.2915	1.0161
$i=2$	-0.6716	1.2644	-0.0794	1.0017
$i=3$	-0.3626	0.7257	-0.5009	0.8697
$i=4$	-0.3386	0.7294	-0.4001	0.2409

TABLE II. List of consequent parameters

	β_1^i	β_2^i	β_3^i	β_0^i
$i=1$	0.8621	0.8049	0.7735	0.8177
$i=2$	0.2246	-0.0149	0.0732	0.3422
$i=3$	0.3086	0.1521	0.2511	0.0221
$i=4$	0.1462	-0.0266	0.0236	0.0136

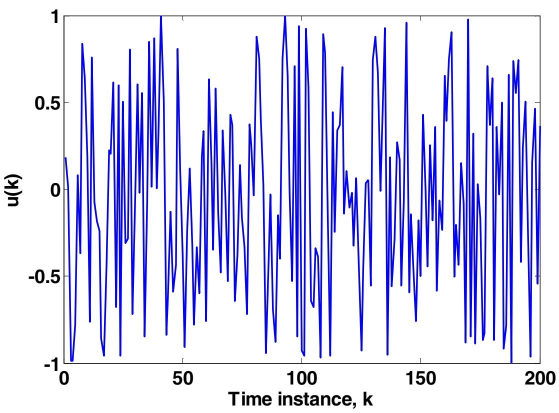

Fig. 1. The training input $u(k)$.

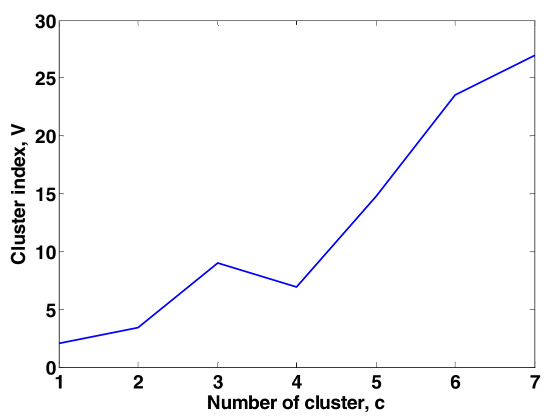

Fig. 2. Plot of the cluster index V vs. the number of cluster c.

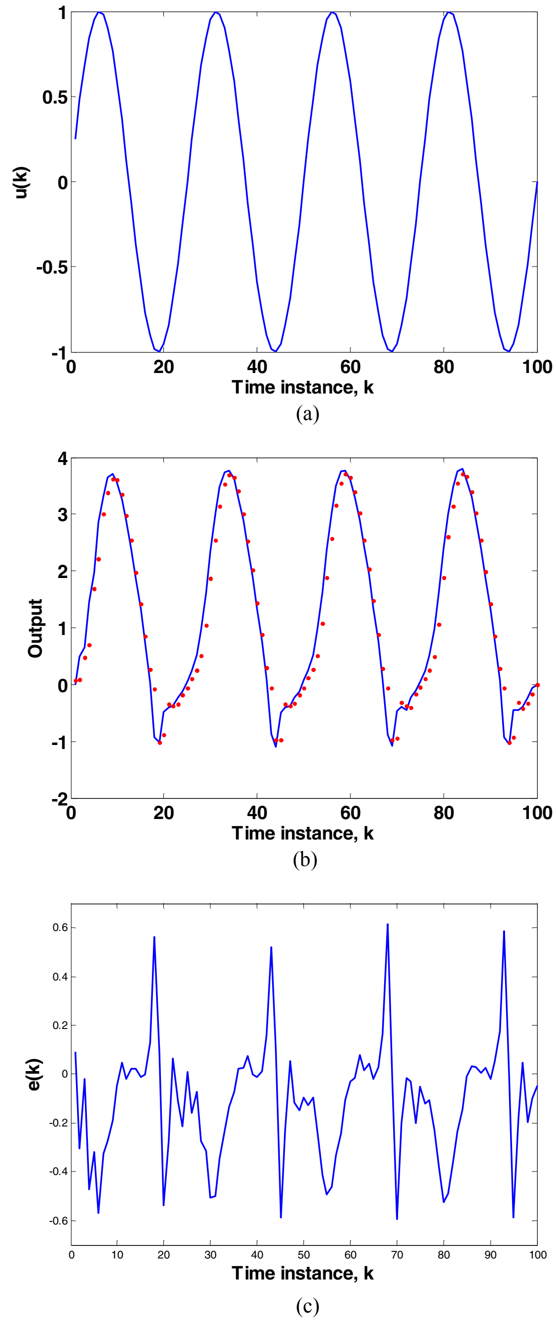

Fig. 3. (a) The testing input $u(k)=\sin(2\pi k/25)$. (b) The output of the T-S affine linear model (dotted line) and the plant (solid line). (c) Errors between the outputs of fuzzy bilinear model and plant for test input with $L = 100$ (MSE = 0.0689).