Proceedings of the 26th Chinese Control Conference
July 26-31, 2007, Zhangjiajie, Hunan, China

Online Multivariable Identification of a MIMO Distillation Column Using Evolving Takagi-Sugeno Fuzzy Model

Molazem Sanandaji Borhan [1], Salahshoor Karim [1, 2]

1. Automation and Instrumentation Engineering Department, Petroleum University of Technology, Tehran, Iran
E-mail: borhan_molazemsanandaji@yahoo.com
2. E-mail: salahshoor@put.ac.ir

Abstract: In this paper, an evolving Takagi-Sugeno (eTS) fuzzy model has been utilized for online identification of a multi-input, multi- output (MIMO) distillation column. In this approach, the rule-base structure and the model parameters of the consequent parts of fuzzy IF-THEN rules gradually evolve during the online identification process. In addition, an exponential time-varying weight is included in the original rule generation condition in order to control the rate of rule generation at the start of the training process and consequently reduce the total number of generated rules in comparison with the original MIMO eTS algorithm. Recursive-Least Squares (RLS) algorithm is employed to estimate the consequent part of each rule. The results show that the modified condition reduces the total number of generated rules for a certain data set with lower RMSE error in comparison with the original eTS method.
Key Words: Systen Identification,Evolving Takagi-sugeno,Fuzzy Systems, Distillation Column

1 INTRODUCTION

Fuzzy model identification is an effective tool for the approximation of nonlinear dynamical systems on the basis of measured data [2]. This approach has been popular for the past years due to its ability to utilize heuristic knowledge to provide quantitative model which can accurately represent complex nonlinear systems.

Among different fuzzy modeling techniques, Takagi-Sugeno (TS) model [1] has received a great deal of attention and has been employed in many applications in nonlinear system identification. This is mainly due to its good results in different applications and also because it employs mathematical functions as rule consequent parts. This model consists of IF-THEN rules with fuzzy antecedents and mathematical functions in the consequent parts. This structure gives the ability to utilize the input-output data in an efficient way.

In the offline mode all of the data are available at the start of the training process. In the online mode, however, we do not have the whole data at the start of the training process. Thus learning of the TS fuzzy models should be started with the first data sample [4,9]. In this condition, the model structure is not known a priori, but instead it evolves gradually during the identification process. Then we come to the concept of Evolving Takagi-Sugeno (eTS) fuzzy models [4,9]. In eTS, the potential of the new data sample is used as a trigger to update the rule-base. It is important to note that learning could start without a priori information and only with a single data sample. This interesting feature makes the approach potentially very useful in many smart adaptive systems [4,9].

In this paper, the original rule generation condition is modified so as to efficiently control the rate of rule generation especially at the start of the training process. This leads to a reduction in the total number of generated rules for a certain data history with a better accuracy, expressed by lower RMSE error, in comparison with the original eTS algorithm.

This technique is applied to a simulated nonlinear MIMO distillation column as a popular benchmark problem in nonlinear identification [3, 5, 7]. The paper is organized as follows. In section 2, the MIMO Takagi-Sugeno fuzzy model is introduced. In section 3, the Recursive-Least Squares (RLS) algorithm is presented. The eTS method is described in section 4. The distillation column model is given in section 5. In section 6, our modification for eTS is presented and then these methods are applied to the distillation column. Some brief conclusion remarks are given in section 7.

2 MIMO TS FUZZY MODEL

In this paper, a MIMO extension of the TS fuzzy model is considered [6, 9]. Therefore, the i^{th} rule of the MIMO TS fuzzy model has the following form:

R^i :

IF x_1 *is* A_1^i *AND* x_2 *is* A_2^i *AND...AND* x_n *is* A_n^i

THEN

$$y^i = X_e^{\mathrm{T}}\pi^i \quad \text{i=1,2,…,M} \tag{1}$$

where,

$x_1, x_2, ..., x_n$ are input variables;

n is the total number of input variables;

$y^i = \left[y_1^i, y_2^i, ..., y_m^i \right]$ is the output of i^{th} rule;

m is the total number of output variables;

M is the total number of rules;

$A_j^i \left(A_1^1 ... A_n^M \right)$ are the linguistic terms of fuzzy sets; In (1), R^i denotes the i^{th} fuzzy rule; X_e^{T} is the extended input vector, $X_e^{\mathrm{T}} = \left[1, X^{\mathrm{T}} \right]$, which is formed by appending the input vector $X = \left[x_1, x_2, ..., x_n \right]^{\mathrm{T}}$ with 1 to define a free bias parameter for each rule[4,9].

As a MIMO extension of the TS model, y^i in the consequent part of the fuzzy rule denotes the multidimensional vector of the i^{th} linear sub-system [9]. Similarly,

the parameters in the consequent parts will be denoted by [4], [9]:

$$\pi^i = \begin{bmatrix} a_{01}^i & a_{02}^i & \dots & a_{0m}^i \\ a_{11}^i & a_{12}^i & \dots & a_{1m}^i \\ \dots & \dots & \ddots & \dots \\ a_{n1}^i & a_{n2}^i & \dots & a_{nm}^i \end{bmatrix} \quad (2)$$

Symmetrical Gaussian functions are used as antecedent fuzzy membership functions as follows:

$$\mu^i(x) = \exp\left(-4\left(\frac{x - x^{i*}}{r}\right)^2\right), \text{i=1,2…M} \quad (3)$$

where r is a positive constant which defines the radius of the antecedent and the zone of influence of the i^{th} model; x^{i*} is the focal point associated with the center of the i^{th} rule antecedent. A value of r in the range of [0.3; 0.5] has been recommended [9].

In order to calculate the crisp output, for the fuzzifier, Mamdani MIN operator can be used, and defuzzifiction may be obtained using weighted average method as follows:

$$y = \frac{\sum_{i=1}^{M} y^i \alpha^i}{\sum_{i=1}^{M} \alpha^i} \quad \text{(Crisp Output)} \quad (4)$$

where

$$\alpha^i = \prod_{j=1}^{n} \mu_{A_j^i}(x_j) \quad (5)$$

α^i is the fulfillment degree of the i^{th} rule.

Equation (4) can be rewritten as follows. Defining:

$$\xi^i = \frac{\alpha^i}{\sum_{i=1}^{M} \alpha^i} \quad (6)$$

Equation (4) becomes:

$$y = \sum_{i=1}^{M} y^i \xi^i \quad (7)$$

Then, by substituting (1) in (7), it gives:

$$y = \sum_{i=1}^{M} \left(X_e^{\mathrm{T}} \pi^i\right) \xi^i \quad (8)$$

3 RECURSIVE−LEAST SQUARES (RLS) ALGORITHM

Recursive-Least Squares (RLS) algorithm is used to estimate the parameters in matrix π^i, defined in (2), which are dependent on the values of the membership functions. In order to use RLS identification algorithm, (8) can be expanded as follows[9]:

$$y = \xi^1 X_e^{\mathrm{T}} \pi^1 + \xi^2 X_e^{\mathrm{T}} \pi^2 + \dots + \xi^M X_e^{\mathrm{T}} \pi^M \quad (9)$$

which can be defined in the form of two new terms φ and θ as :

$$\varphi = \left[\xi^1 X_e^{\mathrm{T}}, \xi^2 X_e^{\mathrm{T}}, \dots, \xi^M X_e^{\mathrm{T}}\right]^{\mathrm{T}} \quad (10)$$

$$\theta = \left[\left(\pi^1\right)^{\mathrm{T}}, \left(\pi^2\right)^{\mathrm{T}}, \dots, \left(\pi^M\right)^{\mathrm{T}}\right]^{\mathrm{T}} \quad (11)$$

Consequently, (8) can be rewritten as follows:

$$y = \varphi^{\mathrm{T}} \theta \quad (12)$$

For a given set of input-output data $\left(X_k^{\mathrm{T}}, y_k\right)$ where k= [1, N] and N is the number of training data, θ should minimize the following globally optimal objective function [8]:

$$J = \sum_{k=1}^{N} \left(y_k - \varphi_{k-1}^{\mathrm{T}} \theta_{k-1}\right)^2 \quad (13)$$

4 EVOLVING TAKAGI−SUGENO MODEL

In Online mode, the training data are collected continuously, rather than being a fixed set and the identification process starts with the first data point. Online learning of eTS models includes online clustering under assumption of a gradual change of the rule-base. Therefore, the numbers of fuzzy rules grow during the identification process and are not fixed beforehand. So a modified Recursive-Least Squares (RLS) algorithm will be required [4, 9].

4.1 Online Potential Clustering Approach

The online clustering procedure starts with the first data point. This point is also considered as the center of the first cluster. Its coordinates are used to form the antecedent part of the fuzzy rule (1) using Gaussian membership functions (3). Its potential is assumed to 1.

Starting from the next data point onwards the potential of new data points is calculated recursively. As a measure of potential, we use a Cauchy type function of first order as follows [4,9]:

$$P_k(z_k) = \frac{1}{1 + \frac{1}{(k-1)} \sum_{i=1}^{k-1} \sum_{j=1}^{n+m} \left(d_{ik}^j\right)^2}; k = 2,3,\dots \quad (14)$$

where $P_k(z_k)$ denotes the potential of the data point z_k calculated at time k; $d_{ik}^j = z_i^j - z_k^j$,denotes the distance between two data points at different sample times i, k for different inputs(x^j for j=1,2,…,n) and outputs(y^j for j=n+1,n+2,…, n+m).

After the new data are available in online mode, they influence the potentials of the centers of the previous clusters ($z^{i*}, i = 1,2,\dots,M$). The reason is that by definition the potential depends on the distance to all data points, including the new ones. So the potential of the existing rule centers should be updated when a new data point is added. This can be done by using the following equation recursively:

$$P_k(z^{i*}) = \frac{(k-1) P_{k-1}(z^{i*})}{k - 2 + P_{k-1}(z^{i*}) + P_{k-1}(z^{i*}) \sum_{j=1}^{n+m} \left(d_{k(k-1)}^j\right)^2} \quad (15)$$

where $P_k(z^{i*})$ is the potential of the existing rule centers which is updated at the sample time k.

Now after calculating the potential of the new data point and also updating the potential of the existing rule centers, the potential of the new data point should be compared to the updated potential of the existing rule centers whether to add a new rule or to modify the existing rule-base by replacement. The evolution of the rule-base is conducted by the following two basic principles:

CASE 1 (REPLACMENT):
IF the potential of the new data point is higher than the maximum of the updated potential of the existing rule centers:

$$P_k(z_k) > \max_{i=1}^{M} P_k(z^{i*}) \tag{16}$$

AND z_k is close to an existing rule center:

$$\frac{P_k(z_k)}{\max_{i=1}^{M} P_k(z^{i*})} - \frac{\delta_{\min}}{r} > 1 \tag{17}$$

where $\delta_{\min} = \min_{i=1}^{M} \left\| z_k - z^{i*} \right\|$ denotes the distance between the new data point and the closest rule center of the existing rule-base.
THEN the new data point z_k replaces this closest center which we assume its index h as follows:

$$z^{h*} = z_k; P_k(z^{h*}) = P_k(z_k) \tag{18}$$

CASE 2 (ADD):
IF only (16) is satisfied but not (17)
THEN the new data point is added to the rule-base as a new rule center as follows:

$$M = M + 1; z^{M*} = z_k; P_k(z^{M*}) = P_k(z_k) \tag{19}$$

4.2 Online Recursive Estimation of Consequent Parameters of eTS

Because in eTS, the rule-base gradually evolves, the straightforward application of the RLS is not applicable. A resetting of the covariance matrices and parameters of the RLS should be done each time a new rule is added to the rule-base. In the case when the globally optimal objective function is minimized (13), the RLS algorithm is conducted in the following steps [8]:
1) Form the new φ_k^{T} at each sample time and then evaluate the following gain vector named as Kalman gain vector:

$$K_k = C_k \varphi_k = \frac{C_{k-1}\varphi_k}{1 + \varphi_k^{\mathrm{T}} C_{k-1} \varphi_k} \tag{20}$$

2) Update the parameter matrix θ_k:

$$\theta_k = \theta_{k-1} + K_k e_k \tag{21}$$

where

$$e_k = y_k - \varphi_k^{\mathrm{T}} \theta_{k-1}$$

3) Update the covariance matrix C_k:

$$C_k = C_{k-1} - \frac{C_{k-1}\varphi_k \varphi_k^{\mathrm{T}} C_{k-1}}{1 + \varphi_k^{T} C_{k-1} \varphi_k} = \left[\mathrm{I} - K_k \varphi_k^{\mathrm{T}} \right] C_{k-1} \tag{22}$$

where $\theta_1 = \left[\left(\pi^1\right)^{\mathrm{T}}, \left(\pi^2\right)^{\mathrm{T}}, ..., \left(\pi^M\right)^{\mathrm{T}} \right]^{\mathrm{T}} = 0; C_1 = \Omega \mathrm{I}$

In the case when a new rule is added to the rule-base, the RLS will be reset in the following way[4,9]:
1) Parameters of the new rule ($\hat{\pi}_k^{M+1}$) are determined by the weighted average of the parameters of the other existing rules. The weights are the normalized firing levels of the existing rules. Parameters of the other rules are inherited from the previous step[4]:

$$\hat{\theta}_k = \left[\left(\hat{\pi}_{k-1}^{1}\right)^{\mathrm{T}}, \left(\hat{\pi}_{k-1}^{2}\right)^{\mathrm{T}}, ..., \left(\hat{\pi}_{k-1}^{M}\right)^{\mathrm{T}}, \left(\hat{\pi}_{k}^{M+1}\right)^{\mathrm{T}} \right]^{\mathrm{T}} \tag{23}$$

where $\hat{\pi}_k^{M+1} = \sum_{i=1}^{M} \xi_{k-1}^{i} \hat{\pi}_{k-1}^{i}$

2) Covariance matrices are reset as follows:

$$C_k = \begin{bmatrix} \rho\zeta_{1,1} & \cdots & \rho\zeta_{1,R(n+1)} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho\zeta_{R(n+1),1} & \cdots & \rho\zeta_{R(n+1),R(n+1)} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \Omega & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \Omega \end{bmatrix} \tag{24}$$

where ζ_{ij}, i=1,2,…,$R(n$+1); j=1,2,…,$R(n$+1) is an element of the covariance matrix at previous sample time; $\rho = \frac{R^2+1}{R^2}$ [4].
In the case when a rule is replaced with another one, the covariance matrices and parameters are inherited from the previous time step [4,9]. Finally, in the same loop and after estimating and updating the parameters, the next value of the outputs can be predicted online as follows:

$$\hat{y}_{k+1} = \varphi_k^{\mathrm{T}} \hat{\theta}_k \quad k=2,3\ldots \tag{25}$$

5 DISTILLATION COLUMN CASE STUDY

The process to be identified is a first-principle model of a binary distillation column (see Fig. 1).The column is referred to as "column A" which has been studied in several papers [3]. The simulated system covers the most important effects for the dynamic of a real distillation column. Further details of the simulated process model are described in [3], [5], and [7].
To identify the simulated process, an appropriate model should be selected. It is assumed that the process under study can be represented by the following one-step prediction model [5]:

$$\hat{y}(k) = f\left[y(k-1), y(k-2), u(k-1) \right] \tag{26}$$

This structure is selected due to its intrinsic nonlinear characteristics and its generality.

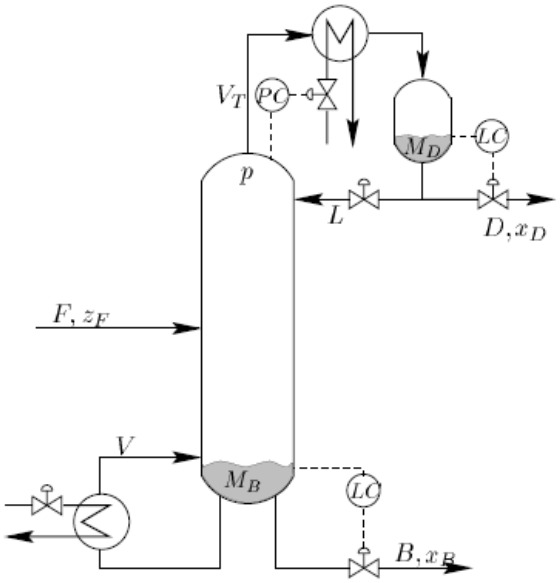

Fig.1 Distillation Column

In the identification procedure, L(reflux flow rate) and V (boilup flow rate) are used as inputs while M_B (reboiler holdup) and M_D (condenser holdup) are used as outputs.

Fig. 2 shows the block diagram of the assumed input-output variables for the fuzzy model identification of the simulated system. As shown, by considering (26), there are 6 inputs and 2 outputs for the identification scheme.

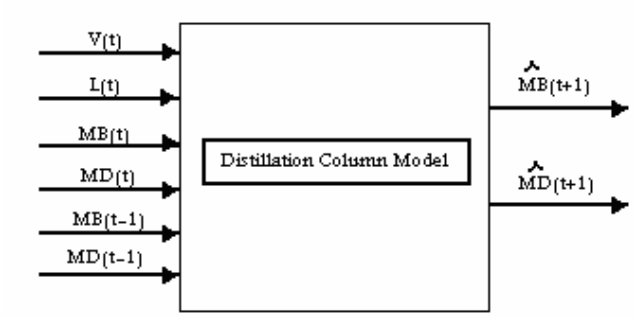

Fig.2 Block diagram of distillation column fuzzy model

6 APPLICATION TO DISTLLATION COLUMN

Signals that have applied as input changes in boil up flow V and reflux flow L are chosen such that they will generate data with good enough signal to noise ratio but will not disturb the product quality. The steady state values of the proposed distillation column are given in [3]. These values are used for the process initial conditions. From 600 samples of data, 300 are used for identification of fuzzy parameters and the rest will be used for evaluation purpose. The only pre-specified parameters in the algorithm are $r = 0.5$ and $\Omega = 10^3$.

6.1 Application of the Original eTS Method

In Fig. 3 and Fig. 4, the corresponding outputs M_B and M_D obtained using original eTS fuzzy model, are depicted. The absolute error for each of the outputs is also depicted in Fig. 5 and Fig.6 The evolution of the generated rules during the online identification process is also shown in Fig. 7. As a measure of performance, the Root Mean Square Error is used as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(y_k - \hat{y}_k\right)^2}{N}} \qquad (27)$$

where N is the total number of training data samples, y_k is the actual output at the sample time k and $\hat{y}_k$ is the estimated output at the sample time k.

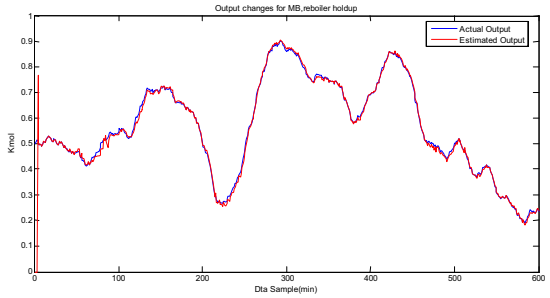

Fig.3 Output changes for MB (*t*), reboiler holdup

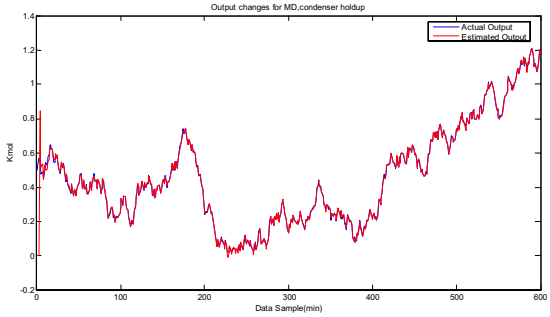

Fig.4 Output changes for MD (*t*), condenser holdup

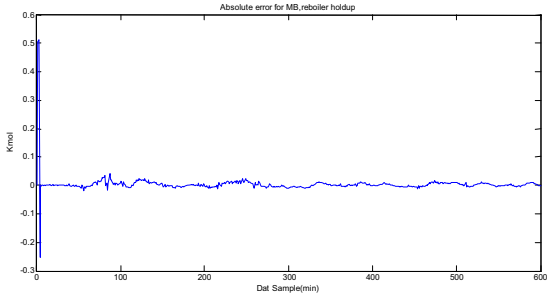

Fig.5 Absolute error for MB (*t*), reboiler holdup

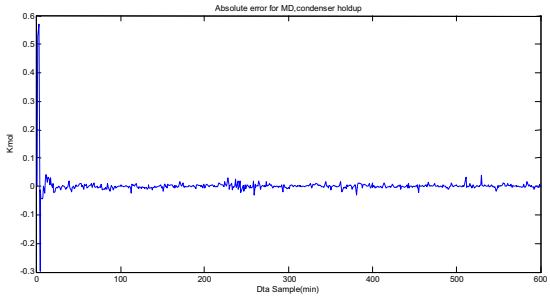

Fig.6 Absolute error for MD (*t*), condenser holdup

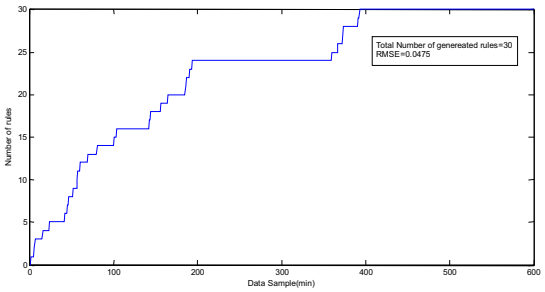

Fig.7 Evolution of the rule-base and rule generation

6.2 Application of the Modified eTS Method

The following exponential time-varying weight (29) is included in (16) in order to control the rate of rule generation efficiently. This modification makes the algorithm add new rules with more caution in the initial identification phases. Then, as the identification algorithm progresses and more dynamic knowledge are captured, the rule generation condition gets back exponentially to its original lower decision level. So (16) changes as follows:

$$P_k(z_k) > \alpha(k) \times \max_{i=1}^{M} P_k(z^{i*}) \tag{28}$$

where

$$\alpha(k) = 1 + (\alpha_0 - 1) \times \exp\left(-\frac{k}{\tau}\right) \tag{29}$$

where α_0 and τ are chosen as follows:

$\alpha_0 = 1.05$; $\tau = \frac{N}{5}$ (N is the total number of training data samples). Of course, this strategy can be resumed whenever a significant dynamic change, detected by the residual error, is occurred. The time evolution of the generated rules has been illustrated in Figure 8.

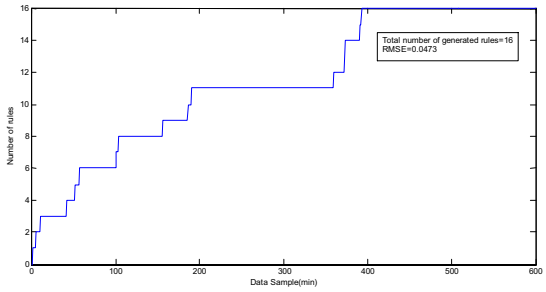

Fig.8 Evolution of rule-base by exerting the modified condition

As shown, the generation rate becomes more efficient leading to a reduction in the total number of generated rules. Tab. 1 shows a comparison between the original and the modified method. The results demonstrate that the obtained reduction in the number of generated rules does not attenuate the accuracy.

Tab.1 RMSE and Total Number of Generated Rules Comparison

	Total Number of Generated Rules	RMSE
Original Method	30	0.0475
Modified Method	16	0.0473

7 CONCLUSION

This paper presents the application of the eTS fuzzy model for online identification of a MIMO distillation column benchmark problem. An exponential time-varying weight has been proposed to enhance the rule generation mechanism. The modified approach adds rules cautiously at the start of identification and gradually becomes more flexible when more dynamic knowledge is captured. The simulation results indicate the efficient performance of the resulting identification method with lower total generated rules.

REFERENCES

[1] Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control[J]. IEEE Transactions on Systems, Man and Cybernetics,1985, 15(1):116-132.

[2] Hellendoorn H, Driankov D, Fuzzy Model Identification: Selected Approaches[M]. Berlin: Springer, Germany, 1997.

[3] Skogestad S, MATLAB Distillation Column Model ("ColumnA"). Available: http://www.nt.ntnu.no/users/skoge/book/matlab_m/cola/cola.html.

[4] Angelov P P, Filev D P. An approach to online identification of Takagi-Sugeno models[J]. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, 2004, 34(1).

[5] Skogestad S, Morari M. Understanding the dynamic behavior of distillation columns[J]. Industrial and Engineering Chemistry Research, 1988, 27:1848-1862.

[6] Mollov S, Babuska R, Verbruggen H B. Analysis of Interactions in MIMO Takagi-Sugeno Fuzzy Models[J]. In Proc. FUZZ- IEEE, 2001: 769-773.

[7] Skogestad S.Dynamics and control of distillation columns. Chem. Eng. Res. Des. (Trans IChemE), 1997, 75:539-562.

[8] Ljung L, System Identification: Theory for the User[M]. Prentice-Hall, Englewood Cliffs, N.J, 1987.

[9] Angelov P, Xydeas C S, Filev D. Online identification of MIMO evolving Takagi-Sugeno fuzzy models. Proceedings of the IJCNN-FUZZ-IEEE, Budapest, Hungary, July, 2004: 55-60.