HUNGARIAN JOURNAL
OF INDUSTRIAL CHEMISTRY
VESZPRÉM
Vol. 29. pp. 129 - 134 (2001)

IDENTIFICATION OF NONLINEAR SYSTEMS USING GAUSSIAN MIXTURE OF LOCAL MODELS

J. ABONYI, T. CHOVAN and F. SZEIFERT

(Department of Process Engineering, University of Veszprém, P.O. Box 158, H-8201, HUNGARY)

Received: October 8, 2001

Identification of operating regime based models of nonlinear dynamic systems is addressed. The operating regimes and the parameters of the local linear models are identified directly and simultaneously based on the Expectation Maximization (EM) identification of Gaussian Mixture Model (GMM). The proposed technique is demonstrated by means of the identification of a neutralization reaction in a continuously stirred tank reactor.

Keywords: Operating regime based model, expectation maximization, Takagi-Sugeno fuzzy model, nonlinear system, neutralization reaction

Introduction

The problem of a successful model based control application arises from difficulties in system modeling [1, 2]. This difficulty stems from lack of knowledge or understanding of the process to be controlled [3]. While it may not be possible to find process information that is universally applicable, it would certainly be worthwhile to examine what types of process-knowledge would be most relevant for specific operating points of the process. This type of local understanding, in fact, will be a key to identifying reliable local models with a limited amount of data. The model that has a range of validity less than the operating regime of the process is called *local model*, as opposed to a *global model* that is valid in the full range of operation. Global modeling is a complicated task because of the need to describe the interactions between a large number of phenomena that appear globally. Local modeling, on the other hand, may be considerably simpler, because locally there may be a smaller number of phenomena that are relevant, and their interactions are simpler [4]. The modeling framework that is based on combining a number of local models, where each local model has a predefined operating region in which the local model is valid is called *operating regime based model* [5], where the local models are combined into a global model using an interpolation technique as it is illustrated in *Fig.1*.

The main advantage of this framework is its transparency. Both the concept of operating regimes and the model structure are easy to understand. This is important, since the model structure can be interpreted in terms of operating regimes, but also quantitatively in terms of individual local models.

The operating regime of the local models can be also represented by fuzzy sets [6]. This representation is appealing, since many systems change behaviors smoothly as a function of the operating point, and the soft transition between the regimes introduced by the fuzzy set representation captures this feature in an elegant fashion.

Fuzzy modeling and identification proved to be effective tools for the approximation of uncertain nonlinear systems because of the ability to combine expert knowledge and measured data. Fuzzy models use if-then rules to describe the process through a collection of locally valid relationships. The antecedents (if-parts) of the rules divide the input space into several fuzzy subspaces, while the consequents (then-parts) describe the local behavior of the system in these fuzzy subspaces [7]. In this paper the local models are linear.

The contribution of this paper is two-fold:

- A new method for the identification of operating regime based models is proposed based on EM identification of Gaussian Mixtures Model.
- Method to transform the obtained model into Takagi-Sugeno fuzzy model is presented.

The paper is organized as follows. Section 2 presents the structure of the operating regime based model along with the methods for its transformation into a fuzzy model. In Section 3, the identification algorithm of the

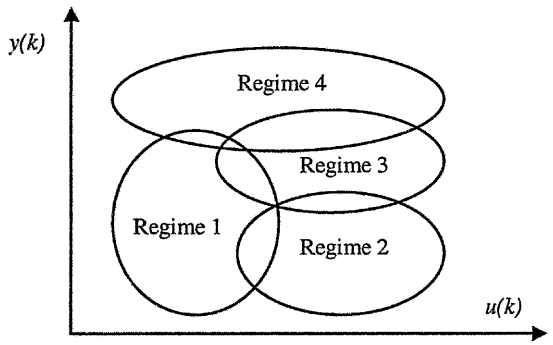

Fig.1 Example for an operating regime based model. The operating region defined by the current input $u(k)$ and output $y(k)$ of the system is decomposed into four regimes.

model is proposed. An application example - the identification of a pH process is given in Section 3. Conclusions are given in Section 4.

Operating Regime based Modeling of Dynamical System

Nonlinear dynamic systems are often represented in the Nonlinear AutoRegressive with eXogenous input (NARX) model form, which establishes a nonlinear relationship between the past inputs and outputs and the predicted output:

$$y(k+1)=f\big(y(k),\ldots,y(k-n_y),u(k-n_d),\ldots,u(k-n_u)\big) \quad (1)$$

Here, n_y and n_u denote the maximum lags considered for the output, and input terms, respectively, $n_d < n_u$ is the discrete dead time, and f represents the mapping of the NARX model.

The aim of this paper is to develop an algorithm for the identification of a transparent and easily interpretable model

$$y_k = f(\mathbf{x}_k) \quad (2)$$

based on some available training pattern y_k and $\mathbf{x}_k=[x_{1,k},\ldots,x_{n,k}]^T$, where $k=1,\ldots,N$ denotes the index of the k-th data that can be used for identification. When *Eq.(2)* is used to represent a NARX model, the training pattern is formed to have a model that gives one-step-ahead prediction: $y_k = y(k+1)$,

$\mathbf{x}_k=[y(k),\ldots,y(k-n_y),u(k-n_d),\ldots,u(k-n_u)]^T$.

The operating regime based model of the system given by *Eq.(2)* is formulated as:

$$\hat{y}_k = \sum_{i=1}^{c} \phi_i(\mathbf{x}_k)\,(\mathbf{a}_i\,\mathbf{x}_k + b_i) \quad (3)$$

where the $\phi_i(\mathbf{x}_k)$ function describes the operating regime of the i-th local linear model defined by the $\boldsymbol{\theta}_i=[\mathbf{a}_i^T, b_i]^T$ parameter vector.

The operating regime of the local models can also be represented by fuzzy sets [6]. Hence, the entire global model *Eq.(3)* can be conveniently represented by Takagi-Sugeno fuzzy rules:

$$R_i: \text{ If } \mathbf{x}_k \text{ is } \mathbf{A}_i(\mathbf{x}_k) \text{ then } \hat{y}_k = \mathbf{a}_i\,\mathbf{x}_k + b_i, \quad i=1,\ldots,c \quad (4)$$

where $\mathbf{A}_i(\mathbf{x}_k)$ represents the multivariable membership function that describes the fuzzy set $\mathbf{A}_i$ while $\mathbf{a}_i$ and b_i are the parameters of the local linear model.

Usually, the antecedent proposition "$\mathbf{x}_k$ is $\mathbf{A}_i(\mathbf{x}_k)$" is expressed as a logical combination of simple propositions with univariate fuzzy sets defined for the individual components of $\mathbf{x}_k$, often in a conjunction form:

$$R_i: \text{ If } x_{1,k} \text{ is } A_{i,1}(x_{1,k}) \text{ and } \ldots \text{ and } x_{n,k} \text{ is } A_{i,n}(x_{n,k}) \text{ then } \hat{y}_k = \mathbf{a}_i\,\mathbf{x}_k + b_i \quad (5)$$

In this case, the degree of fulfillment of a rule is calculated as the product of the degree of fulfillment of the fuzzy sets in the rule

$$\beta_i(\mathbf{x}_k) = \mathbf{A}_i(\mathbf{x}_k) = \prod_{j=1}^{n} A_{i,j}(x_{j,k}) \quad (6)$$

The rules of the fuzzy model are aggregated using the fuzzy mean formula

$$\hat{y}_k = \frac{\sum_{i=1}^{c} w_i\,\beta_i(\mathbf{x}_k)\,(\mathbf{a}_i\mathbf{x}_k + b_i)}{\sum_{i=1}^{c} w_i\,\beta_i(\mathbf{x}_k)} \quad (7)$$

where $w_i=[0,1]$ is the weight of the rule that represents the desired impact of the rule. The value of w_i is often chosen by the designer of the fuzzy system based on his or her belief in the goodness and accuracy of the i-th rule. When such knowledge is not available w_i is set as $w_i=1, \forall i$.

As *Eq.(7)* and *Eq.(3)* show, fuzzy models are identical to operating regime based models as the operating region of the local linear models are defined by the normalized rule fulfillments:

$$\phi_i(\mathbf{x}_k) = \frac{w_i\,\beta_i(\mathbf{x}_k)}{\sum_{i=1}^{c} w_i\,\beta_i(\mathbf{x}_k)} \quad (8)$$

To represent the $A_{i,j}(x_{j,k})$ fuzzy set, in this paper Gaussian membership function is used

$$A_{i,j}(x_{j,k}) = \exp\left(-\frac{1}{2}\frac{(x_{j,k}-v_{i,j})^2}{\sigma_{i,j}^2}\right) \quad (9)$$

where $v_{i,j}$ represents the center and $\sigma_{i,j}^2$ the variance of the Gaussian function.

The use of Gaussian membership functions allows the compact formulation of *Eq.(6)*:

$$\beta_i(\mathbf{x}_k) = \mathbf{A}_i(\mathbf{x}_k) = \exp\left(-\frac{1}{2}(\mathbf{x}_k - \mathbf{v}_j^x)^T (\mathbf{F}_i^{xx})^{-1} (\mathbf{x}_k - \mathbf{v}_j^x)\right) \quad (10)$$

where $\mathbf{v}_j^x = [v_{1,j}, \ldots, v_{n,j}]$ represents the center of the i-th multivariate Gaussian and $(\mathbf{F}_i^{xx})^{-1}$ stands for the inverse of a diagonal matrix that contains the variances:

$$\mathbf{F}_i^{xx} = \begin{bmatrix} \sigma_{1,i}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2,i}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{n,i}^2 \end{bmatrix} \quad (11)$$

This compact formulation of the operating regime based model suggests that the model is not only equivalent to a TS fuzzy model but it is functionally identical to Generalized Radial Basis Function Network (GBFN) [8].

When there is a correlation among input variables of the model, $\mathbf{F}_i^{xx}$ is not a diagonal matrix as it was shown in *Eq.(11)*. In this case the decomposition of $\mathbf{A}_i(\mathbf{x}_k)$ to $A_{i,j}(x_{j,k})$ fuzzy sets by

$$\mathbf{A}_i(\mathbf{x}_k) = \prod_{j=1}^{n} A_{i,j}(x_{j,k}) = \prod_{j=1}^{n} \exp\left(-\frac{1}{2}\frac{(x_{j,k} - v_{i,j})^2}{\sigma_{i,j}^2}\right) \quad (12)$$

is not possible directly. The proposed method to solve this problem is based on the eigenvector projection [9] or transformed input-domain approach [10]. The approach is based on the calculation of the eigenvalues λ_j^i and the eigenvectors $\mathbf{t}_j^i$, of the $\mathbf{F}_i^{xx}$ matrix, where $j = 1, \ldots, n$. Using the eigenvalues and the eigenvectors the following fuzzy model that has no correlation in its transformed input space can be obtained:

$$R_i: \text{ If } (\mathbf{t}_1^i)^T \mathbf{x}_k \text{ is } A_{i,1} \text{ and} \ldots \text{and } (\mathbf{t}_n^i)^T \mathbf{x}_k \text{ is } A_{i,n} \quad (13)$$
$$\text{then } \hat{y}_k = \mathbf{a}_i \mathbf{x}_k + b_i$$

where the Gaussian membership functions are defined as

$$A_{i,j}(x_{j,k}) = \exp\left(-\frac{1}{2}\frac{\left((\mathbf{t}_j^i)^T \mathbf{x}_k - (\mathbf{t}_j^i)^T \mathbf{v}_i^x\right)^2}{(\lambda_j^i)^2}\right) = \quad (14)$$
$$= \exp\left(-\frac{1}{2}\frac{(\tilde{x}_{j,k} - \tilde{v}_{i,j})^2}{\tilde{\sigma}_{i,j}^2}\right)$$

where $\tilde{x}_{j,k} = (\mathbf{t}_j^i)^T \mathbf{x}_k$, $\tilde{v}_{i,j} = (\mathbf{t}_j^i)^T \mathbf{v}_i^x$ and $\tilde{\sigma}_{i,j}^2 = (\lambda_j^i)^2$ denote the transformed input variable, the cluster center and variance, respectively.

The aim of the remaining part of the paper is to propose a new identification technique for the identification of the model presented above.

EM algorithm for Identification of Mixture of Gaussians Model

The Expectation Maximization (EM) algorithm is an iterative algorithm for the computation of maximum likelihood parameter estimates when the observations can be viewed as incomplete data. The EM algorithm is widely used for parameter estimation of the mixture of models, in particular the mixture of Gaussians model [11].

The basics of EM are the following. Suppose we know the observed values of a random variable $\mathbf{z}$ and we wish to model the density of $\mathbf{z}$ using a model parameterized by η.

Each observation consists of $n+1$ measured variables, grouped into an $n+1$-dimensional column vector $\mathbf{z}_k = [z_{1,k}, \ldots, z_{n+1,k}]^T$. A set of N observations is denoted by $\mathbf{Z}$ and represented as a matrix:

$$\mathbf{Z} = \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,N} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n+1,1} & z_{n+1,2} & \cdots & z_{n+1,N} \end{bmatrix} \quad (15)$$

As the identification is performed on the available identification data, $\mathbf{Z}$ is divided into a regression data matrix $\mathbf{X}$ and a regression vector $\mathbf{y}$

$$\mathbf{X} = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (16)$$

$$\mathbf{Z}^T = [\mathbf{X}\mathbf{y}] \quad (17)$$

In the pattern recognition terminology, the columns of $\mathbf{Z}$ called patterns or objects, the rows are called the features or attributes, and $\mathbf{Z}$ is called the pattern matrix.

EM obtains parameter estimates $\hat{\eta}$ which maximize the likelihood $L(\eta) = p(\mathbf{z}|\eta)$ of the data. The EM assumes that this estimation is intractable and the values of a missing or hidden random variable $\mathbf{h}$ would make the problem more tractable. Let $p(\mathbf{z},\mathbf{h}|\eta)$ denote the joint probability of $\mathbf{z}$ and $\mathbf{h}$ parameterized by η. It is assumed that $\mathbf{z}$ and $\mathbf{h}$ are such that maximizing the complete data likelihood $L_c(\eta) = p(\mathbf{z},\mathbf{h}|\eta)$ is more tractable than maximizing $L(\eta)$. However, the values of $\mathbf{h}$ are not known. The EM algorithm tackles this problem by iteratively generating a probability over the values $\mathbf{h}$ and estimating the parameters which maximize the expected value of $L_c(\eta)$ with respect to $\mathbf{h}$.

The mixture of Gaussians model represents the $p(\mathbf{z}|\eta)$ probability density function that is expanded in a sum over the c clusters

$$p(\mathbf{z}|\eta) = \sum_{i=1}^{c} p(\mathbf{z},\eta_i) = \sum_{i=1}^{c} p(\mathbf{z}|\eta_i)\,p(\eta_i) \qquad (18)$$

where η is the set of the parameters $\eta = \{\eta_i | i = 1,\ldots,c\}$ of the model and $p(\eta_i)$ denotes the unconditioned cluster probabilities normalized to satisfy $\sum_{i=1}^{c} p(\eta_i) = 1$.

The $p(\mathbf{z}|\eta_i)$ distribution generated by the i-th cluster is represented by Gaussian like

$$p(\mathbf{z}|\eta_i) = \frac{1}{(2\pi)^{\frac{n+1}{2}}\sqrt{|\mathbf{F}_i|}} \exp\left(-\frac{1}{2}(\mathbf{z}-\mathbf{v}_i)^T (\mathbf{F}_i)^{-1}(\mathbf{z}-\mathbf{v}_i)\right) \qquad (19)$$

where η_i represents the parameters of the i-th cluster, $\eta_i = \{\mathbf{v}_i, \mathbf{F}_i | i = 1,\ldots,c\}$. As each data point $\mathbf{z}_k$ is generated by one and only one of the component Gaussians, the hidden random variable $\mathbf{h}_k$ for the EM algorithm it is the label of the component Gaussian to which $\mathbf{z}_k$ belongs.

Based on this assumption, the EM algorithm for maximum likelihood parameter estimation is the following. The algorithm starts with an initial guess $\eta^{(0)}$ of the parameters and repeatedly applies the following two steps to generate successively better parameter estimates:

Initialization: Initialize the means $\mathbf{v}_i$ to randomly picked data points from $\mathbf{Z}$ and the covariance matrices $\mathbf{F}_i$ to unit matrices. Set $p(\eta_i) = 1/c$ for all i.

Expectation (E) step: In the E-step we assume the current cluster parameter to be correct and evaluate the posterior probabilities that relate each data point in the conditional probability $p(\eta_i|\mathbf{z})$. These posterior probabilities can be interpreted as the probability that a particular piece of data was generated by a particular cluster. By using Bayes Theorem,

$$p(\eta_i|\mathbf{z}) = \frac{p(\mathbf{z}|\eta_i)\,p(\eta_i)}{p(\mathbf{z})} = \frac{p(\mathbf{z}|\eta_i)\,p(\eta_i)}{\sum_{i=1}^{c} p(\mathbf{z}|\eta_i)\,p(\eta_i)} \qquad (20)$$

This step is identical to

$$p(\eta_i|\mathbf{z}_k) = \frac{\frac{p(\eta_i)}{(2\pi)^{\frac{n+1}{2}}\sqrt{|\mathbf{F}_i|}} \exp\left(-\frac{1}{2}(\mathbf{z}_k-\mathbf{v}_i)^T \mathbf{F}_i^{-1}(\mathbf{z}_k-\mathbf{v}_i)\right)}{\sum_{i=1}^{c} \frac{p(\eta_i)}{(2\pi)^{\frac{n+1}{2}}\sqrt{|\mathbf{F}_i|}} \exp\left(-\frac{1}{2}(\mathbf{z}_k-\mathbf{v}_i)^T \mathbf{F}_i^{-1}(\mathbf{z}_k-\mathbf{v}_i)\right)} \qquad (21)$$

Maximization (M) step: In the M-step we assume the current data distribution to be correct and find the parameters of the clusters that maximize the likelihood of the data.

According to this, the unconditional probabilities are calculated as

$$p(\eta_i) = \frac{1}{N}\sum_{k=1}^{N} p(\eta_i|\mathbf{z}_k) \qquad (22)$$

Next we compute the remaining parameters of the cluster, the mean

$$\mathbf{v}_i = \int \mathbf{z}\,p(\mathbf{z}|\eta_i)\,d\mathbf{z} = \int \mathbf{z}\frac{p(\eta_i|\mathbf{z})}{p(\eta_i)}p(\mathbf{z})\,d\mathbf{z}$$
$$= \frac{1}{N\,p(\eta_i)}\sum_{k=1}^{N}\mathbf{z}_k\,p(\eta_i|\mathbf{z}_k) = \frac{\sum_{k=1}^{N}\mathbf{z}_k\,p(\eta_i|\mathbf{z}_k)}{\sum_{k=1}^{N}p(\eta_i|\mathbf{z}_k)} \qquad (23)$$

and in a similar way the cluster weighted covariance matrices

$$\mathbf{F}_i = \frac{\sum_{k=1}^{N}(\mathbf{z}_k-\mathbf{v}_i)(\mathbf{z}_k-\mathbf{v}_i)^T\,p(\eta_i|\mathbf{z}_k)}{\sum_{k=1}^{N}p(\eta_i|\mathbf{z}_k)} \qquad (24)$$

In the above it has been shown that the Gaussian Mixture Model models the $p(\mathbf{z}) = p(\mathbf{x}, y)$ joint density of the response variable y_k and the regressors $\mathbf{x}_k$ as a mixture of c multivariate $n+1$ dimensional Gaussian functions. The conditional density $p(y|\mathbf{x})$ is also a mixture of Gaussians model and the regression $E[y|\mathbf{x}]$ is

$$E[y|\mathbf{x}] = \int y\,p(y|\mathbf{x})\,dy = \frac{\int y\,p(y,\mathbf{x})\,dy}{p(\mathbf{x})} = \frac{\int y\,p(y,\mathbf{x})\,dy}{\int p(\mathbf{x},y)\,dy}$$
$$= \sum_{i=1}^{c}\frac{[[\mathbf{x}^T\ 1]\theta_i]\,p(\mathbf{x}|\eta_i)\,p(\eta_i)}{p(\mathbf{x})} = \sum_{i=1}^{c} p(\eta_i|\mathbf{x})[\mathbf{x}^T\ 1]\theta_i] \qquad (25)$$

where θ_i denotes the parameter vector of the local model and $p(\eta_i|\mathbf{x})$ denotes the probability that the i-th Gaussian component generated the regression vector $\mathbf{x}$:

$$p(\eta_i|\mathbf{x}) = \frac{\frac{p(\eta_i)}{(2\pi)^{n/2}\sqrt{|\mathbf{F}_i^{xx}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{v}_i^x)^T (\mathbf{F}_i^{xx})^{-1}(\mathbf{x}-\mathbf{v}_i^x)\right)}{\sum_{i=1}^{c}\frac{p(\eta_i)}{(2\pi)^{n/2}\sqrt{|\mathbf{F}_i^{xx}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{v}_i^x)^T (\mathbf{F}_i^{xx})^{-1}(\mathbf{x}-\mathbf{v}_i^x)\right)} \qquad (26)$$

where the $\mathbf{F}_i^{xx}$ is obtained by the partitioning of the $\mathbf{F}_i$ covariance matrix

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{F}_i^{xx} & \mathbf{F}_i^{xy} \\ \mathbf{F}_i^{yx} & \mathbf{F}_i^{yy} \end{bmatrix} \qquad (27)$$

The optimal θ_i parameter vector of the local models can be obtained as:

$$\begin{aligned} a_i &= \left(\mathbf{F}_i^{xx}\right)^{-1}\mathbf{F}_i^{xy} \\ b_i &= v_i^y - \mathbf{a}_i^T\mathbf{v}_i^x \end{aligned} \qquad (28)$$

or in a more compact form:

$$\theta_i = \left[\mathbf{F}_i^{yx}\left(\mathbf{F}_i^{xx}\right)^{-1},\ v_i^y - \mathbf{F}_i^{yx}\left(\mathbf{F}_i^{xx}\right)^{-1}\mathbf{v}_i^x\right] \qquad (29)$$

as $\mathbf{F}_i^{yx} = \left(\mathbf{F}_i^{xy}\right)^T$ and $\mathbf{F}_i^{xx} = \left(\mathbf{F}_i^{xx}\right)^T$.

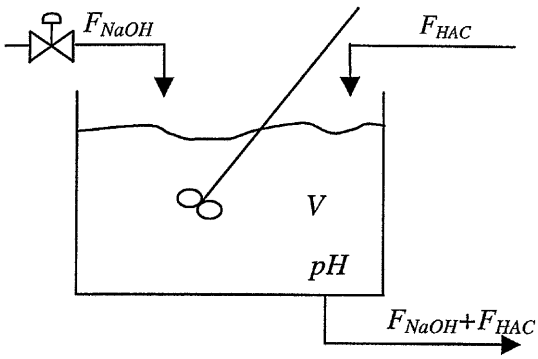

Fig.2 The considered continuous stirred tank reactor.

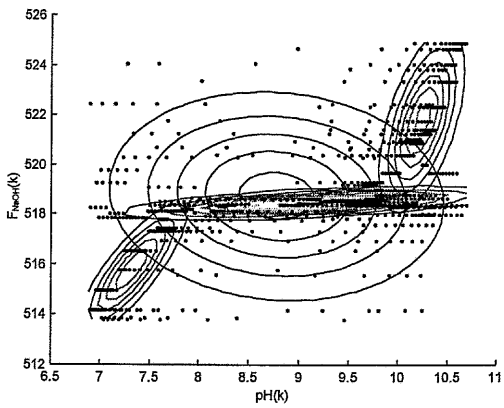

Fig.3 Orientation of the identified clusters.

This is identical to the weighted, also called local, parameter estimation approach that does not estimate all parameters simultaneously, because the parameters of the local models are estimated separately using a set of local estimation criteria

$$\min_{\theta_i} \frac{1}{N}(\mathbf{y}-\mathbf{X}_e\theta_i)^T \mathbf{\Phi}_i(\mathbf{y}-\mathbf{X}_e\theta_i) \qquad (30)$$

where $\mathbf{X}_e$ denotes the extended regression matrix obtained by adding a unitary column to $\mathbf{X}$, $\mathbf{X}_e=[\mathbf{X}\,1]$, and $\mathbf{\Phi}_i$ denote a diagonal matrix having membership degrees in its diagonal elements.

$$\mathbf{\Phi}_i=\begin{bmatrix}\mu_{i,1} & 0 & \cdots & 0\\ 0 & \mu_{i,2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \mu_{i,N}\end{bmatrix} \qquad (31)$$

The weighted least-squares estimate of the consequent rule parameters is given by

$$\theta_i=\left(\mathbf{X}_e^T\mathbf{\Phi}_i\mathbf{X}_e\right)^{-1}\mathbf{X}_e^T\mathbf{\Phi}_i\mathbf{y} \qquad (32)$$

Note that the resulted Gaussian Mixture of Local Models defined by *Eq.(25)* is identical to the operating regime based model given by *Eq.(3)* when the weighting function is chosen as $\phi_i(\mathbf{x})=p(\eta_i|\mathbf{x})$. Furthermore, the model is also identical to TS fuzzy models, when the membership functions are identified as it was shown in Section 2.

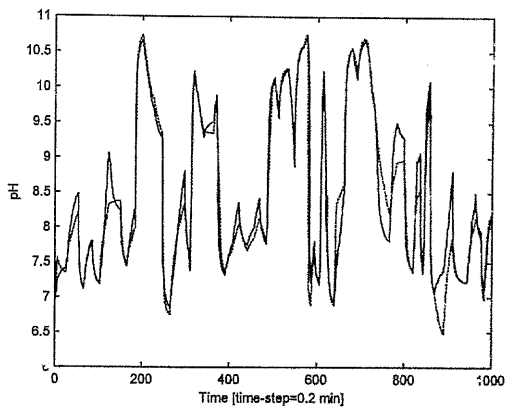

Fig.4 Free-run test of the model on the validation data.

Application to the identification of a pH process

The identification of the pH (the concentration of hydrogen ions) in a continuous stirred tank reactor (CSTR) is a well-known benchmark problem. The CSTR depicted in *Fig.2.* has two input streams: sodium hydroxide and acetic acid. The dynamic model for the pH in the tank is given in the Appendix. For collection of the data, a sampling interval of 0.2 min was used.

As the process can be modeled as a first order dynamic system [12], the fuzzy model consists of the following rules:

$$R_i:\ \text{If } [pH(k),F_{NaOH}(k)] \text{ is } A_i \quad \text{then } pH(k+1)=a_{i,1}\cdot pH(k)+a_{i,2}\cdot F_{NaOH}(k)+b_i \qquad (33)$$

By using transformed input variables:

$$R_i:\ \text{If } [t_{1,1}^i\cdot pH(k)+t_{1,2}^i\cdot F_{NaOH}(k)] \text{ is } A_{i,1} \text{ and } [t_{2,1}^i\cdot pH(k)+t_{2,2}^i\cdot F_{NaOH}(k)] \text{ is } A_{i,2} \quad \text{then } pH(k+1)=a_{i,1}\cdot pH(k)+a_{i,2}\cdot F_{NaOH}(k)+b_i \qquad (34)$$

Four local models were identified. The number of local models were determined by cross-validation. The operating region (the orientation of $\mathbf{F}_i^{xx}$ matrices) of these models are shown in *Fig.3.*

The identified model was tested with a validation data set. The model was used for one-step-ahead and simulation (infinite-step-ahead) prediction of the pH. The later experiments is depicted in *Fig.4.*

The results were compared with the performance of models obtained by using Fuzzy Model Identification Toolbox [9]. As *Table 1* shows, the proposed method shows superior performance over this advanced tool developed for identification of nonlinear dynamic systems.

Table 1 Comparison of Model Performances (mean squares of prediction errors).

Method	One-step	Simulation
FMID [9]	0.0241	0.2835
Proposed	0.009	0.0956

Conclusions

A new algorithm for the identification of nonlinear systems is proposed that is based on the Expectation Maximization identification of Gaussian Mixtures Model. A method to extract Takagi-Sugeno fuzzy models from Gaussian Mixtures Model is presented. The resulted fuzzy models are based on the transformed input-domain approach, which allows the effective partition of the input space and enables the interpretability of the model. The performance of the proposed modeling technique was demonstrated in the identification of the a pH process.

Acknowledgement

The financial support of the Hungarian Ministry of Culture and Education (FKFP-0023/2000, FKFP-0073/2001) and the Hungarian Science Foundation (TO23157) is greatly acknowledged. Janos Abonyi is grateful for the financial support of the Janos Bolyai Research Fellowship of the Hungarian Academy of Science. Tibor Chovan is supported by the Hans Pape foundation.

Model of the pH process

A dynamic model of the pH in a tank can be obtained by considering the material balances on $[Na^+]$ and the total acetate $[HAC+AC^-]$ and assuming that acid-base equilibrium and electroneutrality relationships hold [12].

$$F_{HAC}\cdot[HAC]_{in}-(F_{HAC}+F_{NaOH})\cdot[HAC+AC^-]=$$
$$=V\frac{d[HAC+AC^-]}{dt}$$

Sodium ion balance:

$$F_{NaOH}\cdot[NaOH]_{in}-(F_{HAC}+F_{NaOH})\cdot[Na^+]=V\frac{d[Na^+]}{dt}$$

HAC equilibrium:

$$\frac{[AC^-]\cdot[H^+]}{[HAC]}=K_a$$

Water equilibrium:

$$[H^+]\cdot[OH^-]=K_w$$

Electroneutrality:

$$[Na^+]+[H^+]=[OH^-]+[AC^-]$$

The pH can be calculated from the previous equations as

Table 2 Parameters used in the simulations.

Parameter	Description	Nominal Value
V	Volume of the tank	1000 [l]
F_{HAC}	Flow rate of acetic acid	81 [l/min]
F_{NaOH}	Flow rate of NaOH	515 [l/min]
$[NaOH]_{in}$	Inlet concentration of NaOH	0.05 [mol/l]
$[HAC]_{in}$	Inlet concentration of acetic acid	0.32 [mol/l]
$[Na^+]$	Initial concentration of sodium in the CSTR	0.0432 [mol/l]
$[HAC+AC^-]$	Initial concentration of acetate in the CSTR	0.0432 [mol/l]
K_a	Acid equilibrium constant	$1.753\cdot10^{-5}$
K_w	Water equilibrium constant	10^{-14}

$$[H^+]^3+[H^+]^2(K_a+[Na^+])+[H^+]([Na^+]K_a-$$
$$-[HAC+AC^-]K_a-K_w)-K_wK_a=0$$
$$pH=\log[H^+]$$

The parameters used in our simulations are taken from [12] and are given in *Table 2*.

REFERENCES

1. LEITCH R.: Comput. Control Eng. J., *1992*, July, 153-163
2. COTT B. J., DURHAM R. G. and SULLIVAN G. R.: Comput. Chem. Eng., *1989*, 13, 973-984
3. SJÖBERG J., ZHANG Q., LJUNG L., BENVENISTE A., DEYLON B., GLORENNEC P-Y., HJALMARSSON H. and JUDITSKY A.: Automatica, *1995*, 31, 1691-1724
4. MURRAY-SMITH R.: A Local Model Network Approach to Nonlinear Modelling, PhD Thesis, University of Strathclyde, Computer Science Department, 1994
5. MURRAY-SMITH R. and JOHANSEN T. A. (Ed): Multiple Model Approaches to Nonlinear Modeling and Control, Taylor & Francis, London, UK, 1997
6. BABUSKA R. and VERBRUGGEN H. B.: Fuzzy Set Methods for Local Modeling and Identification, in Multiple Model Approaches to Nonlinear Modeling and Control (Ed. MURRAY-SMITH R., JOHANSEN, T. A.), Taylor & Francis. London, UK, 75-100, 1997
7. TAKAGI T. and SUGENO M.: IEEE Trans. Syst. Man Cybern., *1985*, 15(1), 116-132
8. HUNT K.J., HAAS R. and MURRAY-SMITH R.: IEEE Trans. Neural Netw., *1996*, 7(3), 776-781
9. BABUSKA R.: Fuzzy Modeling for Control, Kluwer Academic Publishers, Boston, MA, 1998
10. KIM E., PARK M., KIM S. and PARK M.: IEEE Trans. Fuzzy Syst., *1998*, 6, 596-604
11. SCHONER B.: Probabilistic Characterization and Synthesis of Complex Driven Systems, Ph.D. Thesis, Massachusetts Institute of Technology, 2000
12. BHAT N. V. and MCAVOY T. J.: Comput. Chem. Eng., *1992*, 16, 271-281