Applied Mathematical Modelling 35 (2011) 4083–4090

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Fuzzy confidence interval for pH titration curve

Igor Škrjanc *

Faculty of Electrical Engineering, University of Ljubljana, Slovenia Tržaška 25, 1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 19 December 2008
Received in revised form 26 January 2011
Accepted 8 February 2011
Available online 16 February 2011

Keywords:
Fuzzy model
Interval fuzzy model
Confidence interval

ABSTRACT

In this paper we present a new method of confidence interval identification for Takagi–Sugeno fuzzy models in the case of the data with regionally changeable variance. The method combines a fuzzy identification methodology with some ideas from applied statistics. The idea is to find, on a finite set of measured data, the confidence interval defined by the lower and upper bounds. The confidence interval which defines the band that contains the measurement values with certain confidence. The method can be used when describing a family of uncertain nonlinear functions or when the systems with uncertain physical parameters are observed. In our example the proposed method is applied to model the pH-titration curve.

1. Introduction

The problem of the function approximation from a finite set of measured data using an optimality criterion has received a great deal of attention in the scientific community. A lot of different approaches appeared to approximate functions from data as: continuous piecewise linear (PWL) approach where it is possible to uniformly approximate any Lipschitz continuous function defined on a compact domain [1], neural-network approach [2], the fuzzy model approach, which in Takagi–Sugeno (TS) form, approximates the nonlinear system by smoothly interpolating affine local models [3]. Each local model contributes to the global model in a fuzzy subset of the space characterized by a membership function.

In this paper we look at the development of an interval function approximation methodology problem. This results in a lower and upper fuzzy model or a fuzzy functions. It is well known that the structure and shape of if-part fuzzy sets have a significantly effect on the fuzzy-model approximation of continuous functions [4]. Therefore, the proposed approach will exhibit an extra degree of flexibility in the domain partition as well as in the use of different membership functions compared to other function approximation technique.

The interval fuzzy model identification is a methodology for approximating the functions of a finite set of input and output measurements that can also be used to compress information in the case of a nonlinear function family approximation to obtain the confidence interval or band which contains the whole set or a certain amount, a certain enough big part of measurements. The interval fuzzy model approach is shown in [5], where the linear programming approach is used to obtain the fuzzy confidence interval, and in [6], where the confidence band is obtained using least-square optimization and a constant variance is assumed in the whole problem domain. A frequently used approach of modeling systems with uncertainties is called type-2 fuzzy modeling. This approach is described in Turksen [7], Liang and Mandel [8] and in Karnik and Mandel [9]. The problem of measurement uncertainties is captured by the uncertainties in membership functions. The type-2 fuzzy model gives the third dimension to the membership functions. This third dimension of membership function is called a possibility grade. The estimation of the fuzzy model parameters in this case requires a linear programming method. Our method, in comparison with type-2 fuzzy model, offers easier interpretation and understanding of the confidence interval.

* Tel.: +386 1 47 68 311; fax: +386 1 42 64 631.
E-mail address: igor.skrjanc@fe.uni-lj.si

0307-904X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.apm.2011.02.033

In our approach we are dealing with the problem where the variance of the data depends on the region of the input space and is therefore different in each subspace. This is of great importance in many technological areas, e.g., the modeling of nonlinear time-invariant systems with uncertain physical parameters. In our example the methodology of confidence interval modeling is used to define the nonlinear confidence band for pH-titration curve.

The paper is organized as follows: Section 2 provides the background to the fuzzy modeling; Section 3 describes the idea of fuzzy confidence interval model identification; Section 4 introduces the confidence interval of local linear model; and Section 5 presents an application to the pH-titration curve modeling.

2. Nonlinear model described in fuzzy form

A typical fuzzy model [3] is given in the form of rules

$$\mathbf{R}_j: \textbf{if } x_{p1} \textbf{ is } \mathbf{A}_{1,k_1} \textbf{ and } x_{p2} \textbf{ is } \mathbf{A}_{2,k_2} \textbf{ and} \ldots \textbf{and } x_{pq} \textbf{ is } \mathbf{A}_{q,k_q} \textbf{ then } y = \phi_j(\boldsymbol{x}) \quad j = 1,\ldots,m, \quad k_1 = 1,\ldots,f_1, \\ k_2 = 1,\ldots,f_2 \ldots \; k_q = 1,\ldots,f_q. \tag{1}$$

The q-element vector $\boldsymbol{x}_p^T = [x_{p1},\ldots,x_{pq}]$ denotes the input or variables in premise, and the variable y is the output of the model. With each variable in premise x_{pi} $(i = 1,\ldots,q)$, f_i fuzzy sets $(\mathbf{A}_{i,1},\ldots,\mathbf{A}_{i,f_i})$ are connected, and each fuzzy set $\mathbf{A}_{i,k_i}$ $(k_i = 1,\ldots,f_i)$ is associated with a real-valued function $\mu_{A_{i,k_i}}(x_{pi}) : \mathbb{R} \to [0,1]$, that produces the membership grade of the variable x_{pi} with respect to the fuzzy set $\mathbf{A}_{i,k_i}$. To make the list of fuzzy rules complete, all possible variations of fuzzy sets are given in Eq. (1), yielding the number of fuzzy rules $m = f_1 \times f_2 \times \cdots \times f_q$. The variables x_{pi} are not the only inputs of the fuzzy system. Implicitly, the n-element vector $\boldsymbol{x}^T = [x_1,\ldots,x_n]$ also represents an input to the system. It is usually referred to as the consequence vector. The functions $\phi_j(\cdot)$ can be arbitrary smooth functions in general, although linear or affine functions are normally used.

The system in Eq. (1) can be described in closed form if the intersection of the fuzzy sets is previously defined. The generalized form of the intersection is the so-called *triangular norm* (T-norm). In our case, the latter was chosen as an algebraic product providing the output of the fuzzy system

$$y = \frac{\sum_{k_1=1}^{f_1}\sum_{k_2=1}^{f_2}\cdots\sum_{k_q=1}^{f_q}\mu_{A_{1,k_1}}(x_{p1})\mu_{A_{2,k_2}}(x_{p2})\cdots\mu_{A_{q,k_q}}(x_{pq})\phi_j(\boldsymbol{x})}{\sum_{k_1=1}^{f_1}\sum_{k_2=1}^{f_2}\cdots\sum_{k_q=1}^{f_q}\mu_{A_{1,k_1}}(x_{p1})\mu_{A_{2,k_2}}(x_{p2})\cdots\mu_{A_{q,k_q}}(x_{pq})} \tag{2}$$

It should be noted that there is a slight abuse of notation in Eq. (2), since j is not explicitly defined as a running index. From Eq. (1) it is evident that each j corresponds to the specific variation of indexes k_i, $i = 1,\ldots,q$.

To simplify Eq. (2), a partition of unity is considered where the functions $\beta_j(\boldsymbol{x}_p)$, defined by

$$\beta_j(\boldsymbol{x}_p) = \frac{\mu_{A_{1,k_1}}(x_{p1})\mu_{A_{2,k_2}}(x_{p2})\cdots\mu_{A_{q,k_q}}(x_{pq})}{\sum_{k_1=1}^{f_1}\sum_{k_2=1}^{f_2}\cdots\sum_{k_q=1}^{f_q}\mu_{A_{1,k_1}}(x_{p1})\mu_{A_{2,k_2}}(x_{p2})\cdots\mu_{A_{q,k_q}}(x_{pq})}, \quad j = 1,\ldots,m, \tag{3}$$

give information about the fulfilment of the respective fuzzy rule in the normalized form. It is obvious that $\sum_{j=1}^{m}\beta_j(\mathbf{x}_p) = 1$ irrespective of $\boldsymbol{x}_p$ as long as the denominator of $\beta_j(\mathbf{x}_p)$ is not equal to zero (this can be easily prevented by stretching the membership functions over the whole potential area of $\boldsymbol{x}_p$). Combining Eqs. (2) and (3) and changing the summation over k_i to a summation over j we arrive at the following equation:

$$y = \sum_{j=1}^{m}\beta_j(\boldsymbol{x}_p)\phi_j(\boldsymbol{x}). \tag{4}$$

From Eq. (4) it is evident that the output of a fuzzy system is a function of the premise vector $\boldsymbol{x}_p$ (q-dimensional) and the consequence vector $\boldsymbol{x}$ (n-dimensional). The dimension of the input space may be lower than $(q + n)$ since it is very common to have the same variables present in vectors $\boldsymbol{x}_p$ and $\boldsymbol{x}$. Vector $\boldsymbol{z}$ (d-dimensional) is composed of the elements of $\boldsymbol{x}_p$, and those of $\boldsymbol{x}$ that are not present in $\boldsymbol{x}_p$.

Very often, the output value is defined as a linear combination of consequence states

$$\phi_j(\boldsymbol{x}) = [1\boldsymbol{x}^T]\boldsymbol{\theta}_j, \quad j = 1,\ldots,m, \quad \boldsymbol{\theta}_j^T = [\theta_{j0},\theta_{j1},\ldots,\theta_{jn}]^T, \tag{5}$$

by augmenting 1 to the vector $\boldsymbol{x}$.

Eq. (4) consists of m local linear models and can be written as

$$y = \sum_{j=1}^{m}\boldsymbol{\psi}_j^T\boldsymbol{\theta}_j, \quad j = 1,\ldots,m, \tag{6}$$

where $\boldsymbol{\psi}_j^T = \beta_j(\boldsymbol{x}_p)[1\boldsymbol{x}^T]$, $j = 1,\ldots,m$.

If the matrix of the coefficients for the whole set of rules is written as $\boldsymbol{\Theta}^T = [\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_m]$, and the fuzzy regression matrix

$$\boldsymbol{\psi}^T = \left[\boldsymbol{\psi}_1^T,\ldots,\boldsymbol{\psi}_m^T\right], \tag{7}$$

then Eq. (4) can be rewritten in the matrix form

$$y = \boldsymbol{\psi}^T \boldsymbol{\Theta}. \tag{8}$$

The fuzzy model in the form given in Eq. (8) is referred to as the affine Takagi–Sugeno model and can be used to approximate any arbitrary function that maps the compact set $\mathbf{C} \subset \mathbb{R}^d$ to $\mathbb{R}$ with any desired degree of accuracy [4,10,11]. The generality can be proven with the Stone-Weierstrass theorem which suggest that any continuous function can be approximated by a fuzzy basis function expansion [2].

3. Fuzzy confidence interval model identification

In this section we discuss our approach to define the fuzzy confidence interval. We assume a set of premise vectors $\boldsymbol{X}_p = \{\boldsymbol{x}_{p1}, \boldsymbol{x}_{p2}, \ldots, \boldsymbol{x}_{pN}\}$ and a set of antecedent (or consequence) vectors $\boldsymbol{x} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$, from which a set $\boldsymbol{Z} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N\}$ can be constructed that represents the input measurement data, collected from the compact set $\boldsymbol{S} \subset \mathbb{R}^d$. A set of corresponding outputs is also defined as $\boldsymbol{Y} = \{y_1, y_2, \ldots, y_N\}$. The measurements satisfy the nonlinear equation of the system

$$y_i = g(\boldsymbol{z}_i), \quad i = 1, \ldots, N. \tag{9}$$

According to the Stone–Weierstrass theorem, for any given real continuous function g on a compact set $\mathbf{U} \subset \mathbf{R}^d$ and arbitrary $\epsilon > 0$, there exist a fuzzy system f such that

$$\max_{\boldsymbol{z}_i \in \boldsymbol{Z}} |g(\boldsymbol{z}_i) - f(\boldsymbol{z}_i)| < \epsilon, \quad \forall i. \tag{10}$$

This implies the approximation of any given real continuous function with a fuzzy function from class $\mathscr{F}^d$ defined in Eq. (8). However, it has to be pointed out that lower values of ϵ imply higher values of m that satisfy Eq. (10). The answer lies in the proper arrangement of membership functions. This is a well-known problem in fuzzy systems. It can be overcome with a cluster analysis [12,13] or other approaches. The details will not be discussed in this paper.

Taking into account Eqs. (8)–(10) the set of data samples can be written as follows:

$$y_i = \boldsymbol{\psi}^T(\boldsymbol{z}_i)\boldsymbol{\Theta} + e_i, \quad i = 1, \ldots, N, \tag{11}$$

where stands e_i, $i = 1, \ldots, N$, for the noise of normal distribution, with zero mean value and variance which is regionally dependant and is written as $\boldsymbol{e} = \mathcal{N}(0, \sigma^2(\boldsymbol{x}_p))$ where $\boldsymbol{e} = [e_1, \ldots, e_N]^T$.

The error between the measured values and the fuzzy function outputs can be defined as

$$e_i = y_i - \boldsymbol{\psi}^T(\boldsymbol{z}_i)\boldsymbol{\Theta}, \quad i = 1, \ldots, N. \tag{12}$$

By having the membership functions defined, the structure of the model is known and only the fuzzy model parameters have to be defined. The parameters of $\boldsymbol{\Theta}$ are calculated separately for each local model. This means that we split Eq. (12) into m equations of the form

$$e_{i,j} = y_{i,j} - \boldsymbol{\psi}_{i,j}^T \boldsymbol{\theta}_j, \quad i = 1, \ldots, N, \quad j = 1, \ldots, m \tag{13}$$

where $e_{i,j} = \beta_{i,j} e_i, y_{i,j} = \beta_{i,j} y_i$, $\boldsymbol{\psi}_{i,j}^T = \beta_{i,j} \boldsymbol{\psi}^T(\boldsymbol{z}_i)$ and where $\beta_{i,j} = \beta_j(\boldsymbol{x}_{pi})$, $i = 1, \ldots, N$.

In matrix form the equation is written as follows

$$\boldsymbol{e}_j = \boldsymbol{y}_j - \boldsymbol{\Psi}_j^T \boldsymbol{\theta}_j, \quad j = 1, \ldots, m, \tag{14}$$

where $\boldsymbol{e}_j = [e_{1,j}, \ldots, e_{N,j}]^T$, $\boldsymbol{y}_j = [y_{1,j}, \ldots, y_{N,j}]^T$ and $\boldsymbol{\Psi}_j = [\boldsymbol{\psi}_{1,j}, \ldots, \boldsymbol{\psi}_{N,j}]^T$.

The vector of the estimated local model parameters is the minimizing argument which can be expressed as

$$\hat{\boldsymbol{\theta}}_j = \arg\min V_j(\boldsymbol{\theta}_j), \quad j = 1, \ldots, m, \tag{15}$$

where V_j reads as $V_j = \boldsymbol{e}_j^T \boldsymbol{e}_j$. The idea of an approximation can be interpreted as the most representative local fuzzy function to describe the local domain of outputs $\boldsymbol{y}_j$ as a function of inputs $\boldsymbol{z}$. The estimation of the local fuzzy model parameters is given by the *minimum least-square* optimization as follows:

$$\hat{\boldsymbol{\theta}}_j = \left(\boldsymbol{\Psi}_j \boldsymbol{\Psi}_j^T\right)^{-1} \boldsymbol{\Psi}_j \boldsymbol{y}_j \tag{16}$$

and the estimated output of the jth local fuzzy model is therefore written as $\hat{\boldsymbol{y}}_j = \boldsymbol{\psi}_j^T \hat{\boldsymbol{\theta}}_j$.

In particular case the estimated parameters of the fuzzy model, by taking into account Eq. (14), become $\hat{\boldsymbol{\theta}}_j = \boldsymbol{\theta}_j + \tilde{\boldsymbol{\theta}}_j$ where

$$\tilde{\boldsymbol{\theta}}_j = (\boldsymbol{\Psi}_j \boldsymbol{\Psi}_j^T)^{-1} \boldsymbol{\Psi}_j \boldsymbol{e}_j. \tag{17}$$

The expected bias of the local model parameters is then described as follows:

$$E\{\hat{\boldsymbol{\theta}}_j\} = \boldsymbol{\theta}_j + E\left\{\left(\boldsymbol{\Psi}_j \boldsymbol{\Psi}_j^T\right)^{-1} \boldsymbol{\Psi}_j \boldsymbol{e}_j\right\}, \quad j = 1, \ldots, m. \tag{18}$$

The right term in Eq. (18) equals zero, because of uncorrelated regression matrix $\boldsymbol{\Psi}_j$ and vector $\boldsymbol{e}_j$ and the zero mean value $E\{\boldsymbol{e}_j\} = 0$. This can be explained by taking into account the statistical property of the noise $E\{\boldsymbol{e}\} = 0$, what implies that also the noise of the jth local linear model, i.e., the weighted mean value $E\{\boldsymbol{e}_j\} = 0$ equals zero, when assuming enough big amount of measurements inside one fuzzy partitioning. This means that the estimation of the model parameters is unbiased. The weighted mean value is calculated as follows:

$$\bar{e}_j = E\{\boldsymbol{e}_j\} = \frac{1}{v}\sum_{i=1}^{N}\beta_{i,j}e_i, \quad v = \sum_{i=1}^{N}\beta_{i,j}, \quad j = 1,\ldots,m. \tag{19}$$

The expected covariance of the estimated parameters is calculated in the following way:

$$cov(\boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j) = E\left\{\tilde{\boldsymbol{\theta}}_j\tilde{\boldsymbol{\theta}}_j^T\right\}, \tag{20}$$

taking into account that $E\{\tilde{\boldsymbol{\theta}}_j\} = 0$. Using Eq. (17), the covariance matrix of model parameters is written as follows

$$cov(\boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j) = E\left\{\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\boldsymbol{\Psi}_j\boldsymbol{e}_j\boldsymbol{e}_j^T\boldsymbol{\Psi}_j^T\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\right\} = \left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\boldsymbol{\Psi}_j E\left\{\boldsymbol{e}_j\boldsymbol{e}_j^T\right\}\boldsymbol{\Psi}_j^T\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1} \tag{21}$$

and by taking into account the following notation $E\left\{\boldsymbol{e}_j\boldsymbol{e}_j^T\right\} = \hat{\sigma}_j^2\boldsymbol{I}$ the covariance matrix is written as follows:

$$cov(\boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j) = \hat{\sigma}_j^2\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}, \tag{22}$$

where $\hat{\sigma}_j^2$ stands for variance of $\boldsymbol{e}_j$

$$\hat{\sigma}_j^2 = \frac{1}{\mu - (n+1)}\sum_{i=1}^{N}\beta_{i,j}^2(e_i - \bar{e}_j)^2, \quad \mu = \sum_{i=1}^{N}\beta_{i,j}^2, \quad j = 1,\ldots,m \tag{23}$$

and $n + 1$ stands for the number of the estimated parameters of the fuzzy model.

The expected covariance of the residuals between the observed data and the model output is given as follows

$$cov(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j) = E\{(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j - E\{\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j\})(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j - E\{\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j\})^T\}. \tag{24}$$

Taking into account that $E\{\boldsymbol{e}_j\} = 0$ the expected value of the residue between measured output and estimated output becomes $E\{\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j\} = 0$.

The covariance matrix of the residuals, in Eq. (24), can be written as:

$$cov(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j) = E\{(\boldsymbol{e}_j - \boldsymbol{\Psi}_j^T\tilde{\boldsymbol{\theta}}_j)(\boldsymbol{e}_j - \boldsymbol{\Psi}_j^T\tilde{\boldsymbol{\theta}}_j)^T\} \tag{25}$$

and by taking into account Eq. (17), it is written as follows

$$cov(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j) = \hat{\sigma}_j^2\boldsymbol{I} - \hat{\sigma}_j^2\boldsymbol{\Psi}_j^T(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T)^{-1}\boldsymbol{\Psi}_j. \tag{26}$$

4. Confidence interval of local linear model

Let us define a confidence interval for a new set of data, given by the same function as in the case of identification, $g \in \mathscr{G}$. The corresponding set of measured output values $\boldsymbol{Y}^* = \{y_1^*,\ldots,y_M^*\}$ over the set of inputs $\boldsymbol{Z}^*$, i.e., $y_i^* = g(\boldsymbol{z}_i^*), g \in \mathscr{G}, \boldsymbol{z}_i^* \in \boldsymbol{S}, i = 1,\ldots,M$ is called the validation data set.

The idea of confidence interval fuzzy modeling is to find a lower fuzzy function $\underline{f}$ and an upper fuzzy function $\bar{f}$ satisfying

$$\underline{f}(\boldsymbol{z}_i^*) \leqslant g(\boldsymbol{z}_i^*) \leqslant \bar{f}(\boldsymbol{z}_i^*), \quad \forall \boldsymbol{z}_i^* \in \boldsymbol{S}. \tag{27}$$

In this sense, a function from class $\mathscr{G}$ can be found with a certain confidence in the band defined by the upper and the lower fuzzy function. The main request in defining the band is that it is as narrow as possible and should contain a certain percentage of data. The problem has been treated in the literature using the piecewise linear function approximation [1]. Our approach using the fuzzy function approximation can be viewed as a generalization of the piecewise linear approach and gives a better approximation, or at least a much narrower approximation band.

The measured output values of the jth local linear model are now defined as

$$\boldsymbol{y}_j^* = \boldsymbol{\Psi}_j^{*T}\boldsymbol{\theta}_j + \boldsymbol{e}_j^*, \tag{28}$$

where $\boldsymbol{\Psi}_j^{*T}$ stands for the regression matrix of the jth local linear model and $\boldsymbol{y}_j^* = [y_1^*,\ldots,y_M^*]^T$. The model output of the jth local linear model is in the case of validation data set defined as follows:

$$\hat{\boldsymbol{y}}_j^* = \boldsymbol{\Psi}_j^{*T}\hat{\boldsymbol{\theta}}_j. \tag{29}$$

To calculate the confidence interval in the case of validation set, we have to calculate the expected covariance of the residual between the model output and the new set of data in each local domain

$$cov\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right) = E\left\{\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^* - E\left\{\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right\}\right)\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^* - E\left\{\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right\}\right)^T\right\}. \tag{30}$$

Taking into account the same statistical properties of the noise for the data in validation data set $\left(E\{\mathbf{e}_j^*\} = 0\right)$ and for the identification set ($E\{\mathbf{e}_j\} = 0$), the expected value of the error between measured output and estimated output becomes $E\left\{\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right\} = 0$.

The covariance matrix in Eq. (30) can be rewritten as:

$$cov\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right) = E\left\{\left(\mathbf{e}_j^* - \boldsymbol{\Psi}_j^{*T}\tilde{\boldsymbol{\theta}}_j\right)\left(\mathbf{e}_j^* - \boldsymbol{\Psi}_j^{*T}\tilde{\boldsymbol{\theta}}_j\right)^T\right\} \tag{31}$$

and further on as follows:

$$cov\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right) = E\left\{\mathbf{e}_j^*\mathbf{e}_j^{*T}\right\} - E\left\{\boldsymbol{\Psi}_j^{*T}\tilde{\boldsymbol{\theta}}_j\mathbf{e}_j^{*T}\right\} - E\left\{\mathbf{e}_j^*\tilde{\boldsymbol{\theta}}_j^T\boldsymbol{\Psi}_j^*\right\} + E\left\{\boldsymbol{\Psi}_j^{*T}\tilde{\boldsymbol{\theta}}_j\tilde{\boldsymbol{\theta}}_j^T\boldsymbol{\Psi}_j^*\right\}. \tag{32}$$

Taking into account Eq. (17) and assuming that both noise signals have identical statistical properties, $E\left\{\mathbf{e}_j\mathbf{e}_j^T\right\} = E\left\{\mathbf{e}_j^*\mathbf{e}_j^{*T}\right\} = \hat{\sigma}_j^2$, and are uncorrelated $E\left\{\mathbf{e}_j\mathbf{e}_j^{*T}\right\} = E\left\{\mathbf{e}_j^*\mathbf{e}_j^T\right\} = 0$, Eq. (32) is written as follows:

$$cov\left(\mathbf{y}_j^* - \hat{\mathbf{y}}_j^*\right) = \hat{\sigma}_j^2\boldsymbol{I} + \hat{\sigma}_j^2\boldsymbol{\Psi}_j^{*T}\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\boldsymbol{\Psi}_j^* \tag{33}$$

The lower and the upper confidence interval of the local linear model are therefore defined as

$$\underline{f}_j(\mathbf{z}_i^*) = \boldsymbol{\psi}_{i,j}^{*T}\boldsymbol{\theta}_j - t_{\alpha,M-n}\hat{\sigma}_j\left(1 + \boldsymbol{\psi}_{i,j}^{*T}\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\boldsymbol{\psi}_{i,j}^*\right)^{\frac{1}{2}}, \quad i = 1,\ldots,M \tag{34}$$

and

$$\bar{f}_j(\mathbf{z}_i^*) = \boldsymbol{\psi}_{i,j}^{*T}\boldsymbol{\theta}_j + t_{\alpha,M-n}\hat{\sigma}_j\left(1 + \boldsymbol{\psi}_{i,j}^{*T}\left(\boldsymbol{\Psi}_j\boldsymbol{\Psi}_j^T\right)^{-1}\boldsymbol{\psi}_{i,j}^*\right)^{\frac{1}{2}}, \quad i = 1,\ldots,M, \tag{35}$$

where $t_{\alpha,M-n}$ stands for percentile of t-distribution for $100(1-2\alpha)$ percentage confidence interval with $M-n$ degrees of freedom.

The interval fuzzy modeling can be used efficiently in the case of fault detection where the data set of normal operating systems is modeled by interval fuzzy model to obtain the band of normal functioning. During operations this band is calculated on-line and it is checked if a measurement corresponds to the normal functioning band or not. If the measurement violates the tolerance band, one can assume that a malfunction might have occurred. The proposed model can also be used for the case of robust control design as described in [14].

5. Fuzzy confidence interval model of titration curve

The proposed approach is used to define the fuzzy confidence interval for titration curve of pH neutralization process. A mathematical model of a pH neutralization process was adopted from [15]. The example consists of a neutralization reaction between a strong acid (HA) and a strong base (BOH) in the presence of a buffer agent (BX). The neutralization takes place in a continuous stirred tank reactor (CSTR) with a constant volume V. It is a well-known fact that the pH processes are extremely difficult to deal with due to their highly nonlinear behavior with respect to different titration curves.

Fig. 1 shows a scheme of the continuous pH neutralization process. An acidic solution with a time-varying volumetric flow $q_A(t)$ of a composition $x_{1i}(t)$ is neutralized using an alkaline solution with volumetric flow $q_B(t)$ of known composition consisting base x_{2i} and buffer agent x_{3i}. Due to the high reaction rates of the acid–base neutralization, chemical equilibrium conditions are instantaneously achieved. Moreover, under the assumption that the acid, base and buffer are strong enough, total dissociation of the three compounds takes place. The process-dynamics model can be obtained by considering the electroneutrality condition (which is always preserved) and through mass balances of equivalent chemical species (known as chemical invariants). For this specific case, the dynamic behavior of the process can be described considering the state variables: $x_1 = [A^-]$; $x_2 = [B^+]$; $x_3 = [X^-]$.

Therefore, the mathematical model of the process can be written in the following way:

$$\begin{aligned}
\dot{x}_1 &= \frac{1}{\theta}\cdot(x_{1i} - x_1) - \frac{1}{V}\cdot x_1\cdot u,\\
\dot{x}_2 &= -\frac{1}{\theta}\cdot x_2 + \frac{1}{V}\cdot(x_{2i} - x_2)u,\\
\dot{x}_3 &= -\frac{1}{\theta}\cdot x_3 + \frac{1}{V}\cdot(x_{3i} - x_3)u,
\end{aligned} \tag{36}$$

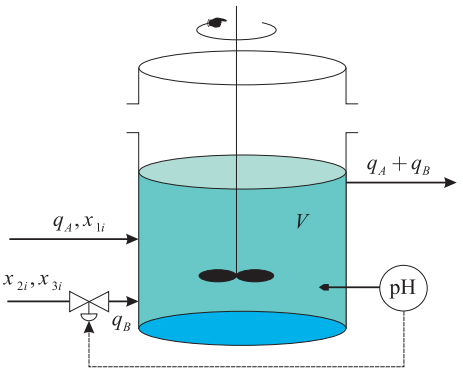

Fig. 1. pH-neutralization process.

$$g(x,\xi) = \xi + x_2 + x_3 - x_1 - \frac{K_w}{\xi} - \frac{x_3}{1 + \frac{K_x\xi}{K_w}} = 0, \tag{37}$$

where $\xi = 10^{-\mathrm{pH}}$, $\theta = V/q_A$, and $u = q_A/q_B$. K_w and K_x are the dissociation constants of the buffer and water, respectively. The parameters of the system represented by Eqs. (36) and (37) are $x_{2i} = 0.0020$ mol NaOH/L, $x_{3i} = 0.0025$ mol NaHCO3/L, $K_x = 10^{-7}$ mol/L, $K_w = 10^{-14}$ mol^2/L^2 and $V = 2.5$ L. Eq. (37) takes the standard form of the widely used implicit expression that connects pH, defined as pH $= -\log\xi$ and denoted as y in our example, with the states of the process, and it can also be rewritten to a third-order polynomial form:

$$g(x,\xi) = \xi^3 + (K_w/K_x + x_2 + x_3 - x_1)\xi^2 + (x_2 - x_1 + K_x)\xi - K_w^2/K_x = 0. \tag{38}$$

For the neutralization reactor, due to the instantaneous character of the acid–base reactions, where equilibrium conditions can be justified, it seems that the only dynamics involved is associated with the mixing phenomena. This implies that the concentrations of different chemical species that take part in the reaction vary from zero to a limit value. The goal of our approach is to model the titration curve of the pH process with a fuzzy confidence interval. The titration curve is defined with the relation between the measured input concentration $u_s(k)$ and corresponding pH value denoted as $y_s(k)$ in steady-state. In Fig. 2 the measured samples of the u_s and y_s are shown which can be used to identify the titration curve. The process input was assumed to be bounded by the interval $0 \leqslant u(k) \leqslant 1$. The input domain was divided into $m = 7$ subsets of gauss shape using fuzzy Gustafson–Kessel clustering algorithm as shown in Fig. 3. The obtained parameters of the fuzzy model are the following:

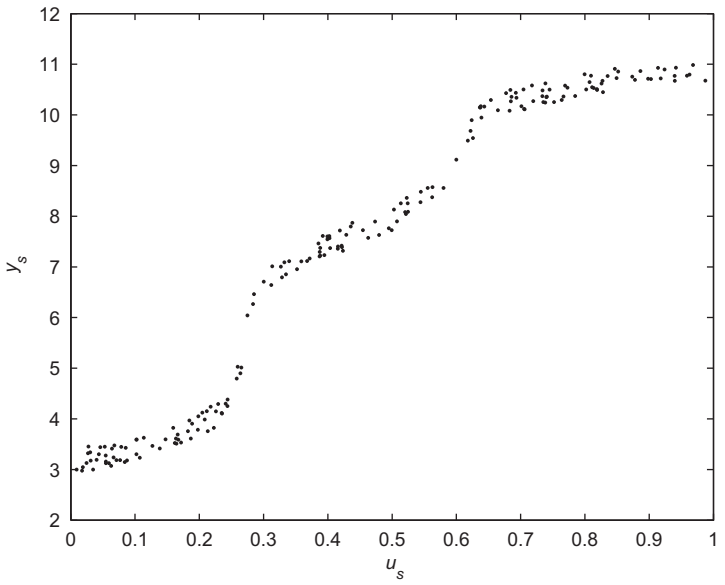

Fig. 2. The data samples of u_s and y_s.

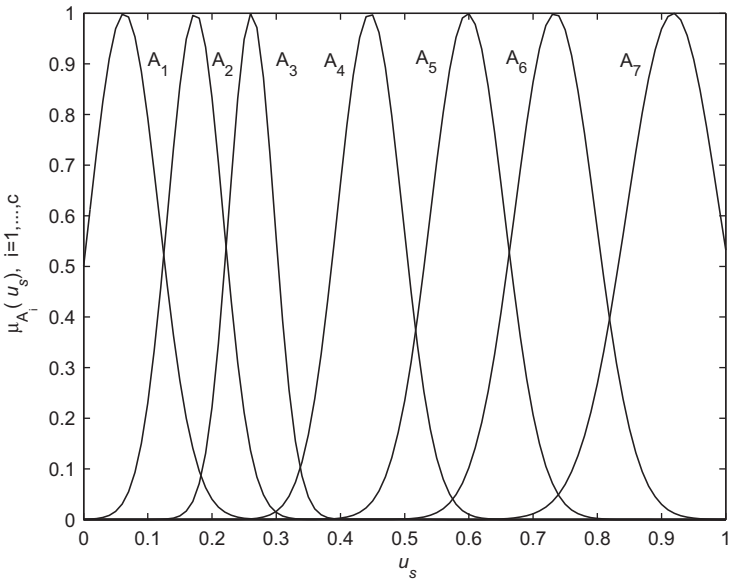

Fig. 3. The distribution of membership functions.

$$\widehat{\boldsymbol{\Theta}}^T = \begin{bmatrix} 3.0494 & 2.6517 & -1.6938 & 5.3845 & -0.8304 & 6.8539 & 9.5597 \\ 4.0357 & 6.3593 & 26.4716 & 5.1181 & 16.7560 & 4.9920 & 1.3354 \end{bmatrix}. \tag{39}$$

The corresponding standard deviations $\hat{\sigma}_j,\ j = 1, \ldots, m$ of the data are given in vector $\hat{\boldsymbol{\sigma}}$ as follows:

$$\hat{\boldsymbol{\sigma}}^T = [\,0.1454 \quad 0.1329 \quad 0.4502 \quad 0.1450 \quad 0.1835 \quad 0.1633 \quad 0.1495\,]. \tag{40}$$

The fuzzy confidence interval is given as a $100(1-2\alpha)\%$ confidence interval, where α stands for the confidence probability and is in our example equal to 0.025. This means that we have a 95% confidence interval and a 95% of all samples belong to the confidence interval. The percentile of t-distribution $t_{\alpha,M-n}$, with the degree of freedom $M - n = 190$, is equal to $t_{\alpha,M-n} = 1.96$. The lower and the upper bound of the confidence interval and the original data are given in Fig. 4. The percentile can be viewed as the parameter of the method, because it defines the percentage of confidence for the fuzzy confidence interval for a given data set. The percentile is constant for the whole data set.

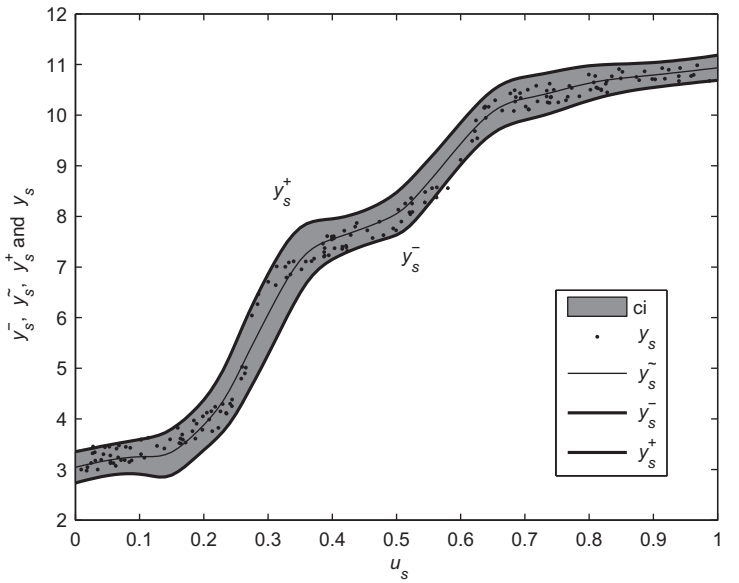

Fig. 4. The fuzzy confidence interval of pH titration curve.

6. Conclusion

A new method of fuzzy confidence interval identification has been proposed that is applicable when a finite set of measurement data is available. The idea is extended to the modeling of the optimal lower and upper bound functions that define the band that contains a certain amount, $(1-2\alpha)$, of the measurement samples. This results in the lower and upper fuzzy confidence bar what can be of great importance in the case of families of functions where the parameters of the observed system vary in certain intervals. Our approach can also be used in data mining to compress the information or in robust system identification. In the example the proposed method is applied to define a fuzzy confidence for highly nonlinear pH-titration curve.

References

[1] P. Julian, M. Jordan, A. Desages, Canonical piecewise linear approximation of smooh functions, IEEE Trans. Circ. Syst. 45 (1998) 567–571.
[2] C.-H. Lin, Siso nonlinear system identification using a fuzzy-neural hybrid system, Int. J. Neural Syst. 8 (3) (1997) 325–337.
[3] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modelling and control, IEEE Trans. Syst. Man Cybern. 15 (1985) 116–132.
[4] B. Kosko, Fuzzy systems as universal approximators, IEEE Trans. Comput. 43 (11) (1994) 1329–1333.
[5] I. Škrjanc, S. Blažič, O. Agamennoni, Identification of dynamical systems with a robust interval fuzzy model, Automatica (Oxf.) 41 (2005) 327–332.
[6] S. Jakubek, N. Keuthb, A local neuro-fuzzy network for high-dimensional models and optimization, Eng. Appl. Artificial Intell. 19 (6) (2006) 705–717.
[7] B. Turksen, Type I and type II fuzzy system modeling, Fuzzy Sets Syst. 106 (1999) 11–34.
[8] Q. Liang, J. Mendel, Interval type-2 fuzzy logic systems: Theory and design, IEEE Trans. Fuzzy Syst. 8 (5) (2000) 535–550.
[9] N. Karnik, J. Mendel, Operation on type-2 fuzzy sets, Fuzzy Sets Syst. 122 (2) (2001) 327–348.
[10] H.G.C. Ying, Necessary conditions for some typical fuzzy systems as universal approximators, Automatica 33 (1997) 333–1338.
[11] L.-X. Wang, J.M. Mendel, Fuzzy basis functions, universal approximation, and orthogonal least-squares learning, IEEE Trans. Neural Netw. 3 (5) (1992) 807–814.
[12] J.C. Bezdek, C. Coray, R. Gunderson, J. Watson, Detection and characterization of cluster substructure. I. Linear structure: Fuzzy *c*-lines, SIAM J. Appl. Math. 40 (2) (1981) 339–357.
[13] J.C. Bezdek, C. Coray, R. Gunderson, J. Watson, Detection and characterization of cluster substructure. II: Linear structure: Fuzzy *c*-varieties and convex combination thereof, SIAM J. Appl. Math. 40 (2) (1981) 358–372.
[14] J. Ackermann, Robust Control, Springer, Verlag, 1993.
[15] O. Galán, J.A. Romagnoli, A. Palazoglu, Robust h control of nonlinear plants based on multi-linear models: an application to a bench-scale ph neutralization reactor, Chem. Eng. Sci. 55 (2000) 4435–4450.