Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

On Identifying Envelop Type Nonlinear Output Error Takagi-Sugeno Fuzzy Models for Dynamic Systems with Uncertainties

Salman Zaidi, Andreas Kroll

Measurement and Control Department, Faculty of Mechanical Engineering, University of Kassel, Moenchebergstrasse 7, D-34125, Kassel, Germany, (e-mail: {salman.zaidi, andreas.kroll}@mrt.uni-kassel.de)

Abstract: In modeling of a stochastic nonlinear dynamic system from input-output data, it may be of interest to model uncertainty in the underlying system besides predicting a most likely or average response of the system. Due to stochasticity in the system behavior, the data obtained for identification can be considered as one realization of the underlying stochastic phenomenon. In order to effectively deal with the identification of such systems, it may be advantageous to repeat the identification experiment multiple times under similar conditions. The multiple input-output time series generated in this way thus contain information about stochastic variations within the system. This paper presents one of the possible approaches to effectively deal with identification in such scenario in the framework of Nonlinear Output Error (NOE) Takagi-Sugeno (TS) fuzzy models. Based on extended Chebyshev's inequality for finite samples, the lower and upper boundaries of the output time-series are obtained using (1-α) confidence interval (envelops of the response). The proposed identification algorithm provides a model for predicting the most likely value as well as the boundary models for predicting the envelops of the output signal. The experimental results for an electro-mechanical throttle shows the applicability and validity of the proposed approach.

Keywords: Nonlinear system identification, fuzzy modeling, stochastic modeling, uncertainty modeling, automobile industry.

1. INTRODUCTION

System identification in systems and control theory is concerned with building mathematical models of dynamic systems from measured input-output data, assuming inputs and model order to be known. The accuracy of the developed model is thus highly dependent upon the information content and quality of the data used for identification. The identification is carried out in two steps. In the first step, a model structure with unknown parameters is selected, and in the next step, the parameters of the model are determined by using parameter estimation techniques. To date, several methodologies have been successfully used for nonlinear dynamic system identification, such as artificial neural networks [Narendra and Parthasarathy (1990), Norgaard et al. (2003)], piecewise affine systems [Ferrari-Trecate et al. (2003), Daafouz et al. (2009)], and Takagi-Sugeno (TS) fuzzy systems [Takagi and Sugeno (1985)], to name a few. This paper focuses on TS fuzzy system identification for nonlinear stochastic dynamic systems.

Owing to its capability of approximating any continuous function with arbitrarily high precision [Ying (1998)], the TS fuzzy systems with affine consequents [Babuška (1998), Kroll (1996), Nelles (2001)] have been extensively used in fuzzy modeling and control. In TS fuzzy modeling, the input space is decomposed into a number of fuzzy subspaces. Each fuzzy subspace is characterized by a multi-variate membership function and describes the local behavior of the underlying system by an affine local model. The global nonlinear behavior of the system is obtained by smoothly interpolating these local models.

By repeating the identification experiment multiple times, a number of input-output time series can be generated. The problem then can be posed as to identify a TS fuzzy model that is capable of predicting not only an average or most likely response of the system (considering the variability in all time-series), but to also provide a measure of dispersion of output values around the predicted response. There exists some attempts in the literature regarding the identification of interval dynamic TS fuzzy systems, e.g., the interval fuzzy model (INFUMA) [Škrjanc et al. (2005)]. In INFUMA, the optimal lower and upper bound functions were approximated using linear programming. These bounds were obtained by first considering all possible extreme variations of parameters of the modeled function, which gave rise to a family of functions, and then selecting the minimum and maximum functions out of that family. Xu and Sun [Xu and Sun (2009)] have recently proposed an interval TS fuzzy model, in which the parameters of the consequent parts of the TS fuzzy rule become intervals by using interval regression analysis. They assumed the output signal to be in the interval form, but no explanation was given as how such output values can be obtained by identification. Another approach can be to use the higher type fuzzy sets (uncertain fuzzy sets) for solving the posed problem. As pointed out by Mendel [Mendel (2001)], ordinary Fuzzy Sets (FSs), referred to as Type-1 Fuzzy Sets (T1 FSs), are not fully capable of handling uncertainties present in real-world problems. Thus to improve the uncertainty handling capability of T1 FSs based Fuzzy Logic Systems (FLSs), the author suggested to use Type-2 Fuzzy Sets (T2 FSs) based FLS [Karnik et al. (1999)]. In T2 TS FLS, the uncertainties in systems are translated into uncertain antecedent membership functions and conclusion parameters. However, how these uncertainties can be determined by identification is still an open research question. A prediction interval based interval T2 FLS have been proposed recently by Khosravi et al. (2012). The prediction interval was created by considering uncertainty in the model's output due to two factors 1) an additive noise term, and 2) the model variance. However, the variability (the inherent stochasticity in systems) was not addressed in their approach. Similar to T2 FLSs, the Probabilistic Fuzzy Logic System (PFLS) was proposed by [Liu and Li (2005)]. The PFLS has the capability of modeling a system with stochastic uncertainties. PFLSs use probabilistic fuzzy sets as secondary fuzzy sets. Again, the criterion for determining the probability density of primary membership function values reflecting true uncertainties in data is not clear to date.

This research aims to provide a solution approach to the posed problem. The presented framework endows the classical Nonlinear Output Error (NOE) TS fuzzy identification technique with the ability to capture the stochastic variation in the realm of probability theory. Based on extended Chebyshev's inequality for finite samples, the upper and lower output time series based on $(1-\alpha)$ confidence level are constructed, which are termed as boundary time series in the sequel. These boundary time series are later used for estimating upper and lower boundary NOE TS fuzzy models. Whereas, the average response of the system, which is defined by sample-wise means of output time series, is obtained by averaging out the responses of these boundary models. The variability in reproducing input signals are usually negligible and so is the case in this research and thus neglected in the proposed identification approach. An electro-mechanical throttle is chosen as a case study for identification. The throttle has been studied extensively by Ren et al. (2012, 2013). It has shown stochastic behavior due to friction [Zaidi et al. (2012)]. The results obtained from the throttle case study has demonstrated that the proposed approach is able to estimate the mean and boundary time series with high accuracy as shown by the values of performances indices.

The rest of this paper is organized as follows. The problem statement is formulated in § 2. The proposed identification approach is discussed in detail in § 3. Experimental results on the throttle are recorded in § 4. Finally, the conclusion and outlook are given in § 5.

2. PROBLEM STATEMENT

The following notations have been used throughout in this paper to represent different types of variables.

- `deterministic scalar` : lower case, roman and normal font letter (e.g. v)
- `deterministic vector` : lower case, roman and bold font letter (e.g. $\mathbf{v}$)
- `deterministic matrix` : upper case, roman and normal font letter (e.g. V)
- `random scalar` : lower case, sans-serif and normal font letter (e.g. v)
- `random vector` : lower case, sans-serif and bold font letter (e.g. $\boldsymbol{\mathsf{v}}$)
- `random matrix` : upper case, sans-serif and normal font letter (e.g. V)

Consider a Single-Input-Single-Output (SISO) dynamic system. An extension to Multiple-Input-Single-Output (MISO) systems is straightforward. Furthermore, Multiple-Input-Multiple-Output (MIMO) systems can be decomposed to MISO systems. Assume that an input signal can be exactly reproduced in a given experiment (i.e. the input signal is considered to be deterministic). Application of the input signal to the system leads to the generation of the following data in the form of input-output pairs.

$$(\mathbf{u}^j, \mathbf{y}^j) := \{(u_k^j, y_k^j) | k = 1, \ldots, n\} \tag{1a}$$

$$(U, Y) := \{(\mathbf{u}^j, \mathbf{y}^j) | j = 1, \ldots, m\} \tag{1b}$$

$$T^j := (\mathbf{u}^j, \mathbf{y}^j) \tag{1c}$$

$$T := \{T^j | j = 1, \ldots, m\} \tag{1d}$$

where $u_k^j \in \mathbb{R}$ and $y_k^j \in \mathbb{R}$ denote the k-th input-output sample of the j-th experiment, n is the total number of samples in one experiment, m is the total number of experiments, $\mathbf{u}^j \in \mathbb{R}^{n\times 1}$ and $\mathbf{y}^j \in \mathbb{R}^{n\times 1}$ represent n dimensional column vectors containing the input and output time series for the j-th experiment, $U \in \mathbb{R}^{n\times m}$ and $Y \in \mathbb{R}^{n\times m}$ denote the input and output matrices, $T^j \in \mathbb{R}^{n\times 2}$ concatenates the input and output vectors for the j-th experiment, and $T \in \mathbb{R}^{n\times 2m}$ represents the entire time series for all experiments.

Since the input signal is deterministic, $u_k^p = u_k^q$ and $\mathbf{u}^p = \mathbf{u}^q$ for all k and $p, q \in \{1, \ldots, m\}$. Assuming variability in y_k^j, the underlying process can be viewed as m realizations of a stochastic process at the k-th time instant. Consequently, each y_k^j can be seen as the j-th realization of an independent and identically distributed (i.i.d.) random variable y_k with $\boldsymbol{\mathsf{y}} = [\mathsf{y}_1, \ldots, \mathsf{y}_k, \ldots, \mathsf{y}_n]$ defined on some probability space $(\Omega, \mathcal{B}, \Pr)$, where Ω is the sample space, $\mathcal{B}$ is the Borel sigma-algebra, and Pr is the probability measure. Each y_k is assumed to be stationary. The i.i.d. assumption about y_k is made by considering the fact that each experiment can be performed independently of each other. However, the components of the random vector $\boldsymbol{\mathsf{y}}$ are dependent upon each other through the tapped delay lines of inputs and outputs defined by the system dynamics.

The objective is firstly to estimate a model that describes the expected or most likely response of the considered stochastic dynamics. Secondly, the model should be able to provide an envelop for the expected spread of the output values. It is remarked here that the uncertainty could be modeled differently, e.g. as parametric instead as signal uncertainty.

3. IDENTIFICATION APPROACH

Assuming a discrete time SISO nonlinear system of the form:

$$y(k) = f(y(k-1), \dots, y(k-n_y), u(k-\tau-1), \dots, u(k-\tau-n_u)), \quad (2)$$

where n_y, τ and n_u represent the number of lagged output samples, dead time and number of lagged input samples, respectively. Observation typically assumes an additional additive i.i.d. Gaussian measurement noise term, $e(k) \sim N(0, \sigma_e^2)$, in (2). In case of deterministic input (u) and stochastic output (y), this system can be represented as a stochastic process of the form

$$\{\mathsf{y}_k | \mathsf{y}_{k-1}, \dots, \mathsf{y}_{k-n_y}, u(k-\tau-1), \dots, u(k-\tau-n_u)\}, \quad (3)$$

and the corresponding conditional densities

$$\Pr(\mathsf{y}_k | \mathsf{y}_{k-1}, \dots, \mathsf{y}_{k-ny}, u(k-\tau-1), \dots, u(k-\tau-n_u)), \quad (4)$$

which can be estimated from T for each time instant k. They can be visually inspected histograms. Additionally, normality tests and probability density estimation techniques can be applied for gaining more insight about the distribution.

In the proposed approach, in order to alleviate the problem and to formulate a computationally tractable model, only the upper and lower boundaries of the output time series are considered for identification which are obtained by using the extended Chebyshev's inequality [Chebyshev (1867)] for finite samples. Chebyshev's inequality guarantees that no more than $1/t^2$ of any arbitrary distribution's values can be farther than t standard deviations away from the mean, where $t \geq 1$ [Hazewinkel (2001)]. If μ_{y_k} (finite) and σ_{y_k} (finite and non-zero) denote respectively the mean and standard deviation of the random variable y_k, then the following inequality holds according to Chebyshev

$$\Pr(|\mathsf{y}_k - \mu_{y_k}| \geq t\ \sigma_{y_k}) \leq \frac{1}{t^2}\ . \quad (5)$$

A $(1-\alpha)$ Confidence Interval (CI) of y_k can be described by an interval of the form $[y_k^l, y_k^u]$, where $\Pr(y_k^l \leq \mathsf{y}_k \leq y_k^l) \xrightarrow{P} (1-\alpha)$ which indicates convergence in the probabilistic sense as the sample size goes to infinity. By equating $1/t^2$ to α and using the argument of the probability density function defined in (5), the $(1-\alpha)$ CI of y_k can be formulated as

$$y_k^l = \mu_{y_k} - \frac{1}{\sqrt{\alpha}}\sigma_{y_k}, \quad y_k^u = \mu_{y_k} + \frac{1}{\sqrt{\alpha}}\sigma_{y_k} \quad (6)$$

From (6), the mean of y_k can be obtained by averaging y_k^l and y_k^u, i.e.

$$y_k^m := \mu_{y_k} = \frac{y_k^l + y_k^u}{2} \quad (7)$$

However, μ_{y_k} and σ_{y_k} are the mean and standard deviation of the actual population and thus are unknown. Sample mean (m_{y_k}) and sample standard deviation (s_{y_k}) from the given sample of size m can be approximated as follows:

$$m_{y_k} = \frac{1}{m}\sum_{j=1}^{m} y_k^j, \quad k = 1, \dots, n \quad (8)$$

$$s_{y_k} = \frac{1}{m-1}\sum_{j=1}^{m}(y_k^j - m_{y_k})^2, \quad k = 1, \dots, n \quad (9)$$

Chebyshev's inequality has been extended to cases where the population mean and variance are not known but are instead replaced by their sample estimates [Saw et al. (1984)] and [Kabán (2012)]. According to Kabán, the resulting inequality is

$$\Pr(|\mathsf{y}_k - m_{y_k}| \geq t\ s_{y_k}) \leq \frac{1}{\sqrt{m(m+1)}}\left(\frac{m-1}{t^2} + 1\right). \quad (10)$$

The corresponding $(1-\alpha)$ CIs and mean of y_k defined in (6) and (7), respectively, can be approximated as

$$\hat{y}_k^l = m_{y_k} - s_{y_k}\sqrt{\frac{m-1}{\alpha\sqrt{m(m+1)}}} \quad (11a)$$

$$\hat{y}_k^u = m_{y_k} + s_{y_k}\sqrt{\frac{m-1}{\alpha\sqrt{m(m+1)}}} \quad (11b)$$

$$\hat{y}_k^m := m_{y_k} = \frac{\hat{y}_k^l + \hat{y}_k^u}{2} \quad (12)$$

As $m \to \infty$, the sample estimates defined by (8) and (9) approach the population parameters μ_k and σ_k, and accordingly, the approximated Kabán's CI and mean defined by (11) and (12) converge to the Chebychev's CI and mean defined by (6) and (7). It is noteworthy to mention here that Chebyshev's inequality provides conservative or loose bounds since it does not take into account the underlying distribution of data. Tighter or less conservative bounds can be obtained by considering either the actual distribution or by making some assumptions about the distribution.

Two time series can be constructed from the lower and upper bounds of the CI defined by (11). From hereon, these will be named lower and upper bound time series. The identification data used for estimating these bounding time series are obtained from (11) as follows

$$T^l := (\mathbf{u}, \hat{\mathbf{y}}^l) = \{(u_k, \hat{y}_k^l)\} \quad (13a)$$

$$T^u := (\mathbf{u}, \hat{\mathbf{y}}^u) = \{(u_k, \hat{y}_k^u)\} \quad (13b)$$

The expected values of condition densities defined in (4) are approximated from the sample means of the output time series given by (8) and (12)

$$T^m := (\mathbf{u}, \hat{\mathbf{y}}^m) = \{(u_k, \hat{y}_k^m)\} \quad (14)$$

By using T^l and T^u for identification, two separate NOE TS fuzzy models (called boundary models from hereon) are constructed. There is no need to construct a separate model from the data defined in (14), as from (12) it is evident that the mean response can be obtained by simply averaging the response of the boundary models.

The entire procedure consists of the following steps:

3.1 Experiment design and data generation

The first step for any identification is the design of experiments. The input signal used for identification should be persistently exciting to excite all the amplitude and frequency modes of interest [Ljung (1999)]. For capturing the stochasticity, the experiment are repeated multiple times to generate multiple time series for identification. A single time series in this case can be considered as one realization of the underlying stochastic system.

3.2 Synchronization of time series

The reason for doing this step is to make sure that the time series are synchronized before carrying out sample-wise operations. Asynchronism is not assumed due to the system behavior, but it can occur due to imperfection of the data generating and recording mechanism. Cross-correlation can be used for synchronizing two time series, which differ by an unknown shift along the time-axis. The cross-correlation between each pair of output time series $\{(\mathbf{y}^i, \mathbf{y}^j) | i \neq j | i, j \in \{1, 2, 3, \ldots, m\}\}$ should be examined. The cross-correlation between $\mathbf{y}^i$ and $\mathbf{y}^j$ is defined as

$$(\mathbf{y}^i \star \mathbf{y}^j)[k] \overset{def}{=} \sum_{r=1}^{n} \mathbf{y}^i[r]\mathbf{y}^j[k+r], \quad k = 1, \ldots, n \qquad (15)$$

from which the shift (k_{delay}) between them can be calculated as

$$k_{\text{delay}} = \arg\max_{k} \ ((\mathbf{y}^i \star \mathbf{y}^j)[k]). \qquad (16)$$

3.3 Identification of boundary models

The boundary models are estimated from the data defined by T^l and T^u in (13a) and (13b). The identification consists of the estimation of lower and upper NOE TS fuzzy models for T^l and T^j with set of parameters $\boldsymbol{\theta}^l_{\text{NOE}}$ and $\boldsymbol{\theta}^u_{\text{NOE}}$, respectively.

The following approach is used:

(1) Fuzzy clustering by Fuzzy c-means (FCM) [Bezdek (1981)] in the product space, using the Euclidean norm, to determine the partitioning into local models.
(2) Simultaneous global identification of local models using Ordinary Least Squares (OLS) by minimizing the sum of squared residuals (SSR) for one-step-ahead prediction (serial-parallel or Nonlinear Auto Regressive with eXogenous Input (NARX) model).
(3) Nonlinear optimization (trust-region-reflective algorithm) to minimize SSR for recursive model evaluation (parallel or Nonlinear Output Error (NOE) model). The parameters to be optimized include the cluster prototypes and parameters of local models.

A TS fuzzy model with multidimensional reference fuzzy sets [Kroll (1996)] and affine consequents having c rules is considered. The i-th fuzzy rule can be written as

$$R_i: \quad \text{IF } \mathbf{x} \text{ IS } \mathbf{v_i} \text{ THEN } \hat{y}_i = f_i(\mathbf{x}) \qquad (17)$$

with:
R_i: i-th fuzzy rule,
$\mathbf{x}$: the vector of r crisp inputs, $\mathbf{x} = [x_1, \ldots, x_r]^T \in \mathbb{R}^{r\times 1}$,
$\mathbf{v}_i$: i-th cluster prototype, $\mathbf{v}_i = [v_{1,i}, \ldots, v_{r,i}]^T \in \mathbb{R}^{r\times 1}$,
$\hat{y}_i$: crisp output of the i-th rule, $\hat{y}_i \in \mathbb{R}$,
f_i: affine conclusion function, $f_i(\mathbf{x}) = a_{0,i} + \sum_{j=1}^{r} a_{j,i}x_j$

In case of NARX nonlinear dynamic systems, $\mathbf{x}$ is the vector of lagged inputs and measured outputs, i.e., $\mathbf{x}(k) = [u(k-1), \ldots, u(k-n_u), y(k-1), \ldots, y(k-n_y)]^T$, with $r = n_u + n_y$ and $\hat{y}_i = \hat{y}_i(\mathbf{x}(k))$. The degree of fulfillment for the i-th rule is determined by evaluating the i-th membership function (MF)

$$\mu_i(\mathbf{x}(k)) = \left[\sum_{j=1}^{c}\left(\frac{||\mathbf{x}(k) - \mathbf{v}_i||}{||\mathbf{x}(k) - \mathbf{v}_j||}\right)^{\frac{2}{\nu-1}}\right]^{-1}, \quad \nu > 1 \qquad (18)$$

where ν is the fuzziness parameter. The final crisp output is given as the average of outputs of the c rules according to (17) weighted by their membership values

$$\hat{y}(\mathbf{x}(k)) = \sum_{i=1}^{c} \mu_i(\mathbf{x}(k))\hat{y}_i(\mathbf{x}(k)) \qquad (19)$$

since the MFs defined by (18) are orthogonal, i.e. $\sum_{i=1}^{c} \mu_i(\mathbf{x}(k)) = 1$. The algorithm consists of the identification of:

(1) Premise parameters, i.e. c cluster prototypes lumped into $\mathbf{v} \in \mathbb{R}^{cr\times 1}$, $\mathbf{v} := [\mathbf{v}_1^T, \ldots, \mathbf{v}_c^T]^T$.
(2) c sets of consequent parameters of the local affine models lumped into $\mathbf{a} \in \mathbb{R}^{c(r+1)\times 1}$, $\mathbf{a} := [\mathbf{a}_1^T, \ldots, \mathbf{a}_c^T]^T$, where $\mathbf{a}_i = [a_{0,i}, \ldots, a_{r,i}]^T \in \mathbb{R}^{(r+1)\times 1}$.

The FCM is used for the identification of premise structure. The cluster prototypes ($\mathbf{v}_{\text{NARX}}$) are obtained by minimizing the objective function

$$\mathbf{v}_{\text{NARX}} := \mathbf{v}^* = \arg\min_{\mathbf{v}} (J_{\text{FCM}}(X, P, \mathbf{v})) \qquad (20)$$

with

$$J_{\text{FCM}}(X, P, \mathbf{v}) = \sum_{k=1}^{n}\sum_{i=1}^{c} \mu(\mathbf{x}(k))^v ||\mathbf{x}(k) - \mathbf{v}_i||_2^2 \qquad (21)$$

where X is the input matrix, $X := [\mathbf{x}(1), \ldots, \mathbf{x}(n)]^T \in \mathbb{R}^{n\times r}$ and P is the partition matrix, $P := [\mu_i(\mathbf{x}(k))] \in \mathbb{R}^{c\times n}$. It is clear from the objective function that cluster prototypes are not adjusted to optimally estimate the input-output behavior of the system.

The consequent parameters ($\mathbf{a}_{\text{NARX}}$) are estimated globally by using OLS (NARX model). Denote $M_i \in \mathbb{R}^{n\times n}$, the diagonal matrix having membership grades $\mu_i(\mathbf{x}(k))$ as its k-th diagonal element. Define a matrix $X_e := [X, \mathbf{1}]$, where $\mathbf{1}$ is a unitary column vector in $\mathbb{R}^n$. Moreover, define $X' \in \mathbb{R}^{n\times kn}$ as $X' := [M_1X_e, \ldots, M_kX_e]$. Then $\mathbf{a}_{\text{NARX}}$ is calculated as

$$\mathbf{a}_{\text{NARX}} = [(X')^TX']^{-1}(X')^T\mathbf{y} \qquad (22)$$

The premise and consequent parameters can be lumped into $\boldsymbol{\theta}_{\text{NARX}} \in \mathbb{R}^{c(2r+1)\times 1}$ as follows

$$\boldsymbol{\theta}_{\text{NARX}} := [\mathbf{v}_{\text{NARX}}^T, \mathbf{a}_{\text{NARX}}^T]^T. \qquad (23)$$

Good evaluation properties of NOE or parallel models are important for simulation or for long-range predictions, e.g., in the context of model-based predictive control [Jelali and Kroll (2002)]. The matlab function `lsqnonlin` was used for determining optimal cluster prototypes and local model parameters for parallel mode evaluation. This function uses a trust-region-reflective algorithm based on the interior-reflective Newton method [Coleman and Li (1994, 1996)]. Denoting the lumped parameter vector for NOE model as $\boldsymbol{\theta}_{\text{NOE}} \in \mathbb{R}^{c(2r+1)\times 1}$, $\boldsymbol{\theta}_{\text{NOE}} := [\mathbf{v}_{\text{NOE}}^T, \mathbf{a}_{\text{NOE}}^T]^T$. It is obtained by the minimizing the SSR of NOE model with the input vector $\mathbf{x}(k) = [u(k-1), \ldots, u(k-n_u), \hat{y}(k-1), \ldots, \hat{y}(k-n_y)]^T$ as follows

$$\boldsymbol{\theta}_{\text{NOE}} := \boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \frac{1}{n}\sum_{k=1}^{n}(y(k) - \hat{y}_{\text{NOE}}(\boldsymbol{\theta}, k))^2 \qquad (24)$$

The starting value of $\boldsymbol{\theta}$ is chosen to be equal to $\boldsymbol{\theta}_{\text{NARX}}$. The identification is performed on the two boundary time series T^l and T^u to provide two boundary models with parameters $\boldsymbol{\theta}^l_{\text{NOE}}$ and $\boldsymbol{\theta}^u_{\text{NOE}}$, respectively.

4. EXPERIMENTAL RESULTS

The identification of an electro-mechanical throttle with friction, shown in Fig. 1, is presented as a case study.

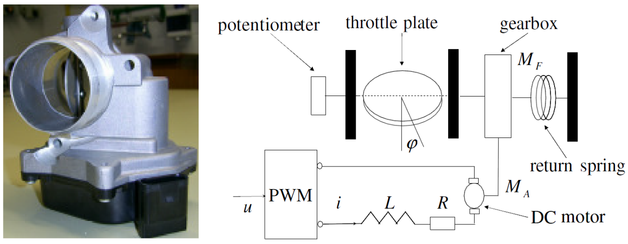

Fig. 1. Typical electro-mechanical throttle and its technology scheme [Ren et al. (2012)]

The collected open-loop data shows that the output signal has randomness due to the presence of uncertainties, mainly due to friction. A multisine signal, having the length of $n = 1000$ samples was chosen as the input signal. This phase optimized multisine signal was so parameterized that the throttle moves within its operation range [Ren et al. (2012, 2013). The sampling time was 1 milli second. The experiment was repeated 80 times to obtain $m = 80$ time series for identification. The input and output time series are shown in Fig. 2 and in Fig. 3, respectively. The observed spread (standard deviation) of output values have time varying characteristics which is shown in Fig. 4.

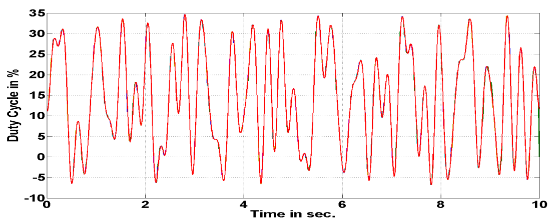

Fig. 2. Input time series (80 experiments)

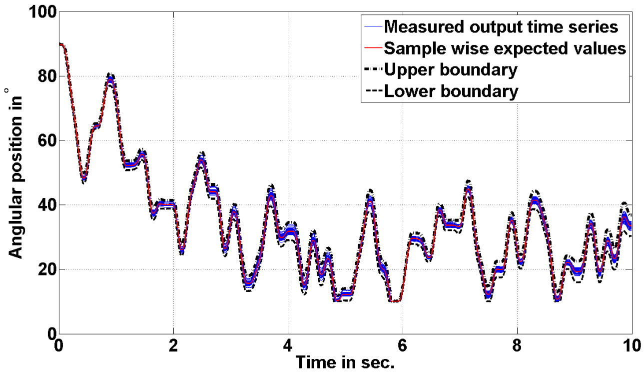

Fig. 3. Output time series with extended Chebyshev's Inequality based CIs and mean curves (80 experiments)

For identification of the throttle, $u(k-1)$, $y(k-1)$ and $y(k-2)$ were used as regressors. The FCM parameters were selected as $c = 4$ and $\nu = 1.2$. For having a parsimonious model, the value of c was selected based on the knee point of J_{NOE} (objective function containing SSR of NOE model), after which no considerable improvement in model performance was observed. The value of ν was selected based on the range suggested by Kroll (2011). The value of α was selected to be 5% (for 95% CI). The data were split into parts 1) identification data which is used for identification 2) test data which is solely used for checking model performance for unseen data. The first 90% of data (1-9 sec.) were used for identification and the remaining 10% (9-10 sec.) for testing the model. The results for identification and test data are shown in Fig. 5. The performance indices of Variance Accounted For (VAF), Root-Mean-Square Error (RMSE) and Maximum Absolute Error (MaxAE) were chosen for assessing model quality and were given by (25), (26) and (27), respectively. Their values are given in Table 1.

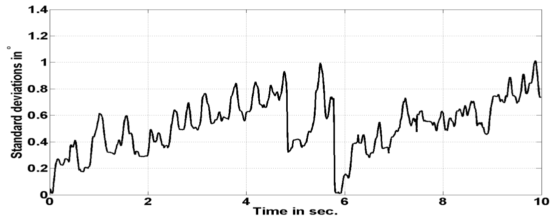

Fig. 4. Instant-wise standard deviation of output time series

$$\mathrm{VAF} = \left(1 - \frac{\mathrm{variance}\ (\mathbf{y} - \hat{\mathbf{y}})}{\mathrm{variance}\ (\mathbf{y})}\right) 100\% \tag{25}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{k=1}^{n} (y(k) - \hat{y}(k))}{n}} \tag{26}$$

$$\mathrm{MaxAE} = \max_{k}(|\hat{y}(k) - y(k)|) \tag{27}$$

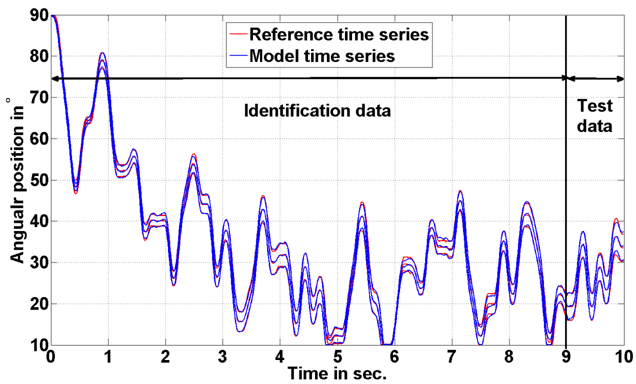

Fig. 5. Reference and model time series for identification and test data

5. CONCLUSION AND OUTLOOK

The experimental results obtained from the throttle case study show that the proposed model is able to provide the boundary models with good accuracy, where the upper and lower bound time series were obtained using $(1-\alpha)$ CI based on extended Chebyshev's inequality for finite sample sizes. The most likely response was inferred by averaging

Table 1. Modeling Performance of Upper Bound (UB), Lower Bound (LB) and Mean Time Series

		Identification data	Test data
LB time series	VAF in %	99.96	99.63
	MaxAE in °	1.11	1.23
	RMSE in °	0.34	0.52
UB time series	VAF in %	99.95	99.69
	MaxAE in °	2.59	1.12
	RMSE in °	0.37	0.40
Mean time series	VAF in %	99.98	99.91
	MaxAE in °	1.28	0.46
	RMSE in °	0.25	0.19

out the responses of these boundary models. The presented approach for building boundary models maybe considered as a first attempt to model the uncertainty observed in output time series in signal by only considering the CIs of output time series. Translating the observed uncertainties in signal space into parameter uncertainties in parameter space and investigating ways to incorporating them in modeling using Type-2 FLS based system will be the topic of future research.

ACKNOWLEDGEMENTS

The authors would like to thank the financial grant from German Academic Exchange Service (DAAD).

REFERENCES

Babuška, R. (1998). *Fuzzy Modeling for Control*. Kluwer Academic Publishers, Boston.

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press.

Chebyshev, P.L. (1867). Des valeurs moyennes. *J. Math. Pures Appl.*, 12(2), 177–184.

Coleman, T.F. and Li, Y. (1994). On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2), 189–224.

Coleman, T.F. and Li, Y. (1996). An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6, 418–445.

Daafouz, J., Benedetto, D.D., Blondel, V.D., Ferrari-Trecante, G., Hetel, L., Johansson, M., Joloski, A.L., Paoletti, S., Pola, G., Santis, E.D., and Vidal, R. (2009). Switched and piecewise affine systems. In *Handbook of Hybrid Systems Control: Theory, Tools, Applications*, 89–137. Cambridge University Press, London.

Ferrari-Trecate, G., Muselli, M., Liberati, D., and Morari, M. (2003). A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2), 205–217.

Hazewinkel, M. (2001). Chebyshev inequality in probability theory. In *Encyclopedia of Mathematics*. Springer.

Jelali, M. and Kroll, A. (2002). *Hydraulic Servo-systems: Modelling, Identification and Control*. Springer.

Kabán, A. (2012). Non-parametric detection of meaningless distances in high dimensional data. *Statistics and Computing*, 22(2), 375–385.

Karnik, N.N., Mendel, J.M., and Liang, Q.L. (1999). Type-2 fuzzy logic system. *IEEE Transactions on Fuzzy Systems*, 7(6), 643–658.

Khosravi, A., Nahavandi, S., Creighton, D., and Naghavizadeh, R. (2012). Prediction interval construction using interval type-2 fuzzy logic systems. In *Proceddings of the 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–7.

Kroll, A. (1996). Identification of functional fuzzy models using multidimensional reference fuzzy sets. *Fuzzy Sets and Systems*, 80(2), 149–158.

Kroll, A. (2011). On choosing the fuzziness parameter for identifying ts models with multidimensional membership functions. *Journal of Artificial Intelligence and Soft Computing Research*, 1(4), 283–300.

Liu, Z. and Li, H.X. (2005). A probabilistic fuzzy logic system for modeling and control. *IEEE Transactions on Fuzzy Systems*, 13(6), 848–859.

Ljung, L. (1999). *System Identification-Theory for the User*. Prentice-Hall.

Mendel, J.M. (2001). *Uncertain rule-based fuzzy logic systems, introduction and new directions*. Prentice Hall.

Narendra, K. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1), 4–27.

Nelles, O. (2001). *Nonlinear System Identification*. Springer, Berlin.

Norgaard, M., Ravn, O., Poulsen, N.K., and Hansen, L.K. (2003). *Neural networks for modelling and control of dynamic systems*. Springer, London.

Ren, Z., Kroll, A., Sofsky, M., and Laubenstein, F. (2012). On identification of piecewise-affine models for systems with friction and its application to electro-mechanical throttles. In *Proceedings of the 16th IFAC Symposium on System Identification*, 1395–1400.

Ren, Z., Kroll, A., Sofsky, M., and Laubenstein, F. (2013). Zur physikalischen und datengetriebenen Modellbildung von Systemen mit Reibung: Methoden und Anwendung auf Kfz-Drosselklappen. *Automatisierungstechnik*, 61(3), 155–171.

Saw, J.G., Yang, M.C.K., and Mo, T.C. (1984). Chebyshev inequality with estimated mean and variance. *The American Statistician*, 38(2), 130–132.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 15(1), 116–132.

Škrjanc, I., Blažič, S., and Agamennoni, O. (2005). Identification of dynamical systems with a robust interval fuzzy model. *Automatica*, 41(2), 327–332.

Xu, Z. and Sun, C. (2009). Interval t-s fuzzy model and its application to identification of nonlinear interval dynamic system based on interval data. In *Proceedings of the 48th IEEE Conference on Decision and Control/Chinese Control Conference*, 4144–4149.

Ying, H. (1998). General SISO Takagi-Sugeno fuzzy systems with linear rule consequent are universal approximators. *IEEE Transactions on Fuzzy Systems*, 6(4), 582–587.

Zaidi, S., Kroll, A., and Sommer, H.J. (2012). On description and identification of uncertainties in system modeling with fuzzy logic. In *Proceedings of the 22nd Workshop Computational Intelligence*, 179–200. KIT Scientific Computing, Dortmund.