NOE TS Fuzzy Modelling of Nonlinear Dynamic Systems with Uncertainties using Symbolic Interval-valued data

Salman Zaidi[a,*], Andreas Kroll[a]

[a] *Department of Measurement and Control, Institute for System Analytics and Control, Faculty of Mechanical Engineering, University of Kassel, Germany.*

Abstract

A novel and generalized approach to Nonlinear Output Error (NOE) modelling using Takagi-Sugeno (TS) fuzzy model for a class of nonlinear dynamic systems having variability in their outputs, owing to the inherent stochasticity, external disturbances and noise, is presented in this article. Assuming the identification method can be repeated offline a number of times under similar conditions, multiple input-output time series can be obtained from the underlying system. These time series are pre-processed using the techniques of statistics and probability theory to generate the envelopes of response (curves outlining the upper and lower extremes of response) at each time instant. Two types of envelopes are proposed in this research: the max-min envelopes and the envelopes based on the confidence intervals provided by extended Chebyshev's inequality. By incorporating interval data in fuzzy modelling and using the theory of symbolic interval-valued data, a TS fuzzy model with interval antecedent and consequent parameters is obtained. The proposed identification algorithm provides a model for predicting the expected response as well as envelopes. It is demonstrated on and validated for an academic simulation study and the real data obtained from an electro-mechanical throttle valve.

Keywords: Nonlinear system identification; Fuzzy modelling; Uncertainty

*Corresponding author

Email addresses: `salman.zaidi@mrt.uni-kassel.de` (Salman Zaidi), `andreas.kroll@mrt.uni-kassel.de` (Andreas Kroll)

modelling; Symbolic interval-valued clustering; Interval-valued linear regression

1. Introduction

1.1. Motivation and literature review

In systems and control theory, System Identification (SI) deals with building mathematical models of dynamic systems from measured input-output data. The accuracy of the developed model is thus highly dependent upon the information content and the quality of data used for identification. The identification is generally carried out in two steps. Firstly, a model structure is selected, and then it is followed by parameter estimation. To date, several methodologies have been successfully used for nonlinear dynamic system identification, such as artificial neural networks [1, 2], piecewise affine systems [3, 4], Takagi-Sugeno (TS) fuzzy systems [5, 6, 7], the Kolmogorov-Gabor polynomial and the parametric Volterra-Series models [8], to name a few. This paper focuses on the TS fuzzy modelling.

The fuzzy model obtained using classical techniques may not adequately represent the true system, on account of different uncertainties, owing to factors such as limited physical insight, unrealistic assumptions, limited quantity and quality of available data, and/or limited model complexity for control-oriented modelling. Consequently, it would be advantageous to have the information on the uncertainty in the model outputs besides a predicted or most likely response. Conventional fuzzy systems lack the ability to effectively handle uncertainties originated during modelling. Typically, ordinary or so-called type-1 fuzzy sets (T1 FSs) and a deterministic rule base are used in such systems, whose crisp membership values make them incapable of handling uncertainties. Mendel [9, 10] suggested to use Type-2 Fuzzy Logic Systems (T2 FLSs), which was originally proposed by Zadeh [11]. T2 FLSs use Type-2 Fuzzy Sets (T2 FSs), in order to permit an uncertainty handling capability of fuzzy systems. In [9, 12, 13, 14, 15, 16], it has been demonstrated that T2 FSs show superior performance (in terms of robustness and error reduction) in the presence of large uncertainties as compared to their counterpart T1 FSs. T2 FSs not only provide a crisp output at the end but also an uncertainty description as additional information, in terms of a type reduced set. The main drawback is the increased computational complexity, especially in type reduction, and the lack of systematic design approach to model effectively uncertainties in the secondary membership function of T2 FSs. Inspired by the T2 FLS, the Probabilistic Fuzzy Logic System (PFLS) [17, 18] was proposed, which is different from the TS FLS in that it uses the probability density function (PDF) in its secondary membership function. The PFLS has the capability of modelling a system with stochastic uncertainties because of probabilistic fuzzy sets as secondary fuzzy sets. The output of the PFLS is a random variable with a certain PDF which actually provides a measure of stochastic uncertainty associated with the output. However, the criterion for determining the probability density of primary membership function values reflecting true uncertainties in data is not clear to date. The prediction interval based T2 FLS was proposed in [19, 20]. The main motivation was to build a model whose response should be associated with the quality tag or confidence measure. In that method, both the validity and informativeness measures were incorporated into the non-continuous and non-differentiable objective function, which was later optimized by a meta-heuristic algorithm. Moreover, the model has no ability to incorporate the inherent stochasticity in the system dynamics.

The aforementioned fuzzy systems use crisp input-output data in which each data point is described by a crisp single-valued number. However, in many stochastic scenarios, the data cannot be pinned down to single-valued numbers due to the presence of uncertainties in the system. One of the possibilities to represent such form of data is to use intervals, possibly associated with weights or probability density function, and deal with them using the theory of Symbolic Data Analysis (SDA) [21, 22]. In the field of fuzzy modelling, only few articles discuss modelling using the interval data. One of the such earliest attempts include the INterval FUzzy MOdel (INFUMO) [23, 24]. The upper and lower bounds of response were obtained by first considering all possible extreme variations of parameters of the modelled function, which gave rise to a family of functions, and then selecting the minimum and maximum functions out of that family. These bounds were then approximated independently using the technique of linear programming. In this method, the estimated consequent parameters were in the form of intervals, whereas the antecedent parameters were crisp. This limitation was overcome by the model proposed by Xu and Sun [25, 26]. They proposed an interval TS fuzzy model in which they used the interval arithmetic to estimate interval consequent parameters. The method starts with the partitioning of the input space by clustering separately the centre and half range values of the antecedent variables. Later, these two memberships are combined using the T-norm operator. For estimating the consequent parameters, they used the interval regression analysis which resulted in interval consequent parameters. They developed a model for the case of crisp input and interval output. Only the case of one step ahead prediction or Nonlinear Auto Regressive with eXogenous inputs (NARX) was discussed in their approach. By simulating all possible combinations of upper and lower bounds of parameters, an interval response was obtained which actually provide the extreme values of the response. However, in system identification, a model is estimated only from input-output data, and thus these interval parameters are not known beforehand. In addition, the model is often required to be derived for the case of recursive evaluation or the NOE case for applications like simulation or model predictive control.

1.2. Scope of this research

This article is an extended and comprehensive version of the previous research of the authors [27, 28, 29] with a clear presentation of developed methods, some comparisons and guidelines, introduction of improved mathematical notations, demonstration of the developed method on an artificial dynamic system and the possible future research directions. A second order single-input-single-output (SISO) nonlinear system with artificially added uncertainty is chosen for demonstrating the developed method in an easy-to-understand way. The key reason for this example is that a known system permits to better access the performance and results of the developed methodology. The main contribution in this research is endowing the classical Nonlinear Output Error (NOE) TS fuzzy modelling with the stochastic theory and the symbolic data analysis for the first time, making it possible to deal with systems having stochastic variations in its output. In the previous research, we first introduced the idea of using the sample-wise max-min bounds as the measure of spread of output time series at each time instant [27]. Two independent TS fuzzy models were identified separately for estimating the max-min bounds of the output time series (so-called envelopes of the response) for the NOE case. The main drawbacks of using max-min bounds include sheer conservativeness, least robust statistics in the presence of outliers, and the fact that a minimum or maximum value of a distribution may not be defined. Subsequently, the approach was enhanced to include the probabilistic bounds using the probability theory in our previous work [28]. The approach assumes that the mathematical expectation is the best point estimate at any given instant of the output sample and uses the fact that the expectation is the arithmetic mean of the random variable coming from any probability distribution. The spread of the distribution is captured using Chebyshev's inequality. The advantage of using this inequality is that it makes no assumption about the distribution and thus assumptions like normality or symmetry etc. are not required. However, this comes with the drawback of extreme conservativeness of the obtained bounds. The drawback can be avoided by using the information of the actual distribution; for instance, if the distribution is known to be normal, its coverage factor can be used to obtain realistic bounds. Assuming the general case of any arbitrary distribution with mathematical expectation as the best point estimator, the method estimates the expected value of the response at each time instant as well as the upper and lower bounds based on the $(1-\alpha)100\%$ (α is usually called the "significance level") confidence intervals (CI) using extended Chebyshev's inequality ([30]) for the finite sample size [31]. Two independent fuzzy models were trained separately for estimating the upper and lower envelopes of response for the NOE case. The expected value is estimated by averaging the response of these two models as permitted by the structure of the Chebyshev's inequality.

The method had the drawback of using two independent TS fuzzy models for the CI based envelopes. In the latest work in [29], instead of using two separate models for envelopes, interval data was directly used in TS modelling to build a single TS model with interval antecedent and consequent parameters. In this article, the developed model is referred to the interval-data based type-1 TS fuzzy model. The modelling procedure uses Fuzzy C-Means clustering for the symbolic interval-valued data by optimizing an adequacy criterion based on suitable squared Euclidean distances between vectors of intervals [32]. As the result of this clustering, cluster prototypes are obtained in the form of intervals. The parameters of the local model are then estimated using the centre and range method suitable for the symbolic interval data ([33]). The model is first estimated for the NARX case and initial values of antecedent and consequent parameters are determined. These initial values are then passed to a nonlinear optimization algorithm to get the parameters of the NOE model.

For demonstrating the effectiveness of the proposed approach, two examples are presented in this article. One of them is an electro-mechanical throttle, which is defined as an experimental benchmark nonlinear stochastic dynamic system in [34]. The motivation behind choosing this test case is that the observed output time series have shown considerable variation when the same input signal applied in a series of experiments. This variability occurs because of different phenomena including friction, nonlinear spring characteristic, and manufacturing imperfections. As a result of experiment repetition, a family (band) of output time series is obtained due to the inherent stochasticity in the system. The second example is a nonlinear dynamic benchmark system described by the second order difference equation [35]. This system is made stochastic by randomizing the output signal. Five random variables of known distributions are added to the output signal at each time instant for achieving randomness. The rest of this article is structured as follows. First, the problem statement is formally and mathematically formulated in Sec. 2. The gist of this research lies in the proposed identification approach described in detail in Sec. 3. Next, the experimental and simulation results are recorded in Sec. 4. Finally, a brief conclusion of the current work and an outlook of the future work are given in Sec. 5.

2. Problem statement

The notations used in this article have been organized as follows. The normal lower case (e.g. y), bold lower case (e.g. $\boldsymbol{y}$) and normal upper case (e.g. Y) letters are used to denote scalars, vectors and matrices, respectively. Additionally, deterministic and random variables are distinguished by the roman italic (e.g. y) and san-serif (e.g. y) letters, respectively. Furthermore, an interval variable is represented by a tilde accent mark (e.g. $\tilde{y}$) in order to distinguish it from a crisp variable. For the ease of illustration, consider the case of a discrete time Single-Input-Single-Output (SISO) dynamic system having variability in its output given by (1). The extension of the SISO to the Multiple-Input-Single-Output (MISO) case is straightforward. Moreover, a Multiple-Input-Multiple-Output (MIMO) system can be thought of being composed of several MISO systems. Supported by the fact that very often the input signal has no uncertainty in practice, the input signal is assumed to be exactly reproducible and thus treated as a deterministic sequence. The system is mathematically described as follows

$$\mathsf{y}_k = \mathcal{F}(\mathsf{x}_k) + \zeta_k, \quad k = 1, \cdots, N \tag{1}$$

where,

- k and N denote respectively the time index and the number of observations
- y_k is a scalar stochastic dependent signal (output/regressent)
- x_k is a vector stochastic independent quantity (regressor) consisting of lagged values of input (deterministic) and output (stochastic) signals, i.e., $\mathsf{x}_k = [\mathsf{y}_{k-1}, \ldots, \mathsf{y}_{k-n_y}, u_{k-\tau-1}, \ldots, u_{k-\tau-n_y}]^\intercal$, where n_y, τ and n_u rep-

resent the number of lagged output samples, dead time and number of lagged input samples, respectively.

– $\mathcal{F}(\cdot)$ is a stochastic function

– ζ_k is additive noise with zero mean and finite variance

$\mathcal{F}(\mathsf{x}_k)$ and ζ_k are assumed to be independent of each other. The underlying process can be viewed as M realizations of a stochastic process. $\mathbf{y} = [\mathsf{y_1}, \ldots, \mathsf{y_k}, \ldots, \mathsf{y_N}]$ is defined on some probability space $(\Omega, \mathcal{B}, \mathrm{Pr})$, where Ω is the sample space, $\mathcal{B}$ is the Borel sigma-algebra, and Pr is a probability measure. Each y_k is assumed to be stationary. The i.i.d. (independent and identically distributed) 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 $\mathbf{y}$ are dependent upon each other through the tapped delay lines of inputs and outputs defined by the system dynamics. Conditional densities $p(\mathsf{y}_k|\mathsf{y}_{k-1}, \ldots, \mathsf{y}_{k-n_y}, u(k-\tau-1), \ldots, u(k-\tau-n_u))$ can be estimated for each instant k. To check the nature of the distribution followed by y_k, the histogram of output values at each time instant can be plotted for visual inspection. As a particular case of normality, the distribution can be tested by statistical techniques, such as Shapiro-Wilk (SW), Kolmogorov-Smirnov (KS), Anderson-Darling (AD) or Lillifors (LF) test, see [36] and the references therein for details and comparisons. This information can then later be used to choose a realistic coverage factor for determining the envelopes of a response.

The expected value ($E(\cdot)$) of y_k in Eq. (1) is given by

$$y(k) = E(\mathsf{y}_k) = E(\mathcal{F}(\mathsf{x}_k)) := f(k) \tag{2}$$

The variance ($\sigma_{(\cdot)}$) of y_k in Eq. (1) can be calculated by the variance sum law

$$\sigma^2_{\mathsf{y}_k} = \sigma^2_{\mathcal{F}(\mathsf{x}_k)} + \sigma^2_{\zeta_k} \tag{3}$$

where σ_{ζ_k} can be estimated experimentally by repeated measurements while holding the inputs constant. For a stochastic system with considerably low measurement noise as compared to its inherent stochasticity, the variance due to stochasticity is significantly greater than the variance of noise, i.e. $\sigma_{\zeta_k} \ll \sigma_{\mathcal{F}(\mathsf{x}_k)}$, which leads to $\sigma_{\mathsf{y}_k} \approxeq \sigma_{\mathcal{F}(\mathsf{x}_k)}$.

The task is to estimate a model that describes the expected response of the considered stochastic dynamics ($f(k)$) in (2)) and provides the envelopes (worst-case and probability based) of response.

3. Identification approach

The proposed identification approach is divided into following steps

3.1. Design of Experiment

The Design of Experiment (DOE) is the first step for estimating a data-driven model. Therefore, the quality of the developed model heavily depends upon the data used for identification. The input signal should be persistently exciting to be able to excite all the amplitudes and frequency modes of interest [37]. For capturing the stochasticity, the experiment is 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 dynamic system.

3.2. Determination of Output Envelopes

The data used for identification can be lumped together in a matrix $Z = [\mathbf{u}, Y] \in \mathbb{R}^{N\times(M+1)}$; where $\mathbf{u} \in \mathbb{R}^N$ and $Y \in \mathbb{R}^{N\times M}$ are, respectively, the input vector and the output matrix; N and M denote the length of one experiment and the total number of experiments, respectively. The same input signal is used in M experiments and thus it is represented by a single vector of values, i.e. $\mathbf{u}$. The output matrix is given by (4)

$$Y = \begin{bmatrix} y_1^1 & y_1^2 & \dots & y_1^{M-1} & y_1^M \\ y_2^1 & y_2^2 & \dots & y_2^{M-1} & y_2^M \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_N^1 & y_N^2 & \dots & y_N^{M-1} & y_N^M \end{bmatrix} = \begin{bmatrix} \mathsf{y}_1 \\ \mathsf{y}_2 \\ \vdots \\ \mathsf{y}_N \end{bmatrix}. \tag{4}$$

In this matrix, each column represents the output of an independent experiment, whereas each row represents the output of each experiment at a given time instant. Given a sample k, the corresponding row of the output matrix can be seen as M realizations of the stochastic output variable y_k at that sampling instant following a certain probability distribution function conditioned by past lagged output and input values, i.e.

$$\Pr(\mathsf{y}_k | \mathsf{y}_{k-1}, \dots, \mathsf{y}_{k-n_y}, u_{k-\tau-1}, \dots, u_{k-\tau-n_u}). \tag{5}$$

For tractability, only the upper and lower output boundaries or envelopes are considered for identification. Two types of envelopes are considered in the following: the max-min and confidence interval based envelopes.

For max-min envelopes, the maximum and minimum values of y_k are considered to be the observed maximum $y_k^{\max}$ and the observed minimum $y_k^{\min}$ as follows:

$$y_k^{\max} = \max_j \left(y_k^j \right), \tag{6}$$ $$y_k^{\min} = \min_j \left(y_k^j \right). \tag{7}$$

The approach of max-min envelopes has, however, some serious limitations. First, it does not provide the expected response, as the average of these bounds doesn't represent the expected response but rather the mid-value of the max-min envelopes. Secondly, the sample maximum and minimum statistics are considered to be the least robust statistics due to their sensitivity to outliers in the data. Thirdly, the true extreme values of the distribution where the data is originating from may not even exist (e.g. it may be extended from $-\infty$ to $+\infty$). Lastly, the max-min values may depend heavily on the given sample especially when the sample size is small.

An alternative and more reliable approach is to use probability theory and calculate the envelopes of the responses based on the upper and lower confidence bounds of y_k around its mean. In case of a distribution of known type, the confidence bounds around the mean can be calculated by adding and subtracting some coverage factor k_{cov} times the standard deviation to the mean value. For instance, the coverage factor in case of a normal distribution for 95% confidence level is $k_{\text{cov}} = 1.96$. However, in many practical situations, the distribution may deviate from the normal distribution significantly and thus using the coverage factor for the normal distribution can be misleading. In general the distribution of y_k can have any shape. Chebyshev's inequality provides the coverage factor for the most general case, i.e. a distribution of any kind (whether unimodal or multi-modal, symmetric or asymmetric etc.). The main drawback is the conservativeness of the obtained bounds because of not making any assumption and not utilizing the data distribution. On the other hand, the advantage is that the obtained bound guarantee that no more than $1/k_{\text{cov}}^2$ of values of a random variable y_k of any arbitrary distribution can be farther than k_{cov} standard deviations (σ_k) from the mean (μ_k), where $k_{\text{cov}} \geq 1$ [30]. Mathematically,

$$\Pr(|\mathsf{y}_k - \mu_k| \geq k_{\text{cov}}\ \sigma_k) \leq \frac{1}{k_{\text{cov}}^2}\ , \quad k = 1, \ldots N. \tag{8}$$

Chebyshev's inequality assumes that the true value of the population mean and standard deviation are known, whereas in the practical world, where the experimental data comes from a complex stochastic system, this is often not the case. Providing the sample size is large enough, the sample mean and standard deviation can be reasonably estimated by the population mean and standard deviation. Kaban ([31]) provided the approximated Chebyshev's inequality in terms of sample parameters as follows:

$$\Pr(|\mathsf{y}_k - m_{y_k}| \geq k_{\text{cov}}\ s_{y_k}) \leq \frac{1}{\sqrt{M(M+1)}} \left(\frac{M-1}{k_{\text{cov}}^2} + 1\right), \tag{9}$$

where m_{y_k} and s_{y_k} are the sample mean and standard deviation at the k-th time instant, respectively, given as follows:

$$m_{y_k} = \frac{1}{M}\sum_{j=1}^{M} y_k^j, \tag{10}$$

$$s_{y_k} = \sqrt{\frac{1}{M-1}\sum_{j=1}^{M}(y_k^j - m_{y_k})^2}. \tag{11}$$

From (9), $(1-\alpha)100\%$ envelopes based on the confidence bounds of y_k are determined in the form of the interval $[\underline{y}_k, \bar{y}_k]$, where the bounds are approximated as

$$\underline{y}_k = m_{y_k} - s_{y_k}\sqrt{\frac{M-1}{\alpha\sqrt{M(M+1)}}}, \tag{12}$$

$$\bar{y}_k = m_{y_k} + s_{y_k}\sqrt{\frac{M-1}{\alpha\sqrt{M(M+1)}}}. \tag{13}$$

By averaging the bounds in (13), the expected response is obtained:

$$y_k^{\exp} := m_{y_k} = 0.5(\underline{y}_k + \bar{y}_k) \tag{14}$$

3.3. Data generation for Identification

Let $\tilde{y}_k = [\underline{y}_k, \bar{y}_k] \in \Im = \{[a,b] : a, b \in \mathbb{R}, a \leq b\}$ $(k = 1, \ldots, N)$ represent the interval-valued output at each sampling instant k. The input-output data pairs are collected in $\tilde{Z} = \{(u_k, \tilde{y}_k)\}_{k=1,\ldots,N}$ which will be used for the identification of the TS fuzzy model using symbolic interval-valued data. The input-output data pairs for the expected values are stored in $Z^{\exp} = \{(u_k, y_k^{\exp})\}_{k=1,\ldots,N}$, which will be used for evaluating the modelling performance for the expected response. In order to avoid the phenomenon of overfitting of the developed model, the data is split into two parts, namely the identification and the validation dataset.

3.4. Model structure

The i-th rule of a TS fuzzy model with c rules having antecedents defined by multidimensional reference fuzzy sets [6] and consequents by affine functions for the SISO and MISO cases is given by:

$$R_i: \quad \text{IF } \mathbf{z} \text{ IS } \mathbf{v}_i \text{ THEN } \hat{y}_i(\mathbf{x}) = \mathbf{a}_i^{\intercal} \cdot [1 \ \mathbf{x}^{\intercal}]^{\intercal}, \tag{15}$$

where:

R_i: i-th fuzzy rule,

$\mathbf{z}$: antecedent variable, $\mathbf{z} \in \mathbb{R}^{r_{\text{ant}}}$,

$\mathbf{v}_i$: i-th cluster prototype, $\mathbf{v}_i \in \mathbb{R}^{r_{\text{ant}}}$,

$\hat{y}_i(\mathbf{x})$: crisp output of the i-th rule, $\hat{y}_i(\mathbf{x}) \in \mathbb{R}$,

$\mathbf{a}_i$: consequent parameters, $\mathbf{a}_i \in \mathbb{R}^{r_{con}+1}$,

$\mathbf{x}$: consequent variable, $\mathbf{x} \in \mathbb{R}^{r_{\text{con}}}$.

In this classical type-1 TS fuzzy model, the cluster prototypes $\mathbf{v}_i$ and the consequent parameters $\mathbf{a}_i$ are defined as crisp numbers. The final crisp output of the model is given by the weighted average of all $\hat{y}_i$'s:

$$\hat{y}(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^{c} \mu_i(\mathbf{z}) \cdot \hat{y}_i(\mathbf{x}), \tag{16}$$

where $\mu_i(\mathbf{z}) \in [0, 1]$ is the membership of the scheduling variable $\mathbf{z}$ defined by the orthogonal membership function of Fuzzy c-Means clustering (FCM) [38]:

$$\mu_i(\mathbf{z}, \mathbf{v}_i|_{i=1,\dots,c}) = \left[\sum_{j=1}^{c} \left(\frac{||\mathbf{z} - \mathbf{v}_i||}{||\mathbf{z} - \mathbf{v}_j||} \right)^{\frac{2}{\nu-1}} \right]^{-1}, \ \forall\, i, \tag{17}$$

where ν is the fuzziness parameter, $\nu > 1$. Since the membership functions defined by (17) are orthogonal, $\sum\limits_{i=1}^{c} \mu_i(\mathbf{z}) = 1$ holds. In case of NARX nonlinear dynamic systems, $\mathbf{z}$ and $\mathbf{x}$ are chosen as the vectors of lagged inputs and measured outputs. In general, the components of $\mathbf{z}$ can be different from $\mathbf{x}$ or they can even be a function of those components.

On account of using symbolic interval-valued output in this research for dynamic system identification, $\mathbf{x}$, $\mathbf{z}$, and $\hat{y}_i$ all become intervals (i.e. interval variables $\tilde{\mathbf{x}}$, $\tilde{\mathbf{z}}$, and $\hat{\tilde{y}}_i$, respectively). In the proposed modelling approach, the i-th rule of the TS fuzzy model is characterized by interval consequent parameters and multivariate fuzzy sets with interval prototypes. The overall model is given by:

$$R_i: \quad \text{IF } \tilde{\mathbf{z}} \text{ IS } \tilde{\mathbf{v}}_i \text{ THEN } \hat{\tilde{y}}_i(\tilde{\mathbf{x}}) = \tilde{\mathbf{a}}_i^{\mathsf{T}} \left[1 \quad \tilde{\mathbf{x}}^{\mathsf{T}}\right]^{\mathsf{T}}, \tag{18}$$

$$\hat{\tilde{y}}(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) = \sum_{i=1}^{c} \mu_i(\tilde{\mathbf{z}}) \cdot \hat{\tilde{y}}_i(\tilde{\mathbf{x}}), \tag{19}$$

where $\mu_i(\tilde{\mathbf{z}}) \in [0, 1]$ is the crisp membership value of the interval scheduling variable ($\tilde{\mathbf{z}}$). Since the membership function is orthogonal, $\sum_{i=1}^{c} \mu_i(\tilde{\mathbf{z}}) = 1$. In this model, the interval data is first clustered using the Fuzzy c-means clustering for symbolic interval-valued data (here referred to as IFCM) [32]. It is followed by parameter estimation of the local models using the centre and range method which is suitable for symbolic interval data [33]. After a NARX model is estimated, a non-linear optimization technique is applied to derive an optimal model for the NOE case.

In order to conceptually compare this model with the other possible alternative modelling frameworks, a visual comparison of the models, with respect to the antecedent and consequent structures and the input & output of the model, is illustrated in Fig. 1. Since only dynamic system identification is ad-

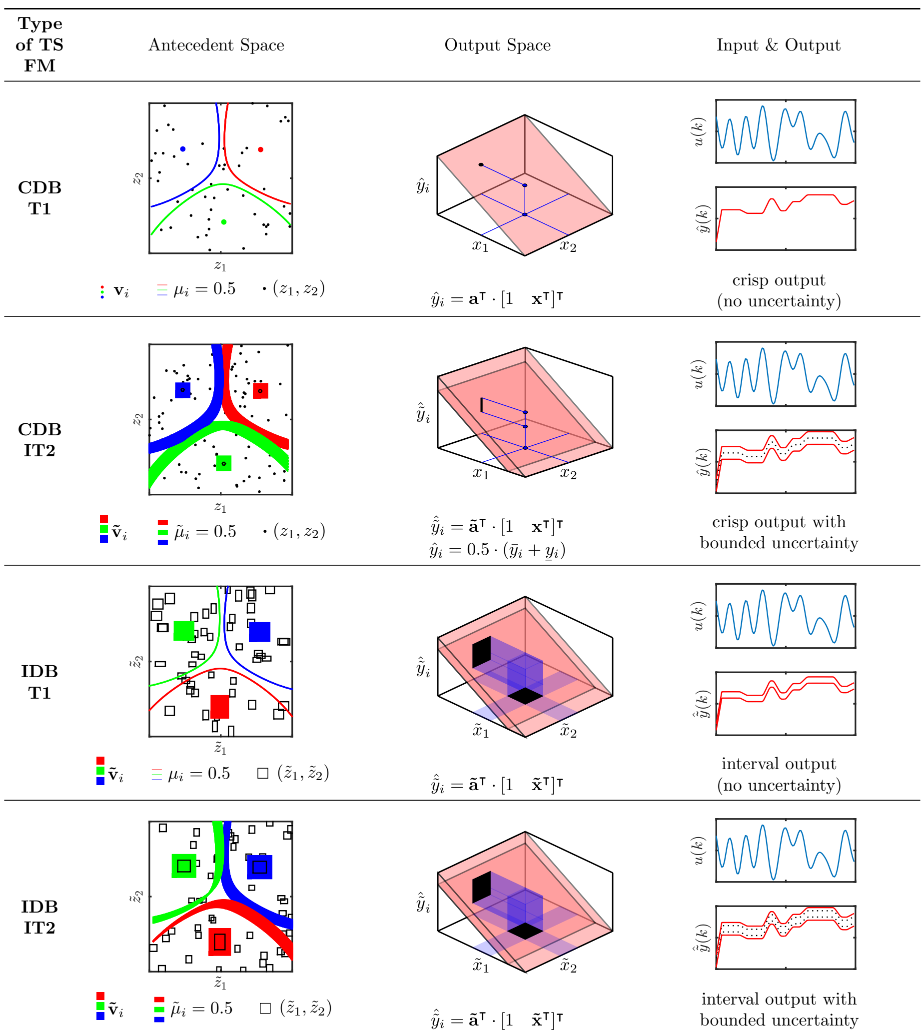

Figure 1: Visual comparison between alternate modelling frameworks

dressed in this article, the antecedent ($\mathbf{z}$ or $\tilde{\mathbf{z}}$) and consequent ($\mathbf{x}$ or $\tilde{\mathbf{x}}$) variables consist of the function of lagged inputs and measured (NARX model) or process (NOE model) outputs. If the NOE model is considered, an example of these variables is $\mathbf{z}(k) = [u(k-1), \hat{y}(k-1)]^\intercal$, $\mathbf{x}(k) = [u(k-1), \hat{y}(k-1)]^\intercal$ for the case of crisp input and crisp output; whereas $\tilde{\mathbf{z}}(k) = [u(k-1), \hat{\tilde{y}}(k-1)]^\intercal$, $\tilde{\mathbf{x}}(k) = [u(k-1), \hat{\tilde{y}}(k-1)]$ for the case of crisp input and interval output. The input signal shown in Fig. 1 is crisp in nature; however, it should not be restrictive, as the modelling procedure remains the same for the interval input. For the ease of illustration, the MF values are plotted only for the value of 0.5. The value of fuzziness parameter $\nu = 2$ was selected so that the membership lines (shown in different colours) of each fuzzy set have some distance or gap between them. As $\nu \to 1$, these membership lines seems to overlap each other. The first of these models is the multi-dimensional reference fuzzy set based fuzzy model presented in [6], and here it is called as the Crisp-Data Based (CDB) Type-1 (T1) TS Fuzzy Model (FM). This model is called CDB because it uses crisp data (input and output) for model estimation. This is the classical case, in which the modal has no uncertainty in membership functions and local model parameters. To add the uncertainty handling capability to the classical FM, interval fuzzy membership functions and interval local model parameters are utilized in the CDB Interval Type-2 (IT2) TS FM, resulting in interval bound with bounded uncertainty (so-called typed-reduced set). When there is stochasticity in the dynamic system to be modelled, the output of such a system can be represented in the form of intervals. The last two models, i.e. the Interval-Data Based (IDB) T1 TS FM, and the IDB IT2 TS FM can be used for modelling such systems with interval data. Mathematically, these models are described for the

Multiple-Input-Single-Output (MISO) case as follows:

$$\begin{aligned}
\text{CDB T1 TS FM:} \quad & \hat{y}(\mathbf{x},\mathbf{z}) = \sum_{i=1}^{c} \mu_i(\mathbf{v}_i|_{\mathbf{v}_1,\ldots,\mathbf{v}_c},\mathbf{z})\left\{\mathbf{a}_i^{\mathsf{T}}\,[1 \quad \mathbf{x}^{\mathsf{T}}]^{\mathsf{T}}\right\}, \\
\text{CDB IT2 TS FM:} \quad & \hat{\tilde{y}}(\mathbf{x},\mathbf{z}) = \sum_{i=1}^{c} \tilde{\mu}_i(\tilde{\mathbf{v}}_i|_{\tilde{\mathbf{v}}_1,\ldots,\tilde{\mathbf{v}}_c},\mathbf{z})\left\{\tilde{\mathbf{a}}_i^{\mathsf{T}}\,[1 \quad \mathbf{x}^{\mathsf{T}}]^{\mathsf{T}}\right\}, \\
\text{IDB T1 TS FM:} \quad & \hat{\tilde{y}}(\tilde{\mathbf{x}},\tilde{\mathbf{z}}) = \sum_{i=1}^{c} \mu_i(\tilde{\mathbf{v}}_i|_{\tilde{\mathbf{v}}_1,\ldots,\tilde{\mathbf{v}}_c},\tilde{\mathbf{z}})\left\{\tilde{\mathbf{a}}_i^{\mathsf{T}}\,[1 \quad \tilde{\mathbf{x}}^{\mathsf{T}}]^{\mathsf{T}}\right\}, \\
\text{IDB IT2 TS FM:} \quad & \hat{\tilde{y}}(\tilde{\mathbf{x}},\tilde{\mathbf{z}}) = \sum_{i=1}^{c} \tilde{\mu}_i(\tilde{\mathbf{v}}_i|_{\tilde{\mathbf{v}}_1,\ldots,\tilde{\mathbf{v}}_c},\tilde{\mathbf{z}})\left\{\tilde{\mathbf{a}}_i^{\mathsf{T}}\,[1 \quad \tilde{\mathbf{x}}^{\mathsf{T}}]^{\mathsf{T}}\right\}.
\end{aligned} \tag{20}$$

As seen from the Fig. 1 and Eq. (20), these model descriptions differ from one another depending upon the crisp or interval nature of data or parameters. Of all these models, the IDB IT2 TS FM, is the most general one. It uses interval data for modelling and incorporates bounded uncertainty in interval cluster prototypes and local model parameters. The output of this model is in the form uncertain interval with bounded uncertainty. Since only the IDB T1 TS FM is discussed in this article, the details of the steps required for estimating this model from the interval data are given in the sequel.

3.5. Partitioning of the antecedent space

The antecedent variable $\tilde{\mathbf{z}}(k)$ which actually contains both the point-valued data (input values) and the symbolic interval-valued data (response) is clustered using IFCM. Since a point-valued data can be seen as a special case of interval-valued data, $\tilde{\mathbf{z}}(k)$ is dealt with the theory of symbolic interval-valued data. The IFCM clustering furnishes the fuzzy partitioning of the space of $\tilde{\mathbf{z}}(k)$ and provides interval prototypes. A brief description of this method is given below, see [32] and the references therein for details.

Let the symbolic interval-valued data to be clustered be given by $\tilde{\mathbf{z}}(k) = (\tilde{z}_1(k),\ldots,\tilde{z}_{r_{\text{ant}}}(k))$, $k = 1,\ldots,N$; where $\tilde{z}_j(k) = [a_k^j, b_k^j] \in \Im = \{[a,b] : a,b \in \mathbb{R}, a \le b\}$, $j = 1,2,\ldots,r_{\text{ant}}$. Let each prototype $\tilde{\mathbf{v}}_i$ of cluster $\tilde{P}_i$ be represented as a vector of intervals, i.e. $\tilde{\mathbf{v}}_i = (\tilde{v}_1^i,\ldots,\tilde{v}_{r_{\text{ant}}}^i)$, $i = 1,\ldots,c$, where $\tilde{v}_j^i = [\alpha_i^j, \beta_i^j] \in \Im = \{[\alpha,\beta] : \alpha,\beta \in \mathbb{R}, \alpha \le \beta\}$, $j = 1,2,\ldots,r_{\text{ant}}$. Let $\nu \in \mathbb{R}$ be the fuzziness parameter. The IFCM minimizes the adequacy criterion based on suitable squared Euclidean distances between vectors of intervals as follows:

$$W = \sum_{i=1}^{c} \sum_{k=1}^{N} \mu_i^{\nu}(\tilde{\mathbf{z}}(k)) \sum_{j=1}^{r_{\text{ant}}} \left[(a_k^j - \alpha_i^j)^2 + (b_k^j - \beta_i^j)^2 \right]. \tag{21}$$

As the standard FCM, the IFCM algorithm starts with the random initialization of either the cluster prototypes or partition matrix, and then subsequently iterates between the representation and allocation steps. In the representation step, the clustering prototypes are updated as follows:

$$\alpha_i^j = \frac{\sum\limits_{k=1}^{N} \mu_i^{\nu}(\tilde{\mathbf{z}}(k)) a_k^j}{\sum\limits_{k=1}^{N} \mu_i^{\nu}(\tilde{\mathbf{z}}(k))} \quad \text{and} \quad \beta_i^j = \frac{\sum\limits_{k=1}^{N} \mu_i^{\nu}(\tilde{\mathbf{z}}(k)) b_k^j}{\sum\limits_{k=1}^{N} \mu_i^{\nu}(\tilde{\mathbf{z}}(k))}. \tag{22}$$

In the allocation step, the values of membership function are updated according to:

$$\mu_i(\tilde{\mathbf{z}}(k)) = \left[\sum\nolimits_{h=1}^{c} \left(\frac{\sum\limits_{j=1}^{r_{\text{ant}}} \left[(a_k^j - \alpha_i^j)^2 + (b_k^j - \beta_i^j)^2 \right]}{\sum\limits_{j=1}^{r_{\text{ant}}} \left[(a_k^j - \alpha_h^j)^2 + (b_k^j - \beta_h^j)^2 \right]} \right)^{\frac{1}{\nu - 1}} \right]^{-1}. \tag{23}$$

The IFCM algorithm has similar convergence properties to the standard FCM algorithm.

3.6. Estimation of the local modal parameters

The centre and range method [33] is used for this purpose. The method applies the weighted linear regression on mid-points (centres) and ranges of the interval valued consequent variable $\tilde{\mathbf{x}}(k)$, which can later be used for determining the output $\tilde{\hat{y}}(k)$.

Let the consequent variable $\tilde{\mathbf{x}}(k)$ be written as $\tilde{\mathbf{x}}(k) = (\tilde{x}_1(k), \ldots, \tilde{x}_{r_{\text{con}}}(k))$ with $\tilde{x}_j(k) = [c_k^j, d_k^j] \in \Im = \{[c, d] : c, d \in \mathbb{R}, c \leq d\}$, $j = 1, \ldots, r_{\text{con}}$ and $k = 1, \ldots, N$. Further assume that $\mathbf{x}^{\text{cen}}(k) = (x_1^{\text{cen}}(k), \ldots, x_{r_{\text{con}}}^{\text{cen}}(k))$, where $x_j^{\text{cen}}(k) = 0.5(c_k^j + d_k^j)$ and $\mathbf{x}^{\text{hr}}(k) = (x_1^{\text{hr}}(k), \ldots, x_{r_{\text{con}}}^{\text{hr}}(k))$, where $x_j^{\text{hr}}(k) =$

$0.5(d_k^j - c_k^j)$; $j = 1, \ldots, r_{\text{con}}$ and $k = 1, \ldots, N$, represent the centre (cen) and half range (hr) values of $\tilde{\mathbf{x}}(k)$. Let the consequent parameters of the i-th rule (the parameters of the i-th local affine model) be represented as $\tilde{\mathbf{a}}_i = (\tilde{a}_0^i, \tilde{a}_1^i, \ldots, \tilde{a}_{r_{\text{con}}}^i)$, where $\tilde{a}_j^i = [\gamma_i^j, \delta_i^j] \in \Im = \{[\gamma, \delta] : \gamma, \delta \in \mathbb{R}, \gamma \leq \delta\}$, with $j = 0, 1, \ldots, r_{\text{con}}$ and $i = 1, \ldots, c$. Moreover, assume $\mathbf{a}_i^{\text{cen}} = (a_0^{i,\text{cen}}, a_1^{i,\text{cen}}, \ldots, a_{r_{\text{con}}}^{i,\text{cen}})$, where $a_j^{i,\text{cen}} = 0.5(\gamma_i^j + \delta_i^j)$ and $\mathbf{a}_i^{\text{hr}} = (a_0^{i,\text{hr}}, a_1^{i,\text{hr}}, \ldots, a_{r_{\text{con}}}^{i,\text{hr}})$, where $a_j^{i,\text{hr}} = 0.5(\delta_i^j - \gamma_i^j)$, $j = 0, 1, \ldots, r_{\text{con}}$ and $i = 1, \ldots, c$, represent the centre (cen) and half range (hr) values of $\tilde{\mathbf{a}}_i$. The vectors lumping all the centre and half range consequent parameters are given by $\mathbf{a}^{\text{cen}} = [(\mathbf{a}_1^{\text{cen}})^\intercal, \ldots, (\mathbf{a}_c^{\text{cen}})^\intercal]^\intercal \in \mathbb{R}^{(r_{\text{con}}+1)c}$, where $\mathbf{a}_i^{\text{cen}} \in \mathbb{R}^{(r_{\text{con}}+1)}$, $i = 1, \ldots, c$ and $\mathbf{a}^{\text{hr}} = [(\mathbf{a}_1^{\text{hr}})^\intercal, \ldots, (\mathbf{a}_c^{\text{hr}})^\intercal]^\intercal \in \mathbb{R}^{(r_{\text{con}}+1)c}$, where $\mathbf{a}_i^{\text{hr}} \in \mathbb{R}^{(r_{\text{con}}+1)}$, $i = 1, \ldots, c$. These vectors are estimated globally by using the Ordinary Least Squares (OLS) method (for the NARX model). Denote $M_i \in \mathbb{R}^{N \times N}$, the diagonal matrix having membership grades $\mu_i(\tilde{\mathbf{z}}(k))$ as its k-th diagonal element with $1 \leq i \leq c$ and $1 \leq k \leq N$. Define the matrices

$$X_e^{\text{cen}} := [X^{\text{cen}}, \mathbf{1}] \in \mathbb{R}^{N \times (r_{\text{con}}+1)}, \tag{24}$$

$$X_e^{\text{hr}} := [X^{\text{hr}}, \mathbf{1}] \in \mathbb{R}^{N \times (r_{\text{con}}+1)}, \tag{25}$$

where $\mathbf{1}$ is a unitary column vector in $\mathbb{R}^n$, X^{cen} and X^{hr} are the input matrices for the centre and radius consequent part, respectively.

$$X^{\text{cen}} := [\mathbf{x}^{\text{cen}}(1), \ldots, \mathbf{x}^{\text{cen}}(n)]^\intercal \in \mathbb{R}^{N \times r_{\text{con}}}, \tag{26}$$

$$X^{\text{hr}} := [\mathbf{x}^{\text{hr}}(1), \ldots, \mathbf{x}^{\text{hr}}(n)]^\intercal \in \mathbb{R}^{N \times r_{\text{con}}}. \tag{27}$$

Moreover, define

$$X_E^{\text{cen}} := [M_1 X_e^{\text{cen}}, \ldots, M_c X_e^{\text{cen}}] \in \mathbb{R}^{N \times (r_{\text{con}}+1)c}, \tag{28}$$

$$X_E^{\text{hr}} := [M_1 X_e^{\text{hr}}, \ldots, M_c X_e^{\text{hr}}] \in \mathbb{R}^{N \times (r_{\text{con}}+1)c}. \tag{29}$$

The reference interval output is defined as $\tilde{\mathbf{y}} = (\tilde{y}(1), \ldots, \tilde{y}(N))$, with $\tilde{y}(k) = [\underline{y}, \bar{y}] \in \Im = \{[a, b] : a, b \in \mathbb{R}, a \leq b\}$, $k = 1, \ldots, N$. The centre and half range values of $\tilde{\mathbf{y}}(k)$ are defined in the same way: $\mathbf{y}^{\text{cen}} = (y^{\text{cen}}(1), \ldots, y^{\text{cen}}(N))$ with $y^{\text{cen}}(k) = 0.5(\underline{y}(k) + \bar{y}(k))$, and $\mathbf{y}^{\text{hr}} = (y^{\text{hr}}(1), \ldots, y^{\text{hr}}(N))$ with $y^{\text{hr}}(k) =$

$(\bar{y}(k) - \underline{y}(k))$, $k = 1, \ldots, N$. The centre and half range values of the parameters of the local models $\mathbf{a}^{\text{cen}}$ and $\mathbf{a}^{\text{hr}}$ are calculated as

$$\mathbf{a}^{\text{cen}} = [(X_E^{\text{cen}})^\intercal X_E^{\text{cen}}]^{-1}(X_E^{\text{cen}})^\intercal \mathbf{y}^{\text{cen}}, \tag{30}$$

$$\mathbf{a}^{\text{hr}} = [(X_E^{\text{hr}})^\intercal X_E^{\text{hr}}]^{-1}(X_E^{\text{hr}})^\intercal \mathbf{y}^{\text{hr}}. \tag{31}$$

3.7. Determination of NOE model

The NOE or Recursive evaluation based prediction model is advantageous for applications like model predictive control or simulation/prognosis ([39]). The cluster prototypes and local model parameters can be optimized for the NOE case. In this case, the parameters obtained for the NARX model are used as the initial parameters for the NOE model and nonlinear optimization is applied to estimate optimal parameters for the NOE case.

The centre and half range of the cluster prototypes $\tilde{\mathbf{g}} = (\tilde{\mathbf{g}}_1, \ldots, \tilde{\mathbf{g}}_c)$ are defined as $\mathbf{g}^{\text{cen}} = [(\mathbf{g}_1^{\text{cen}})^\intercal, \ldots, (\mathbf{g}_c^{\text{cen}}))^\intercal] \in \mathbb{R}^{cr_{\text{ant}}}$, where $\mathbf{g}_i^{\text{cen}} \in \mathbb{R}^{r_{\text{ant}}}$; and $\mathbf{g}^{\text{hr}} = [(\mathbf{g}_1^{\text{hr}})^\intercal, \ldots, (\mathbf{g}_c^{\text{hr}}))^\intercal] \in \mathbb{R}^{cr_{\text{ant}}}$, where $\mathbf{g}_i^{\text{hr}} \in \mathbb{R}^{r_{\text{ant}}}$; $i = 1, \ldots, c$. Denoting the lumped parameter vector of a NARX model as $\boldsymbol{\theta}_{\text{NARX}} := [(\mathbf{g}^{\text{cen}})^\intercal, (\mathbf{g}^{\text{hr}})^\intercal, (\mathbf{a}^{\text{cen}})^\intercal, (\mathbf{a}^{\text{hr}})^\intercal]^\intercal$, with $\boldsymbol{\theta}_{\text{NARX}} \in \mathbb{R}^{2c(r_{\text{ant}}+r_{\text{con}}+1)}$, the optimal set of parameters for the NOE model $\boldsymbol{\theta}_{\text{NOE}}$ is obtained by the minimizing the mean quadratic prediction error of the NOE model for $\tilde{\mathbf{y}} = [\underline{\mathbf{y}}, \bar{\mathbf{y}}]$:

$$\begin{aligned}\boldsymbol{\theta}_{\text{NOE}} := \boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \frac{1}{N}\sum_{k=1}^{N}(\bar{y}(k) - \hat{\bar{y}}_{\text{NOE}}(\boldsymbol{\theta}, k))^2 + \\ (\underline{y}(k) - \hat{\underline{y}}_{\text{NOE}}(\boldsymbol{\theta}, k))^2.\end{aligned} \tag{32}$$

In this research, the Matlab function `lsqnonlin` is used for solving the optimization problem in (32). This function uses a trust-region-reflective algorithm based on the interior-reflective Newton method [40, 41].

3.8. Model assessment criteria

Once the model has been developed its performance and validity are assessed. The three time series: the expected response, the output upper and lower bound time series are assessed separately. Denoting $y(k)$ and $\hat{y}(k)$ as the reference and the model output respectively, the model quality is assessed using Variance Accounted For (VAF), Root-Mean-Square Error (RMSE) and Maximum Absolute Error (MaxAE) [35]:

$$\mathrm{VAF} = \left(1 - \frac{\mathrm{var}\ (\mathbf{y} - \hat{\mathbf{y}})}{\mathrm{var}\ (\mathbf{y})}\right) 100\%, \tag{33}$$

$$\mathrm{var}(\mathbf{y}) = \frac{1}{N-1}\sum_{k=1}^{N}(y(k) - \bar{y}))^2,\ \bar{y} = \frac{1}{N}\sum_{k=1}^{N} y(k),$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y(k) - \hat{y}(k))^2}, \tag{34}$$

$$\mathrm{MaxAE} = \max_{1 \leq k \leq N}(|\hat{y}(k) - y(k)|). \tag{35}$$

4. Experimental results

4.1. Simulation case study with known true model

The academic simulation benchmark system chosen for demonstrating the effectiveness of the developed method is taken from [42]. It is a non-linear SISO dynamic system given by the second order difference equation (36):

$$y(k+1) = \frac{y(k)y(k-1)(y(k)+2.5)}{1 + y^2(k) + y^2(k-1)} + u(k) \tag{36}$$

The input signal $u(k)$ used for identification and validation is shown in Fig. 2. It is an i.i.d random variable distributed uniformly in $[-2, 2]$ and consisted of $N = 2000$ samples. In order to increase stochasticity in the input signal, the 'change probability' - the probability with which the input value changes to a new random value from its previous value - is introduced and the value is set to 0.5. The initial conditions are taken as $y(0) = 0$ and $y(1) = 0$. The data is split into 50% for identification and the remaining 50% for validation. The output of the original system is deterministic. To introduce variability in the output, the output values are perturbed around its mean values by making it a random

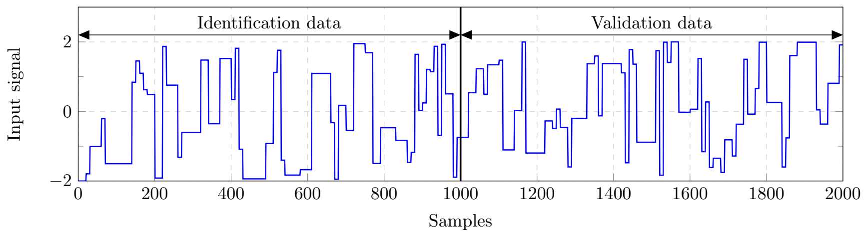

Figure 2: The input signal for the identification (left) and validation (right) data.

variable following a specific pdf. The deterministic output of the system (36) is taken as the true output ($y(k)$). In reality this output time series is not available, but it is demonstrated that the expected time series is a good estimate of the true time series provided that the uncertainty is distributed symmetrically as given in Eq. 37. With the output signal defined as y_k, a random variable at the time instant k, its pdf is computed as follows:

$$\begin{aligned}
\Pr(\mathsf{y}_k) &= 0.5 \cdot \Pr_1 + 0.125 \cdot \Pr_2 \\
&\quad + 0.125 \cdot \Pr_3 + 0.125 \cdot \Pr_4 + 0.125 \cdot \Pr_5, \\
\Pr_l &= \frac{1}{\sqrt{2\pi s^2}} \exp\left(- \frac{(\mathsf{y}_k - \mu_l)^2}{2s^2} \right), \; l = 1, \ldots, 5, \\
\mu_1 &= y(k), \; \mu_2 = 0.9 \cdot y(k), \; \mu_3 = 1.1 \cdot y(k), \\
\mu_4 &= 0.8 \cdot y(k), \; \mu_5 = 1.2 \cdot y(k), \; s = 0.1
\end{aligned} \tag{37}$$

The input signal is repeated $M = 100$ times, resulting in a random sample of output values of the size M at each time instant. For the sake of illustration, the normalized histogram (integral equals unity) of the output sample at $k = 1000$, with the corresponding conditional pdfs ($\Pr(\mathsf{y}_k|\mathsf{y}_{k-1}, \ldots, \mathsf{y}_{k-n_y}, u_{k-\tau-1}, \ldots, u_{k-\tau-n_u})$) for each experiment ($m = 1, \ldots, M$), the mean output value ($y_k^{\exp} = 4.36$) , the true output value ($y(k) = 4.41$), confidence interval based the lower bound ($\underline{y}_k = 3.69$) and the upper bound ($\bar{y}_k = 5.02$), is shown in Fig. 3.

The value of α is chosen to be 0.25, which gives rise to 75% CI bounds and

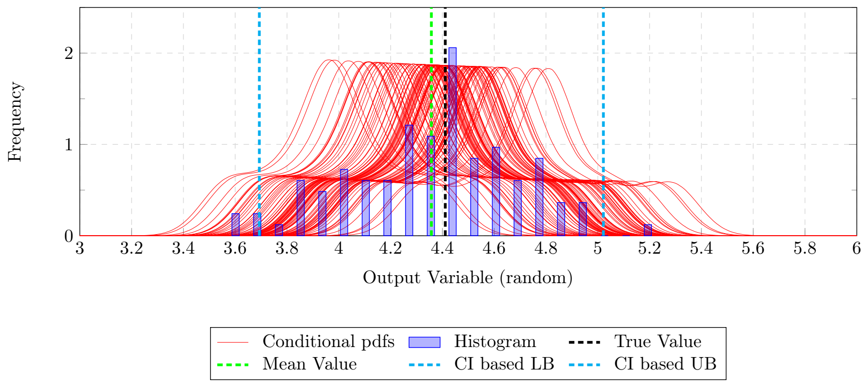

Figure 3: y_k at $k = 1000$

the coverage factor $k_{\text{cov}} = 2.0258$. It is remarked here that the bound obtained using the extended Chebyshev's inequality is conservative because it does not take into account the distribution of data, but it is always guaranteed that $3/4^{\text{th}}$ of the time, the value of the random variable y_k lies within this bound, no matter what the distribution is. The distribution is chosen to be symmetric, for the reason of the assumption that mean is the best point estimator. Recall that, for a symmetric distribution, the most likely value or mode is same as the mean. In the case of a asymmetric distribution, the most likely value is mode of the distribution, which is not considered in this research. The output time series for the the identification and validation data is illustrated in Fig. 4. The task is to estimate the lower and upper bounds of the response. Once they are obtained, the mean time series can be computed by averaging them, according to Eq. (14). The same quantities are used as the antecedent and consequent variables, i.e. $\tilde{\mathbf{x}}(k) = \tilde{\mathbf{z}}(k) = [u(k-1), \tilde{y}(k-1), \tilde{y}(k-2)]^\intercal$. The value of the fuzzy index is chosen as $\nu = 1.3$, which is within the range suggested by [43] for nonlinear identification problems. For having a parsimonious model, six clusters (or local models) are selected for identification; further increasing the number

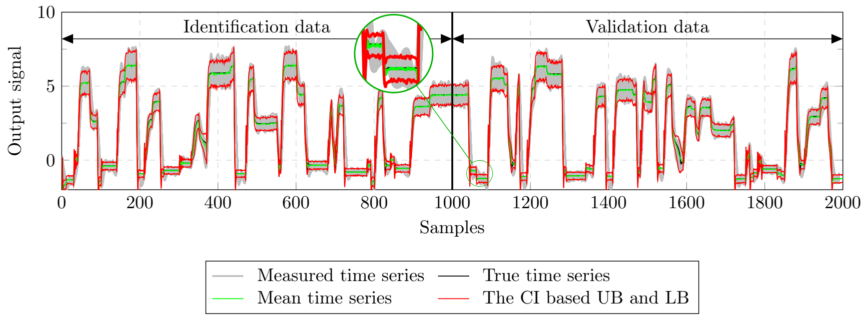

Figure 4: The output signal for the identification and validation data

of clusters improves the modelling performance very slightly. The modelling performance is shown in Table 1. As evident from the table, the model is able to estimate the mean time series and envelopes with good accuracy. To visually inspect the modelling performance, the input signal and the output envelopes for the identification and the validation data are shown in the Fig. 5. Further improvement can be done by either increasing the number of local models, optimal selection of antecedent and consequent structure, and/or using the optimization algorithm to get the (sub) optimal (in the modelling sense) sets of parameters.

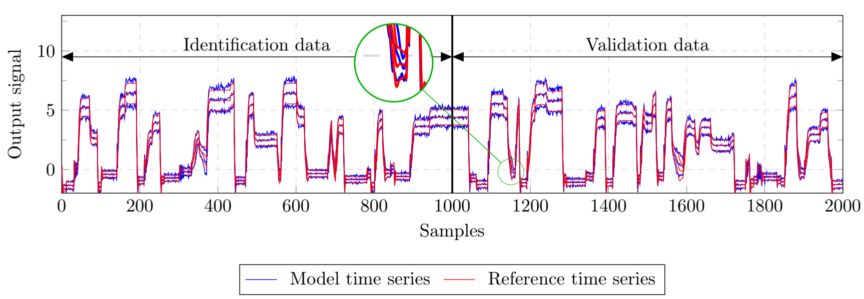

Figure 5: The reference and the model time series for identification and test data

4.2. Electro-mechanical throttle

The developed method is demonstrated for the data recorded from the test stand of electro-mechanical throttle shown in Fig. 6. The system is proposed as a benchmark problem for nonlinear system identification with friction [34]. Electro-mechanical throttles are standard components in diesel and Otto combustion engines and are therefore widespread deployed. A phase optimized multisine input signal proposed in [44], see Fig. 7, is chosen as input signal. The amplitude and the offset of the multisine signal are determined in such a way that they should be able to excite all important operational system characteristics as well as to avoid staying at the mechanical hard stop limits too often. The length of experiment is selected to be $N = 1000$ and the experiment is repeated 80 times ($M = 80$). The sampling time is chosen to be 1 ms which adequately captures the dynamics of the given system. For simplicity, the model structure is utilized as used for T1 TS fuzzy modelling of the throttle in [34], i.e., scheduling variable $\tilde{\mathbf{z}} = [u(k-1),\, \tilde{y}(k-1) - \tilde{y}(k-2)]^\intercal$, consequent variable, $\tilde{\mathbf{x}} = [u(k-1), \tilde{y}(k-1), \tilde{y}(k-2)]^\intercal$, the number of local models $c = 8$, and the fuzziness parameter $\nu = 1.1$. For having a parsimonious model, the value of c is selected based on the knee point of J_{NOE} - objective function containing MSE of NOE model - after which no considerable improvement in model performance is observed. The value of α is selected to be 0.25 (for 75% CI). The first 90 %

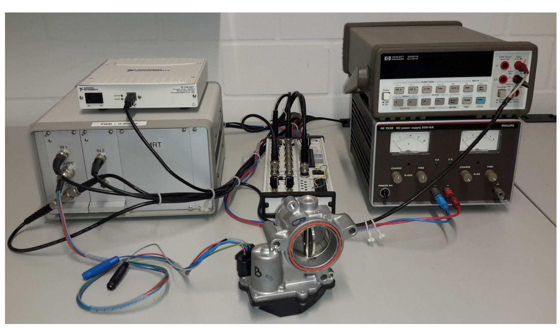

Figure 6: Experimental test stand of electro-mechanical throttle

of data ($1-9$ sec.) is 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. 8 and presented in Table 2. It is evident that the resulting model is able to reasonably capture the bounds and mean response. The maximum error of 1.43° is observed in the case of estimating the upper bound (UB) for the identification data, which is already considered within the reasonable limit.

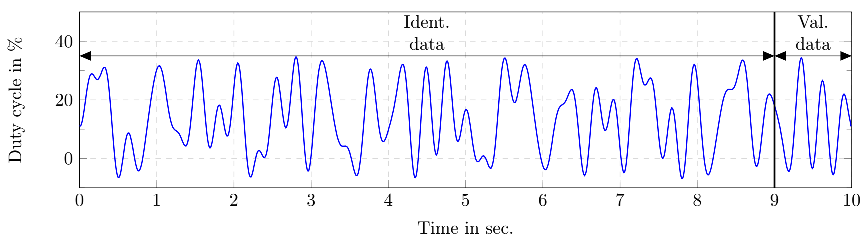

Figure 7: Multisine input signal for the throttle

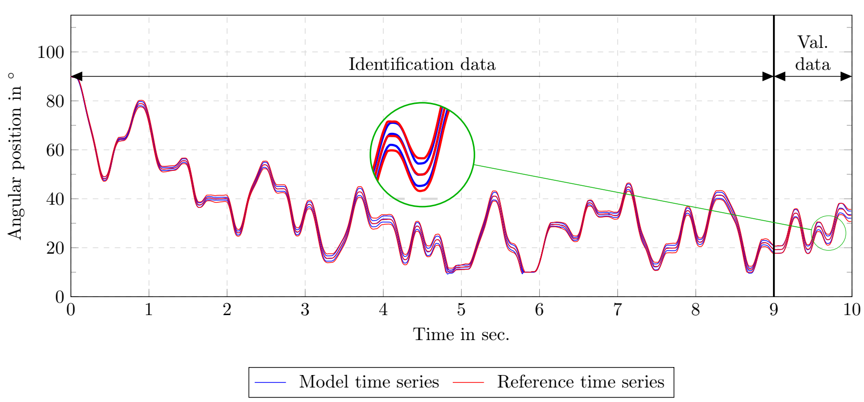

Figure 8: The reference and model time series for the identification (first 9 sec.) and validation data (10-th sec.)

		Identification data	Validation data
	VAF	99.79	99.80
LB	MaxAE	0.85	0.67
	RMSE	0.11	0.11
	VAF	99.89	99.84
UB	MaxAE	0.72	0.69
	RMSE	0.09	0.11
	VAF	99.90	99.86
Mean	MaxAE	0.70	0.70
	RMSE	0.08	0.10

Table 1: Modelling performance for the Upper Bound (UB), the Lower Bound (LB) and the mean time series for the academic example

		Identification data	Validation data
	VAF in %	99.97	99.71
LB	MaxAE in °	1.29	1.36
	RMSE in °	0.62	0.55
	VAF in %	99.98	99.78
UB	MaxAE in °	1.43	0.86
	RMSE in °	0.63	0.46
	VAF in %	99.99	99.94
Mean	MaxAE in °	1.04	0.39
	RMSE in °	0.14	0.15

Table 2: Modelling performance for the Upper Bound (UB), the Lower Bound (LB) and the mean time series for the electro-mechanical throttle

5. Conclusion and outlook

An approach to build models that can provide the expected response of an uncertain nonlinear dynamic system along with the spread around it has been presented and demonstrated for a nonlinear academic benchmark system and an electro-mechanical throttle. The approach can use either the max-min bounds or the extended Chebyshev's inequality to obtain upper and lower bounds of the output time series based on the confidence interval. These bounds are directly used in the TS fuzzy model using the theory of symbolic interval-valued data. The developed model is referred to as the Interval-Data Based (IDB) Type-1 (T1) TS FM and its comparison with the other possible alternative model description are presented in this article. The model provides the estimates of the upper and lower bounds, whereas the mean response can be calculated by averaging them. In the current approach, the IDB T1 TS FM directly uses interval data in the modelling procedure by utilization the techniques of clustering and regression of symbolic interval-valued data. Consequently, the obtained parameters are in the form of interval. The results show that the presented approach is able to adequately model the stochastic effects due to the variability in system output in a unified efficient modelling framework. In future, some other techniques for obtaining less-conservative and robust upper and lower bounds from systems will be explored. Moreover, the extension of the current model to the case of Interval-Data Based (IDB) Interval-Type 2 (IT2) TS FM will be investigated.

Acknowledgement

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the article. The first author is also grateful to the German Academic Exchange Service (DAAD) for the financial grant.

Vitae

M.Sc. Salman Zaidi (1984) is a research associate at the Department of Measurement and Control at the University of Kassel. His research areas are fuzzy identification, intelligent control, intelligent optimization techniques and machine learning. E-Mail: salman.zaidi@mrt.uni-kassel.de

Univ.-Prof. Dr.-Ing. Andreas Kroll (1967) is head of the Department of Measurement and Control at the University of Kassel. His research areas are nonlinear identification and control methods, computational intelligence, and complex systems. E-Mail: andreas.kroll@mrt.uni-kassel.de

References

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

[2] M. Norgaard, O. Ravn, N. K. Poulsen, L. K. Hansen, Neural networks for modelling and control of dynamic systems, Springer-Verlag, London, 2003.

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

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

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

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

[7] R. Babuška, Fuzzy Modeling for Control, Kluwer Academic Publishers, Boston, 1998.

[8] O. Nelles, Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models, Springer-Verlag, Berlin Heidelberg, 2013.

[9] J. M. Mendel, Uncertain rule-based fuzzy logic systems, introduction and new directions, Prentice Hall, 2001.

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

[11] L. Zadeh, The concept of a linguistic variable and its application to approximate reasoning–I, Information Sciences 8(3) (1975) 199–249.

[12] H. Hagras, A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots, IEEE Transactions on Fuzzy Systems 12(4) (2004) 524–539.

[13] Q. Liang, J. Mendel, Interval type-2 fuzzy logic systems: theory and design, IEEE Transactions on Fuzzy Systems 8(5) (2000) 535–550.

[14] D. Wu, On the fundamental differences between interval type-2 and type-1 fuzzy logic controllers, IEEE Transactions on Fuzzy Systems 20(5) (2012) 832–848.

[15] H. F. Zarandi, I. Türkşen, T. Kasbi", Type-2 fuzzy modeling for desulphurization of steel process, Expert Systems with Applications 32(1) (2007) 157 – 171.

[16] O. Castillo, P. Melin, A review on the design and optimization of interval type-2 fuzzy controllers, Applied Soft Computing 12(4) (2012) 1267 – 1278.

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

[18] Z. Liu, H.-X. Li, Probabilistic fuzzy logic system: A tool to process stochastic and imprecise information, in: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Jeju Island, Korea, 2009, pp. 848 – 853.

[19] A. Khosravi, S. Nahavandi, D. Creighton, R. Naghavizadeh, Prediction interval construction using interval type-2 fuzzy logic systems, in: Proceedings of the 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Brisbane, QLD, 2012, pp. 1 – 7.

[20] A. Khosravi, S. Nahavandi, An interval type-2 fuzzy logic system-based method for prediction interval construction, Applied Soft Computing 24 (2014) 222 – 231.

[21] H. H. Bock, E. Diday, Analysis of Symbolic Data. In: Exploratory Methods for Extracting Statistical Information from Complex Data, Springer-Verlag, Berlin Heidelberg, 2000.

[22] L. Billard, E. Diday, Symbolic Data Analysis: Conceptual Statistics and Data Mining, John Wiley & Sons, 2012.

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

[24] I. Škrjanc, S. Blažič, O. Agamennoni, Interval Fuzzy Model Identification Using l_∞ -Norm, IEEE Transactions on Fuzzy Systems 13(5).

[25] Z. Xu, C. Sun, 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, Shanghai, China, 2009, pp. 4144 – 4149.

[26] C. Sun, Z. Xu, Extended t-s fuzzy model based on interval arithmetic and its application to interval nonlinear regression analysis, in: IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Jeju Island, Korea, 2009, pp. 1773 – 1778.

[27] S. Zaidi, A. Kroll, On identifying nonlinear envelop type dynamical T-S fuzzy models for systems with uncertainties: method and application to electro-mechanical throttle, in: 23. Workshop Computational Intelligence, Dortmund, Germany, 2013, pp. 129 – 143.

[28] S. Zaidi, A. Kroll, On identifying envelop type nonlinear output error takagi-sugeno fuzzy models for dynamic systems with uncertainties, in: 19th IFAC World Congress, Cape Town, South Africa, 2014, pp. 3226 – 3231.

[29] S. Zaidi, A. Kroll, A novel approach to t-s fuzzy modeling of nonlinear dynamic systems with uncertainties using symbolic interval-valued outputs, in: Proceedings of the 17th IFAC Symposium on System Identification (SysID), Beijing, China, 2015, pp. 1196 – 1201.

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

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

[32] F. d. A. T. de Carvalho, Fuzzy c-means clustering methods for symbolic interval data, Pattern Recognition Letters 28(4) (2007) 423 – 437.

[33] E. d. A. Lima Neto, F. d. A. de Carvalho, Centre and range method for fitting a linear regression model to symbolic interval data, Computational Statistics & Data Analysis 52(3) (2008) 1500 – 1515.

[34] S. Zaidi, A. Kroll, Electro-mechanical throttle as a benchmark problem for nonlinear system identification with friction, in: 24. Workshop Computational Intelligence, Dortmund, Germany, 2014, pp. 173 – 186.

[35] A. Kroll, H. Schulte, Benchmark problems for nonlinear system identification and control using soft computing methods: Need and overview, Applied Soft Computing 25 (2014) 496 – 513.

[36] N. M. Razali, Y. B. Wah, Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests, Journal of Statistical Modeling and Analytics 2(1).

[37] L. Ljung, System Identification-Theory for the User, 2nd Edition, Prentice-Hall, Upper Saddle River, N.J., 1999.

[38] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

[39] M. Jelali, A. Kroll, Hydraulic Servo-systems: Modelling, Identification and Control, Springer-Verlag, London, 2002.

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

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

[42] K. Narendra, K. Parthasarathy, Adaptive identification and control of dynamical systems using neural networks, in: Proceedings of the 28th IEEE Conference on Decision and Control, Tampa, FL, 1989, pp. 1737 – 1738.

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

[44] Z. Ren, A. Kroll, M. Sofsky, F. Laubenstein, 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, Brussels, Belgium, 2012, pp. 1395–1400.