Knowledge-Based Systems 73 (2015) 124–133

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Identification of uncertain nonlinear systems: Constructing belief rule-based models ☆

Yu-Wang Chen [a,*], Jian-Bo Yang [a], Chang-Chun Pan [b], Dong-Ling Xu [a], Zhi-Jie Zhou [c]

[a] Decision and Cognitive Sciences Research Centre, Manchester Business School, The University of Manchester, Manchester M15 6PB, United Kingdom
[b] Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China
[c] High-Tech Institute of Xi'an, Xi'an, Shaanxi 710025, China

ARTICLE INFO

Article history:
Received 27 September 2013
Received in revised form 10 September 2014
Accepted 26 September 2014
Available online 6 October 2014

Keywords:
Identification
Uncertain nonlinear system
Belief rule base
Interval
Training

ABSTRACT

The objective of this paper is to construct reliable belief rule-based (BRB) models for the identification of uncertain nonlinear systems. The BRB methodology is developed from the evidential reasoning (ER) approach and traditional IF–THEN rule based system. It can be used to model complicated nonlinear causal relationships between antecedent attributes and consequents under different types of uncertainty. In a BRB model, various types of information and knowledge with uncertainties can be represented using belief structures, and a belief rule is designed with belief degrees embedded in its possible consequents. In this paper, we first introduce the BRB methodology for modelling uncertain nonlinear systems. Then we present a comparative analysis of three BRB identification models through combining the BRB methodology with nonlinear optimisation techniques. The novel BRB identification models using l_∞-norm and minimising mean uncertainties in belief rules (MUBR) show remarkable capabilities of capturing the lower and upper bounds of the interval outputs of uncertain nonlinear systems simultaneously. Trade-off analysis between identification accuracy and interval credibility are briefly discussed. Finally, a numerical study of a simplified car dynamics is conducted to demonstrate the capability and effectiveness of the BRB identification models for the modelling and identification of uncertain nonlinear systems.
© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Identification of an uncertain nonlinear system is mainly concerned with characterising the unknown nonlinear system on the basis of measured input–output data in an uncertain environment [26]. It is of fundamental importance in predictive control, fault diagnosis, signal processing and decision analysis, since most real-life systems are nonlinear, and are often associated with uncertainties due to noises, unpredictable disturbance and measurement errors, uncertain physical parameters, incomplete knowledge, etc. [1,43,40,30,20]. Thus a great difficulty in applying traditional identification techniques is dealing with those uncertainties [2,4]. Over past few decades, extensive studies have been conducted for effectively identifying uncertain nonlinear systems, especially with the advent of neural network and fuzzy rule-based system techniques [34,29,17,12]. Neural networks and fuzzy rule-based systems often outperform traditional identification methods in terms of both approximation accuracy and identification reliability. Sjöberg et al. [34] provided a unified overview on nonlinear black-box system identification models with structures based on neural networks and fuzzy rules. Tseng and Chen [40] applied the Takagi–Sugeno fuzzy model to model uncertain nonlinear systems and proposed H_∞ fuzzy filter design for state estimation of nonlinear discrete systems with bounded but unknown disturbance. Nelles [29] provided an in-depth analysis of nonlinear system identification methods, including linear and polynomial approximation, neural networks and fuzzy models. Zheng et al. [49] presented a robust Takagi–Sugeno fuzzy control model for nonlinear systems with both parameter uncertainty and external disturbance. Choi [12] developed an adaptive fuzzy control system for uncertain Takagi–Sugeno fuzzy models with norm-bounded uncertainty on the basis of the variable structure control (VSC) theory. González-Olvera and Tang [17] presented a continuous-time recurrent neuro-fuzzy network for the black-box identification of a class of dynamic nonlinear systems.

However, these aforementioned methods may not be directly applicable to some real-world uncertain nonlinear systems, because they involve some restrictive assumptions, such as Gaussian-distributed noises, deterministic disturbances and

☆ Part of this work was presented at FLINS 2012 Istanbul, Turkey and IUKM2013 Beijing, China.

* Corresponding author.
E-mail address: yu-wang.chen@mbs.ac.uk (Y.-W. Chen).

http://dx.doi.org/10.1016/j.knosys.2014.09.010
0950-7051/© 2014 Elsevier B.V. All rights reserved.

bounded uncertain parameters [5]. In addition, the outputs of uncertain systems are usually represented on the basis of predicted point values. In uncertain systems, any uncertainties usually generate uncertain outputs and the commonly used point prediction delivers no information about different kinds of uncertainties [28]. Therefore, the identification of using prediction intervals (i.e., estimated range of predictions) to model uncertain outputs has attracted much attention in recent years. Shrestha and Solomatine [33] employed prediction intervals to identify the uncertainty of the model outputs using different machine learning techniques, such as locally weighted regression and artificial neural network. Mazloumi et al. [28] discussed how uncertainty arising from both model structure and training data can be quantified through constructing prediction intervals in neural networks. Khosravi et al. [21] proposed a lower upper bound estimation (LUBE) method for construction of neural network-based prediction intervals. The LUBE method uses a neural network with two outputs to construct upper and lower bounds of the prediction intervals, and both interval width and coverage probability are incorporated into the training objective function. However, due to the black-box nature of neural networks, these neural network-based prediction intervals cannot be interpreted in an explicit way and does not reflect uncertain knowledge. Furthermore, on the basis of fuzzy identification methodology, Škrjanc et al. [35,36] presented an interval fuzzy model (INFUMO) to model a class of nonlinear systems with interval parameters. It results in a lower and upper fuzzy model or so called a fuzzy model with lower and upper parameters. Linear programming techniques are used to find the set of lower and upper parameters using the l_∞-norm as the optimality criterion. The idea behind the INFUMO is to find optimal lower and upper bound fuzzy systems that define a prediction interval which encloses all the measured input–output data [37]. However, in INFUMO the lower and upper bound fuzzy models are independent from each other, and it may incur improper identification results with invalid lower and upper bounds if there is no enough training data available in some regions of the input space [36,31].

The motivation of this paper is to construct reliable belief rule-based (BRB) models for the identification of uncertain nonlinear systems. The BRB methodology is developed on the basis of the Dempster-Shafer theory of evidence [14,32], decision theory [47,16], traditional IF–THEN rule-based systems [18,13,24] and relevant artificial intelligence (AI) techniques [39,22,15]. The BRB methodology has an inherent capability of dealing with various types of uncertainty. In a BRB system, various types of information and knowledge with uncertainties can be represented using belief structures, and a belief rule is designed with belief degrees embedded in its possible consequents. The belief structure used in both belief rules and inference processes provides a unified scheme to model uncertain system outputs caused by vagueness, fuzziness, or incompleteness, etc. A belief distribution with incompleteness can be transformed to a prediction interval in a straightforward way. Compared with traditional rule based systems, BRB systems provide a more informative knowledge representation scheme for both quantitative data and qualitative information with uncertainties, and it is also capable of approximating complicated nonlinear causal relationships between antecedent inputs and output [45]. In recent years, it has been successfully applied in various areas, such as fault diagnosis, system identification, forecasting and decision analysis [45,46,42,3,50,6,48,25].

The rest of the paper is organised as follows. In Section 2, the BRB methodology for modelling uncertain nonlinear systems is introduced. In Section 3, a comparative analysis of three identification models for uncertain nonlinear systems is presented through combining the BRB methodology with nonlinear optimisation techniques. The novel BRB interval identification model using l_∞-norm and minimising mean uncertainties in belief rules are capable of capturing the lower and upper bounds of the interval outputs of uncertain nonlinear systems simultaneously. Numerical studies and trade-off analysis are introduced to illustrate the performance of the BRB identification models. In Section 4, a numerical study is conducted to demonstrate the superior capability of the proposed BRB identification models for the identification of uncertain nonlinear systems. The paper is concluded in Section 5.

2. Modelling uncertain nonlinear systems with belief rules

The modelling and identification of an uncertain nonlinear system is basically to characterise an unknown relationship between a set of input variables $\boldsymbol{x} = \{x_i;\ i = 1, \ldots, M\}$ and a dependent variable y using a finite number of input–output datasets $\{\boldsymbol{x}_t;y_t\}$, $t = 1, \ldots, T$, where y_t denotes the measured output at the sampling time t. Correspondingly a BRB identification model can be described as the set of input variables and a vector of parameters which are combined in a nonlinear manner to predict the behaviour of the dependant variable. The model can be represented by

$$y_t = f(\boldsymbol{x}_t; \boldsymbol{P}^*) + \varepsilon_t; \quad t = 1, \ldots, T \tag{1}$$

where $\boldsymbol{P}^*$ is used to denote the optimal values of the set of parameters, and ε_t is the model error which is usually assumed to be independently and normally distributed. The analytical model $f(\boldsymbol{x};\boldsymbol{P}^*)$ is decided by model structure, belief rules and inference process which will be introduced in the following sections.

2.1. Belief rules

To model the uncertain nonlinear relationship between the set of input variables $\boldsymbol{x}$ and the dependent variable y, a belief rule base which is made up of a finite number of belief rules can be constructed. Typically, a belief rule is given in the following form [45,6].

$$R_k: \begin{array}{l} \text{IF } x_1 \text{ is } A_1^k \wedge x_2 \text{ is } A_2^k \wedge \cdots \wedge x_{M_k} \text{ is } A_{M_k}^k \\ \text{THEN}\{(D_1, \beta_{1,k}), (D_2, \beta_{2,k}), \ldots, (D_N, \beta_{N,k})\}, \left(\sum_{n=1}^{N} \beta_{n,k} \leqslant 1\right) \\ \text{with rule weight } \theta_k \\ \text{and weight of variables } \delta_{1,k}, \delta_{2,k}, \ldots, \delta_{M_k,k}, k \in \{1, \ldots, K\}, \end{array} \tag{2}$$

where $x_1, x_2, \ldots, x_{M_k}$ denote the antecedent variable in the kth rule, and these variables belong to the complete set of input variables $\boldsymbol{x} = \{x_i; i = 1, \ldots, M\}$. $A_i^k (i = 1, \ldots, M_k)$ is the referential value taken by the ith antecedent variable in the kth rule and $A_i^k \in \boldsymbol{A}_i$. $\boldsymbol{A}_i = \{A_{i,j};\ j = 1, \ldots, J_i\}$ denotes the set of referential values for the ith antecedent variable and J_i is the number of the referential values. As a set of referential values needs to be defined for each antecedent variable, and the antecedents in a belief rule is a combination of the referential values of antecedent variables, a BRB model essentially decomposes the input space of an uncertain nonlinear system into multiple subspaces. The number of referential values on each antecedent variable decides the granularity and interpretability of the subspaces [7]. $\beta_{n,k}$($n = 1, \ldots, N$; $k = 1, \ldots, K$) represents the belief degree to which the consequent element D_n is believed to be the consequent, given the logical relationship of the kth rule IF x_1 is $A_1^k \wedge x_2$ is $A_2^k \wedge \cdots \wedge x_{M_k}$ is $A_{M_k}^k$. The element D_n in the set of consequent elements $\boldsymbol{D} = \{D_n;\ n = 1, \ldots, N\}$ can either be a conclusion or an action and a subset of elements can also be part of the consequent [42]. The nonlinear inference process of BRB models which will be discussed below is based on the belief distribution $\{(D_n,\beta_{n,k});\ n = 1, \ldots, N\}$. If $\sum_{n=1}^{N} \beta_{n,k} = 1$, the kth rule is said to be complete; otherwise, it is incomplete, and the incomplete belief degree

$\beta_{\mathbf{D},k} = \left(1 - \sum_{n=1}^{N} \beta_{n,k}\right)$. θ_k is the relative weight of the kth rule, and $\delta_{M_k,k}$ represents the relative weight of variables in the kth rule.

On the basis of the unified scheme, a belief rule can represent a functional mapping between antecedent inputs and output possibly with uncertainties. In practice, the belief rules can be constructed by different ways, such as extracted from experts' knowledge, learned from historical data, or just generated randomly. If different ways are jointly used to establish a rule base, information inconsistency detection methods should be employed so as to reduce information partiality, inconsistency, redundancy, etc. [27]. Once a belief rule base is established, the knowledge embedded in all belief rules can be used to perform inference for a specific input vector [45,7].

2.2. Inference process

To estimate the output given an input vector $\boldsymbol{x}_t = \{x_{t,i};\ i = 1, \ldots, M\}$ at the sampling time t, each input $x_{t,i}$ needs to be transformed to the following belief distribution using the referential values defined for the ith antecedent variable x_i [44].

$$S(x_{t,i}) = \{(A_{i,j}, \alpha_{i,j}); j = 1, \ldots, J_i\} \tag{3}$$

where

$$\alpha_{i,j} = \frac{A_{i,j+1} - x_{t,i}}{A_{i,j+1} - A_{i,j}} \text{ and } \alpha_{i,j+1} = 1 - \alpha_{i,j}, \text{ if } A_{i,j} \leqslant x_{t,i} \leqslant A_{i,j+1}$$
$$\alpha_{i,j'} = 0, \text{ for } j' = 1, \ldots, J_i \text{ and } j' \neq j, j+1$$

Here, $\alpha_{i,j}$ represents the similarity degree to which the input value $x_{t,i}$ matches the referential value $A_{i,j}$. After all the inputs are transformed into belief distributions, the activation weight of the kth belief rule can be calculated as follows [45].

$$w_k(\boldsymbol{x}) = \frac{\theta_k \prod_{i=1}^{M_k} \left(\alpha_{i,j}^k\right)^{\bar{\delta}_i}}{\sum_{l=1}^{k} \left[\theta_l \prod_{i=1}^{M_l} \left(\alpha_{i,j}^l\right)^{\bar{\delta}_i}\right]} \text{ and } \bar{\delta}_i = \frac{\delta_i}{\max_{i=1,\ldots,M_k}\{\delta_i\}} \tag{4}$$

Further, the belief degrees on the inference output can be generated through the aggregation of all activated belief rules using the following analytical evidential reasoning (ER) approach [41].

$$\{D_n\} : \beta_n(\boldsymbol{x}) = \mu \times \left[\prod_{k=1}^{K}\left(w_k(\boldsymbol{x})\beta_{n,k} + 1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right) - \prod_{k=1}^{K}\left(1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right)\right] \tag{5}$$

$$\{\boldsymbol{D}\} : \beta_D(\boldsymbol{x}) = \mu \times \left[\prod_{k=1}^{K}\left(1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right) - \prod_{k=1}^{K}(1 - w_k(\boldsymbol{x}))\right] \tag{6}$$

where

$$\mu = \sum_{j=1}^{N}\prod_{k=1}^{K}\left(w_k(\boldsymbol{x})\beta_{j,k} + 1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right) - (N-1)\prod_{k=1}^{K}\left(1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right) - \prod_{k=1}^{K}(1 - w_k(\boldsymbol{x}))$$

$\beta_D(\boldsymbol{x})$ represents the remaining belief degree unassigned to any known D_n. It has been proven that $\sum_{n=1}^{N}\beta_n(\boldsymbol{x}) + \beta_D(\boldsymbol{x}) = 1$ [44]. If all belief rules are complete, i.e., $\sum_{n=1}^{N}\beta_{n,k} = 1; k = 1, \ldots, K$, then $\beta_D(\boldsymbol{x}) = 0$. The final inference output $y(\boldsymbol{x})$ by aggregating the K rules activated by the actual input vector $\boldsymbol{x}$ can be represented as

$$S(\hat{y}(\boldsymbol{x})) = \{(D_n, \beta_n(\boldsymbol{x})); n = 1, \ldots, N\} \tag{7}$$

where $\hat{y}(\boldsymbol{x})$ denotes the output of the BRB identification model.

2.3. BRB prediction interval

Suppose that the utility of each consequent element D_n is denoted by $u(D_n)$. The numerical output of the BRB identification model is given as

$$\hat{y}(\boldsymbol{x}) = \sum_{n=1}^{N} u(D_n)\beta_n(\boldsymbol{x}) \tag{8}$$

In uncertain nonlinear systems, belief rules may not all be complete, and then we may have the remaining belief degree $\beta_D(\boldsymbol{x}) > 0$ in some areas of the output space. In this case, $\beta_D(\boldsymbol{x})$ can be used to quantify the extent of uncertainty associated with each consequent element. Correspondingly, $\beta_n(\boldsymbol{x})$ is used to denote the lower bound of the likelihood that the consequent is assessed to D_n. The upper bound of the likelihood is given by $(\beta_n(\boldsymbol{x}) + \beta_D(\boldsymbol{x}))$ [45]. Without loss of generality, we suppose the utilities satisfy the preference constraints $u(D_{n+1}) \geqslant u(D_n)$. Thus, as a supplement to the belief distribution denoted by (7), a BRB prediction interval can also be established to represent the uncertain output, which is characterised by the minimum, maximum and average utilities of $S(y(\boldsymbol{x}))$ defined as follows

$$\hat{y}_{\min}(\boldsymbol{x}) = u(D_1)(\beta_1(\boldsymbol{x}) + \beta_D(\boldsymbol{x})) + \sum_{n=2}^{N} u(D_n)\beta_n(\boldsymbol{x}) \tag{9-a}$$

$$\hat{y}_{\max}(\boldsymbol{x}) = \sum_{n=1}^{N-1} u(D_n)\beta_n(\boldsymbol{x}) + u(D_N)(\beta_N(\boldsymbol{x}) + \beta_D(\boldsymbol{x})) \tag{9-b}$$

$$\hat{y}_{\text{aver}}(\boldsymbol{x}) = \frac{y_{\max}(\boldsymbol{x}) + y_{\min}(\boldsymbol{x})}{2} \tag{9-c}$$

To express the concepts in a more general sense, we call $\hat{y}_{\min}(\boldsymbol{x})$ and $\hat{y}_{\max}(\boldsymbol{x})$ lower and upper bounds, $\hat{y}_{\text{aver}}(\boldsymbol{x})$ middle point of prediction interval respectively thereinafter. It is evident that $\hat{y}(\boldsymbol{x}) = \hat{y}_{\min}(\boldsymbol{x}) = \hat{y}_{\max}(\boldsymbol{x}) = \hat{y}_{\text{aver}}(\boldsymbol{x})$ if $\beta_D(\boldsymbol{x}) = 0$. Otherwise the prediction uncertainty can be quantified by the width of the prediction interval $[\hat{y}_{\min}(\boldsymbol{x}), \hat{y}_{\max}(\boldsymbol{x})]$.

$$\hat{y}_{\max}(\boldsymbol{x}) - \hat{y}_{\min}(\boldsymbol{x}) = (u(D_N) - u(D_1))\beta_D(\boldsymbol{x}) \tag{10}$$

According to the Stone–Weierstrass theorem [38,11], it has been proven that the BRB model with $\sum_{n=1}^{N}\beta_{n,k} = 1; k = 1, \ldots, K$ can approximate any continuous function on a compact set with arbitrary accuracy [7]. The approximation error can be arbitrarily small by adjusting the design parameters of the BRB model. Allowing belief rules to be incomplete further enhances the flexibility of BRB systems. Therefore, it is intuitively reasonable that the prediction interval $[\hat{y}_{\min}(\boldsymbol{x}), \hat{y}_{\max}(\boldsymbol{x})]$ can be used to approximate the lower and upper bounds of the interval outputs of any uncertain nonlinear system. To determine the optimal parameters of the BRB system, training can be performed to minimise the approximation error over the measured datasets.

2.4. Training of BRB models

As introduced above, there are different types of parameters in constructing a BRB identification model. In this paper, we assume that prior knowledge is known regarding model structure, that is to say, the BRB model has fixed number of referential values for both antecedent input variables and output, and the values of J_i; $i = 1, \ldots, M$ and N are given. Thus, the adjustable parameters include $\boldsymbol{P} = \langle A_{i,j}, \beta_{n,k}, \theta_k, \delta_i, u(D_n)\rangle$, namely referential values for each antecedent variable, basic belief degrees, rule weights, antecedent variable weights, and utilities for consequent elements. These parameters can initially be given by experts. However, in some situations, the experts' domain knowledge may not be sufficient or accurate to design identification models. Hence, training methods can be developed to improve the identification accuracy of BRB models. In the training of BRB models, the parameters in terms of physical meanings must satisfy the following linear equality and inequality constraints [45,6]:

(1) A belief degree must not be less than zero or more than one, i.e.,

$$0 \leqslant \beta_{n,k} \leqslant 1;\ n = 1,\ldots,N;\ k = 1,\ldots,K \tag{11-a}$$

(2) The total belief degree of a belief rule must not be more than one, i.e.,

$$\sum_{n=1}^{N}\beta_{n,k} \leqslant 1;\ k = 1,\ldots,K \tag{11-b}$$

Note that if the kth belief rule is complete, its total belief degree must be equal to one. The equality constraint is required for all the complete belief rules.

(3) A rule weight is normalised, so that it is between zero and one, i.e.,

$$0 \leqslant \theta_k \leqslant 1;\ k = 1,\ldots,K \tag{11-c}$$

(4) A weight of any variable is normalised, so that it is between zero and one, i.e.,

$$0 \leqslant \delta_i \leqslant 1;\ i = 1,\ldots,M \tag{11-d}$$

(5) Without loss of generality, suppose x_i is a "profit" variable, so its referential values must satisfy the preference constraints, i.e.,

$$A_{i,j} - A_{i,j+1} \leqslant 0;\ i = 1,\ldots,M;\ j = 1,\ldots,J_i - 1 \tag{11-e}$$

In some cases, two adjacent referential values for an antecedent variable may be convergent to a local area during training process. It means that one of the two referential values can be eliminated in terms of training data, which can therefore reduce the complexity of the belief rule base. Certainly, the training of model structure cannot just simply take into consideration this situation, and rigorous method should also be developed for training model structure in future research, although it is not the main focus of this paper. Chang et al. [9] recently conducted a comparative study of structuring learning for BRB systems using a set of dimensionality reduction approaches.

In order to improve the accuracy of BRB identification models, reliable training models need to be developed for identifying the optimal values of the set of parameters $\boldsymbol{P}^{*}$ from a finite number of training datasets, which are usually a subset of all measured input–output datasets. Through constructing training models, a set of estimated parameters $\widehat{\boldsymbol{P}}$ can be obtained and thus the trained BRB identification model can be re-written as

$$\hat{y}_t = f(\boldsymbol{x}_t;\widehat{\boldsymbol{P}});\quad t = 1,\ldots,T \tag{12}$$

If $\widehat{\boldsymbol{P}}$ is trained to be close enough to $\boldsymbol{P}^{*}$ within a given time complexity, the model output $\hat{y}_t$ can approximate the measured output y_t with a high accuracy. Nevertheless the BRB model may not necessarily match the real system due to various reasons, such as noises, measurement errors, limited number of training datasets, and fixed complexity of identification model. The measurement of accuracy plays an important role in determining the reliability of identification models. In the point prediction of certain systems, the traditional measures of accuracy, such as mean squared error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE), can be used as a global indication of the reliability of an identification model. However, in the interval prediction of BRB identification models, the global accuracy measures may not be directly applicable, and local accuracy should be used measure how the prediction interval can capture uncertain outputs in certain local areas. In the following section, we start with the general BRB identification model using least MSE, and then construct two novel BRB identification models using l_∞-norm or minimising mean uncertainties in belief rules (MUBR) so as to further enhance its capability of capturing uncertain outputs in the identification of uncertain nonlinear systems.

3. Belief rule based identification models

3.1. BRB identification model using least mean squared error

The BRB identification model using least MSE has been widely used in previous research [46,6]. It is assumed that there is no global uncertainty in all belief rules, i.e., $\sum_{n=1}^{N}\beta_{n,k} = 1; k = 1,\ldots,K$, and so $\hat{y}_{\min}(\boldsymbol{x}) = \hat{y}_{\max}(\boldsymbol{x}) = \hat{y}_{\mathrm{aver}}(\boldsymbol{x})$. Thus, the objective of minimising MSE can be simply defined as follows

$$\min_{\boldsymbol{P}}\xi(\boldsymbol{P}) = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{y}_{\mathrm{aver}}(\boldsymbol{x}_t) - y_t\right)^2 \tag{13}$$

subject to

Constraints (11-a)–(11-e).

In the following a simple nonlinear function with different levels of noise is employed to illustrate the capability of the BRB identification model using least MSE.

Example 1. In this example, a simple nonlinear function is used to generate training datasets [10]. The BRB identification model using MSE is used to simulate this nonlinear function without noise and with two different levels of noise intensity.

$$y = \frac{\sin x}{x} + \varepsilon;\quad x \in [0, 10] \tag{14}$$

where ε is Gaussian noise with zero mean, that is, $\varepsilon = N(0, \delta)$; $\delta \in \{0, 0.05, 0.1\}$. To construct a BRB model to simulate this system with uncertain noises, 4 referential values of the input variable x within the interval [0, 10] are initially defined as {0, 3, 7, 10}. As a result, an initial belief rule base can be constructed in Table 1.

Three training datasets with 100 data points each are generated from the nonlinear function without noise and with two different levels of noise to train the initial belief rule base respectively. Fig. 1 shows the outputs of the BRB identification model trained by the three different datasets [8].

It is evident from Fig. 1 that the trained BRB model can effectively identify the dynamics of the nonlinear function without noise or with different levels of noise. Even if the level of noise intensity shown in the third sub-figure is considerably high, the trained BRB model can still keep track of the dynamics of the system output.

3.2. BRB identification model using l_∞-norm

As discussed previously, in real-world systems there always exists various types of uncertainty, the point prediction of the BRB identification model minimising MSE provides no information about the uncertain outputs, especially if the uncertainty is incurred by uncertain parameters or incomplete knowledge. In this

Table 1
Initial belief rule base for Example 1.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.5, 0, 0.5, 1, 1.5\}$, $\delta_1 = 1$
1	1	0	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 1), (D_5, 0)\}$
2	1	3	$\{(D_1, 0), (D_2, 0.9), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$
3	1	7	$\{(D_1, 0), (D_2, 0.8), (D_3, 0.2), (D_4, 0.0), (D_5, 0)\}$
4	1	10	$\{(D_1, 0.1), (D_2, 0.9), (D_3, 0), (D_4, 0), (D_5, 0)\}$

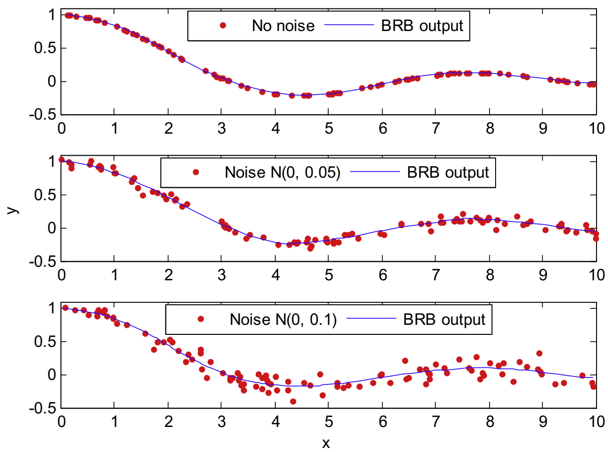

Fig. 1. BRB identification of a nonlinear function with noises.

section, the BRB identification model using l_∞-norm is further presented to construct realistic and reliable BRB prediction intervals. The l_∞-norm is in nature a local accuracy measure and can be used to evaluate the worst-case performance of minimising identification errors on a finite set of measured datasets. The idea of the BRB identification model using l_∞-norm is to find a BRB system satisfying

$$\hat{y}_{\min}(\boldsymbol{x}_t) \leqslant y_t \leqslant \hat{y}_{\max}(\boldsymbol{x}_t) \tag{15}$$

This implies that the output of the uncertain nonlinear system at any sampling time t should always be enclosed by the prediction interval $[\hat{y}_{\min}(\boldsymbol{x}_t), \hat{y}_{\max}(\boldsymbol{x}_t)]$ defined in Eqs. (9-a) and (9-b). On the other hand, the prediction interval $[\hat{y}_{\min}(\boldsymbol{x}_t), \hat{y}_{\max}(\boldsymbol{x}_t)]$ on the whole input space should be as tight as possible under relevant constraints. This can be realised by solving the following *min–max* optimisation problem.

$$\min_{P} \max_{\boldsymbol{x}_t(t=1,\ldots,T)} (|y_t - \hat{y}_{\min}(\boldsymbol{x}_t)|, |\hat{y}_{\max}(\boldsymbol{x}_t) - y_t|) \tag{16}$$

subject to

$$y_t - \hat{y}_{\min}(\boldsymbol{x}_t) \geqslant 0;\ t = 1, \ldots, T \tag{16-a}$$

$$y_t - \hat{y}_{\max}(\boldsymbol{x}_t) \leqslant 0;\ t = 1, \ldots, T \tag{16-b}$$

Constraints (11-a)–(11-e)

Theorem 1. *The min–max optimisation problem (16) can be solved as the nonlinear programming problem of minimising σ subject to the linear constraints 11-a,11-b,11-c,11-d,11-e, the nonlinear constraints (16-a) and (16-b) and the following inequalities.*

$$y_t - u(D_1)(\beta_1(\boldsymbol{x}_t) + \beta_D(\boldsymbol{x}_t)) - \sum_{n=2}^{N} u(D_n)\beta_n(\boldsymbol{x}_t) \leqslant \sigma;\ t = 1, \ldots, T \tag{17-a}$$

$$-y_t + \sum_{n=1}^{N-1} u(D_n)\beta_n(\boldsymbol{x}_t) + u(D_N)(\beta_N(\boldsymbol{x}_t) + \beta_D(\boldsymbol{x}_t)) \leqslant \sigma;\ t = 1, \ldots, T \tag{17-b}$$

$$\sigma \geqslant 0 \tag{17-c}$$

The resulting σ denotes the maximum identification error in the whole sampling input space.

Proof. If we define

$$\sigma = \max_{\boldsymbol{x}_t(t=1,\ldots,T)} (|y_t - \hat{y}_{\min}(\boldsymbol{x}_t)|, |\hat{y}_{\max}(\boldsymbol{x}_t) - y_t|) \tag{18}$$

This directly implies the following inequalities.

$$\left| y_t - u(D_1)(\beta_1(\boldsymbol{x}_t) + \beta_D(\boldsymbol{x}_t)) - \sum_{n=2}^{N} u(D_n)\beta_n(\boldsymbol{x}_t) \right| \leqslant \sigma \tag{19-a}$$

$$\left| \sum_{n=1}^{N-1} u(D_n)\beta_n(\boldsymbol{x}_t) + u(D_N)(\beta_N(\boldsymbol{x}_t) + \beta_D(\boldsymbol{x}_t)) - y_t \right| \leqslant \sigma \tag{19-b}$$

which can be transformed to the Eqs. (17-a)–(17-c). Thus the *min–max* optimisation problem in Eq. (16) can be stated as the minimisation of σ subject to the linear constraints (11-a)–(11-e), and (17-c) and the nonlinear inequalities (17-a) and (17-b). These constraints guarantee that the BRB prediction interval encloses the uncertain outputs of all training datasets.

To test the identification model, we consider a nonlinear function with uncertain but bounded parameters. The uncertain outputs for the same input vector can be obtained by the repeated measurements of the outputs of the systems with different parameters. □

Example 2. There is a class of nonlinear functions $\boldsymbol{G} = \{g(x) = g_{\text{norm}}(x) + \Delta g(x)\}$ with the nominal function $g_{\text{norm}}(x) = \cos(x)\sin(x)$ and the uncertainty part $\Delta g(x) = \gamma\cos(8x)$; $0 \leqslant \gamma \leqslant 0.2$. The function from the class is defined in the input domain $U = \{x| -1 \leqslant x \leqslant 1\}$ and the set of measured input set is $X = \{\boldsymbol{x}_t | \boldsymbol{x}_t = 0.021t;\ t = -47, -46, \ldots, 47\} \subset U$ [35].

In this example, the dimensionality of the input space is $M = 1$. To construct a BRB model, 5 initial referential values for the input variable x are uniformly selected from the interval $[-1, 1]$, and they are $\{-1, -0.5, 0, 0.5, 1\}$. The outputs of the nonlinear function with the parameter $\gamma = 0$, 0.05, 0.1, 0.15 and 0.2 are given in the following matrix.

$$\begin{bmatrix} -0.4546, -0.4619, -0.4692, -0.4765, -0.4837 \\ -0.4207, -0.4534, -0.4861, -0.5188, -0.5515 \\ 0.0000, 0.0500, 0.0100, 0.1500, 0.2000 \\ 0.4207, 0.3881, 0.3554, 0.3227, 0.2900 \\ 0.4546, 0.4474, 0.4401, 0.4328, 0.4255 \end{bmatrix} \tag{20}$$

In the matrix, each row represents the uncertain outputs under a given referential input. Through observing the measured outputs on these referential inputs, the referential utilities $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.8, -0.4, 0, 0.4, 0.8\}$ are defined for measuring the consequent elements of belief rules. As a result, a set of initial belief rules can be constructed as listed in Table 2.

In Table 2, the sum of the belief degree on each individual consequent element D_n is less than 1, and the belief degree on the whole set of consequent elements $\boldsymbol{D}$ is not equal to 0 as usual. This implies that the output of the nonlinear system represented by the belief structures is uncertain. To obtain a prediction interval to enclose all the uncertain outputs of the nonlinear system, the measured input–output dataset above can be used to train the initial belief rules, and the training model can be constructed as discussed above. Using the nonlinear optimisation solver *fmincon* in the Optimisation Toolbox of MATLAB, the simulation results are presented in Fig. 2.

In Fig. 2, the dotted lines are the outputs of the nonlinear function with the above 5 different parameters, and the solid lines are the lower and upper bounds of the uncertain outputs respectively, which are identified by the trained BRB identification model. It is obvious from Fig. 2 that all uncertain output are enclosed in the close BRB prediction intervals, and the trained BRB identification model can generate relatively tight bounds around the family of the nonlinear functions. In the trained BRB models, the referential values of the input variable x are updated to be $\{-1.0000, -0.4521, 0.3202, 0.7118, 1.0000\}$. Table 3 lists the updated parameters in the trained belief rules.

It can be seen in Table 3 that all parameters in the trained belief rule base have been updated significantly, including the remaining belief degree in each rule $\beta_{D,k}$ ($k = 1, \ldots, 5$). In this example, the

Table 2
Initial belief rule base for Example 2.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.8, -0.4, 0, 0.4, 0.8\}$, $\delta_1 = 1$
1	1	−1	$\{(D_1, 0.2), (D_2, 0.7), (D_3, 0), (D_4, 0), (D_5, 0), (\boldsymbol{D}, 0.1)\}$
2	1	−0.5	$\{(D_1, 0.3), (D_2, 0.5), (D_3, 0), (D_4, 0), (D_5, 0), (\boldsymbol{D}, 0.2)\}$
3	1	0	$\{(D_1, 0), (D_2, 0), (D_3, 0.5), (D_4, 0.4), (D_5, 0), (\boldsymbol{D}, 0.1)\}$
4	1	0.5	$\{(D_1, 0), (D_2, 0), (D_3, 0.1), (D_4, 0.6), (D_5, 0.1), (\boldsymbol{D}, 0.2)\}$
5	1	1	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 0.8), (D_5, 0.1), (\boldsymbol{D}, 0.1)\}$

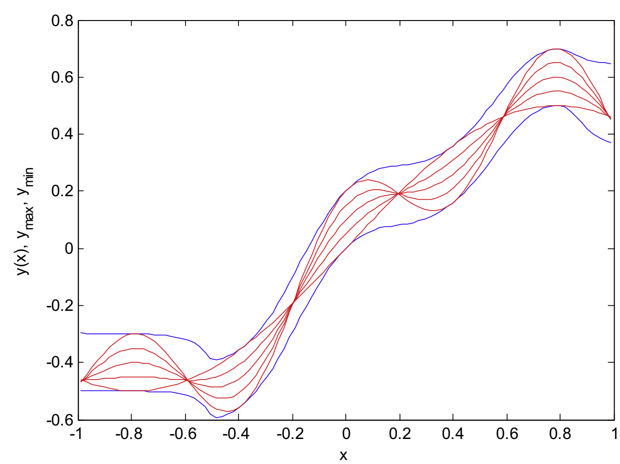

Fig. 2. Lower and upper bounds of BRB prediction interval.

results show that the BRB identification model using l_∞-norm is appropriate and effective in the identification of the uncertainties for the family of nonlinear functions. Even better identification performance can be obtained if we increase the number of referential values for the input variable according to the universal approximation property of BRB systems [7]. With simulations, it is worth noting that the maximum approximation error $\sigma = 0.2005$ with the above 5 referential values for the input variable, and $\sigma = 0.2000$ while increasing to 11 referential values. The maximum identification error does not change too much with the significant increase on the number of referential values, because the local accuracy measure using l_∞-norm is conservative. As shown in Fig. 2, although the interval of the real uncertain outputs in some areas is quite small, the BRB identification model using l_∞-norm only takes into consideration the maximum interval, for example, at around $x = -0.8$. This loose accuracy measure generates conservativeness in the sense that all uncertain outputs are enclosed by the BRB prediction interval completely, but the model is also less reliable in some areas, for example, at around $x = -0.6$. In the following section, a novel BRB identification model is proposed through fully exploiting the characteristics of the BRB methodology.

3.3. BRB identification model minimising mean uncertainties in belief rules

From Eq. (10), the width of the BRB prediction interval is essentially determined by the remaining belief degree $\beta_{\boldsymbol{D}}(\boldsymbol{x})$, given that $u(D_1)$ and $u(D_N)$ are two end-point referential values. In terms of the inference analysis of the BRB methodology [6], it is intuitively true that $\beta_{\boldsymbol{D}}(\boldsymbol{x})$ is positively related to the uncertain part $1 - \sum_{n=1}^{N} \beta_{n,k}; k = 1, \ldots, K$.

Theorem 2. *The remaining belief degree $\beta_D(\boldsymbol{x})$ is not positively related to the basic belief degree $\beta_{n,k}$, i.e., $\partial\beta_D(\boldsymbol{x})/\partial\beta_{n,k} \leqslant 0$.*

Proof. It has been proved the combined belief degree $\beta_n(\boldsymbol{x})$ increases monotonically with the increase of the basic belief degree $\beta_{n,k}$ in the ER inference process [6]. Mathematically, the first-order derivative $\partial\beta_n(\boldsymbol{x})/\partial\beta_{n,k} \geqslant 0$. Similarly, we can prove $\partial\beta_D(\boldsymbol{x})/\partial\beta_{n,k} \leqslant 0$ (see Appendix A).

Theorem 2 implies that $\beta_{\boldsymbol{D}}(\boldsymbol{x})$ is monotonically decreasing as the basic belief degree $\beta_{n,k}$ increases. In other words, reducing the uncertain part $1 - \sum_{n=1}^{N} \beta_{n,k}; k = 1, \ldots, K$ can reduce $\beta_{\boldsymbol{D}}(\boldsymbol{x})$ which further minimises the width of the BRB prediction interval. Therefore, the tightness of the BRB prediction intervals over the whole input space can be achieved by minimising mean uncertainties in belief rules (MUBR) which is defined as follows

$$\min_{\boldsymbol{P}}(\boldsymbol{P}) = \frac{1}{K}\sum_{k=1}^{K}\left(1 - \sum_{n=1}^{N}\beta_{n,k}\right) \tag{21}$$

subject to

Constraints (11-a)–(11-e), (16-a) and (16-b).

To illustrate the capability of the BRB identification model minimising MUBR, a simple nonlinear function with non-uniform output noises is considered in the following example. □

Example 3. A nonlinear function $y = 0.5\sin(1.5\pi x + \pi/2) + 2.0 + v$ [23], where x is the independent variable, y is the dependent variable and v is Gaussian noise with the standard deviation of $\sigma_v = 0.045 + 0.04x$ (non-uniform output noise distribution).

Firstly, an initial belief rule base can be constructed in Table 4. Actually, no prior knowledge about the uncertain part has been assumed in constructing each belief rule.

A training dataset with 200 data points are generated from the nonlinear function with non-uniform Gaussian noise. Fig. 3(a) and (b) show the prediction intervals trained from the BRB identification model using l_∞-norm and minimising MUBU respectively.

With the same initial belief rule base and parameter settings, it can be seen in Fig. 3(a) that the prediction interval does not change significantly, as the noise gets higher and higher as the value of the input variable x increases. In other words, the identification model using l_∞-norm provides unreliable prediction interval in some areas where both the input variable and the degree of uncertainty are low. However, as shown in Fig. 3(b), the identification model minimising MUBR can identify the levels of uncertainty at different local areas accurately. In addition, all data points lie within the prediction interval. Table 5(a) and (b) list the trained belief rule base using l_∞-norm and minimising MUBR for Example 3 respectively.

It can also be observed that there is no significant difference on the uncertain part $\beta_{\boldsymbol{D},k}$ $(k = 1, \ldots, 5)$ in the trained belief rule base using l_∞-norm. On contrast, the uncertain part $\beta_{\boldsymbol{D},k}$ $(k = 1, \ldots, 5)$ in

Table 3
Trained belief rule base for Example 2.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{-0.8, -0.5500, 0.1654, 0.4395, 0.8\}$, $\delta_1 = 0.9959$
1	1.0000	−1.0000	$\{(D_1, 0.0000), (D_2, 0.8112), (D_3, 0.0000), (D_4, 0.0000), (D_5, 0.0614), (\boldsymbol{D}, 0.1274)\}$
2	0.2500	−0.4521	$\{(D_1, 0.7354), (D_2, 0.0000), (D_3, 0.0000), (D_4, 0.0000), (D_5, 0.1372), (\boldsymbol{D}, 0.1274)\}$
3	0.4252	0.3202	$\{(D_1, 0.0000), (D_2, 0.0000), (D_3, 0.5748), (D_4, 0.2946), (D_5, 0.0000), (\boldsymbol{D}, 0.1306)\}$
4	0.4614	0.7118	$\{(D_1, 0.0000), (D_2, 0.0019), (D_3, 0.1641), (D_4, 0.0000), (D_5, 0.6900), (\boldsymbol{D}, 0.1440)\}$
5	0.6293	1.0000	$\{(D_1, 0.0589), (D_2, 0.0407), (D_3, 0.0000), (D_4, 0.0044), (D_5, 0.7069), (\boldsymbol{D}, 0.1891)\}$

Table 4
Initial belief rule base for Example 3.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{1, 1.5, 2, 2.5, 3\}$, $\delta_1 = 1$
1	1	−1	$\{(D_1, 0), (D_2, 0), (D_3, 1), (D_4, 0), (D_5, 0)\}$
2	1	−0.5	$\{(D_1, 0), (D_2, 0.9), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$
3	1	0	$\{(D_1, 0), (D_2, 0), (D_3, 0), (D_4, 1), (D_5, 0)\}$
4	1	0.5	$\{(D_1, 0), (D_2, 0.9), (D_3, 0.1), (D_4, 0), (D_5, 0)\}$
5	1	1	$\{(D_1, 0), (D_2, 0), (D_3, 1), (D_4, 0), (D_5, 0)\}$

the trained belief rule base using minimising MUBR increases from a low level to a relatively high level, which completely reflects the level of uncertainty in the real system.

3.4. Discussion and analysis of identification accuracy and interval credibility

Although the BRB identification model minimising MUBR provide a very good fit for the nonlinear system with changing level of uncertainty, the three identification models discussed above can be applied in different scenarios. In terms of the numerical studies, the model using least MSE shows a good capability of dealing with traditional point prediction of certain systems, or identifying the dynamics of nonlinear systems with noises. The model using l_∞-norm can identify the maximum uncertainty with the *min–max* optimisation technique which is quite straightforward to understand in the identification of uncertain nonlinear systems. In both the model using l_∞-norm and minimising MUBR, the prediction interval can reflect the uncertainty of outputs incurred by whatever sources of uncertainty in a unified way, and the identification results represented by belief structure are also informative. The prediction intervals could be beneficial to the practical user on many aspects, such as (1) taking into account experts' subjective knowledge in the identification of uncertain nonlinear systems; (2) providing a clear picture about the level of uncertainties at different local areas; (3) designing different strategies in terms of the lower and upper bounds of uncertain outputs, for example, designing different rules of activating alarms in fault diagnosis. In the BRB identification models, the parameter constraints (11-a)–(11-e) are linear, and the boundary constraints (16-a) and (16-b) are nonlinear. The objective functions apart from the maximum identification error in the transformed *min–max* training model are nonlinear. The nonlinear BRB identification models can be easily solved by typical nonlinear optimisation algorithms, such as the nonlinear optimisation solver *fmincon* in the Optimisation Toolbox of MATLAB.

Given the fact that training is vital for improving the performance of any identification model, the model accuracy may depend heavily on the quality of training datasets. In some practical systems, there may exist some error data or unusual measured data (e.g., outliers due to erroneous measurements, noisy data from the heavy tails of noise distribution). However in the model using l_∞-norm and minimising MUBR, the constraints (16-a) and (16-b) specify that the prediction interval must envelop all training data points. Any error data or unusual measured data, such as outliers, may deteriorate the identification accuracy as shown in Fig. 4.

Although the outlier only affect the identification accuracy in a local area, given that the referential values decompose the input space and the BRB model is in essence a multi-model approximator [7], the large prediction interval in the local area $[-1, -0.6]$ does not tell useful information. There are a few possible ways to cope with this situation:

(1) Eliminate outliers from training datasets through data preprocessing. There are a wide range of outlier detection methodologies [19].
(2) Make a trade-off between identification accuracy and interval credibility. It is completely credible that the prediction interval generated by the BRB identification model with the constraints (16-a) and (16-a) envelop all data points. But, in order to improve the poor identification accuracy in some local areas as shown in Fig. 4, trade-off needs to be made between identification accuracy and interval credibility. For example, we just make sure that a certain percentage of training data points lies within the prediction interval. The constraints (16-a) and (16-a) can be replaced by setting a minimum value γ to the prediction interval coverage probability (PICP) [21], which can be calculated in a statistical way.

$$\gamma \leqslant PICP = \frac{1}{T}\sum_{t=1}^{T} c_t \tag{22}$$

where $c_t = 1$, if $y_t \in [\hat{y}_{\min}(\boldsymbol{x}), \hat{y}_{\max}(\boldsymbol{x})]$; otherwise, $c_t = 0$. As discussed above, PICP is calculated as the proportion of data points which lie within the prediction interval.

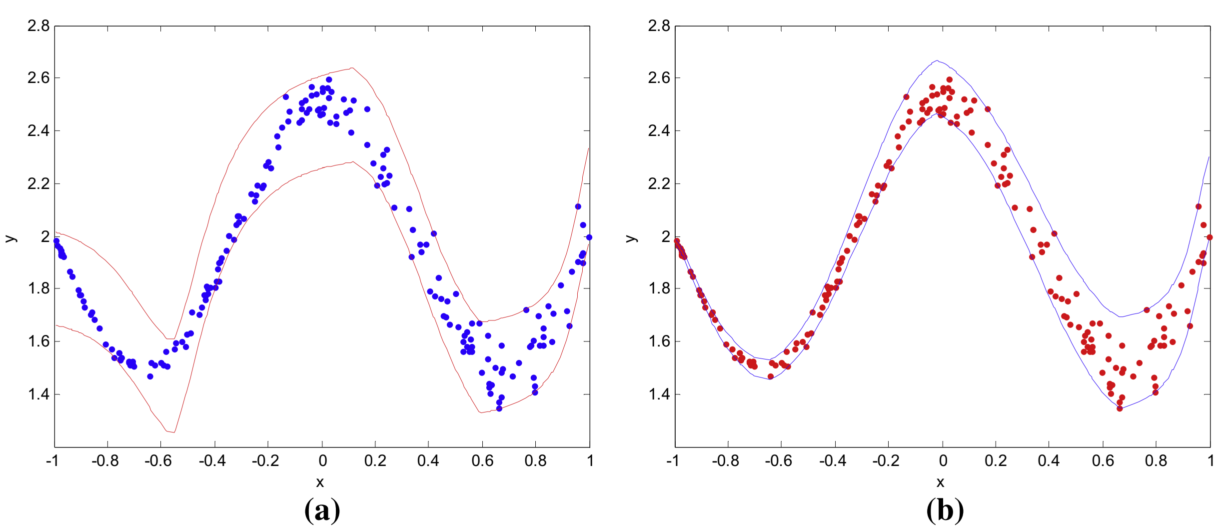

Fig. 3. Comparison of BRB identification models using l_∞-norm and minimising MUBR.

Table 5
(a) Trained belief rule base using l_∞-norm for Example 3.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{1, 1.2500, 1.7910, 2.7289, 3\}$, $\delta_1 = 0.7745$
1	0.9561	−1.0000	$\{(D_1, 0.0317), (D_2, 0.0098), (D_3, 0.6030), (D_4, 0.0195), (D_5, 0.0835), (\boldsymbol{D}, 0.2525)\}$
2	0.9002	−0.7229	$\{(D_1, 0.0194), (D_2, 0.7376), (D_3, 0.0119), (D_4, 0.0465), (D_5, 0.0157), (\boldsymbol{D}, 0.1690)\}$
3	1.0000	0.0659	$\{(D_1, 0.0134), (D_2, 0.0000), (D_3, 0.0523), (D_4, 0.6824), (D_5, 0.0725), (\boldsymbol{D}, 0.1794)\}$
4	0.8808	0.5914	$\{(D_1, 0.0640), (D_2, 0.6758), (D_3, 0.0000), (D_4, 0.0591), (D_5, 0.0286), (\boldsymbol{D}, 0.1725)\}$
5	0.2500	1.0000	$\{(D_1, 0.0000), (D_2, 0.0000), (D_3, 0.5322), (D_4, 0.0000), (D_5, 0.2858), (\boldsymbol{D}, 0.1821)\}$

(b) Trained belief rule base minimising MUBR for Example 3

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{1, 1.7617, 2.0601, 2.7500, 3\}$, $\delta_1 = 0.7915$
1	0.4900	−1.0000	$\{(D_1, 0.0000), (D_2, 0.8070), (D_3, 0.0000), (D_4, 0.0000), (D_5, 0.1822), (\boldsymbol{D}, 0.0109)\}$
2	0.8101	−0.6675	$\{(D_1, 0.3764), (D_2, 0.5859), (D_3, 0.0000), (D_4, 0.0000), (D_5, 0.0000), (\boldsymbol{D}, 0.0377)\}$
3	0.7604	−0.0201	$\{(D_1, 0.0000), (D_2, 0.0000), (D_3, 0.2933), (D_4, 0.2257), (D_5, 0.3792), (\boldsymbol{D}, 0.1018)\}$
4	0.7257	0.6629	$\{(D_1, 0.3740), (D_2, 0.4535), (D_3, 0.0000), (D_4, 0.0000), (D_5, 0.0000), (\boldsymbol{D}, 0.1725)\}$
5	0.2528	1.0000	$\{(D_1, 0.0000), (D_2, 0.0000), (D_3, 0.6820), (D_4, 0.1552), (D_5, 0.0023), (\boldsymbol{D}, 0.1605)\}$

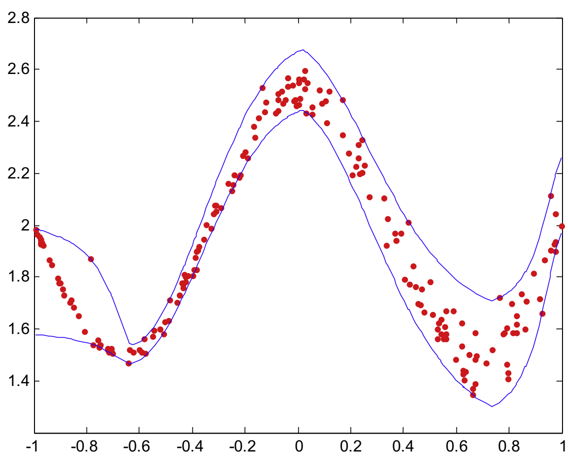

Fig. 4. BRB prediction interval deteriorated by a outlier.

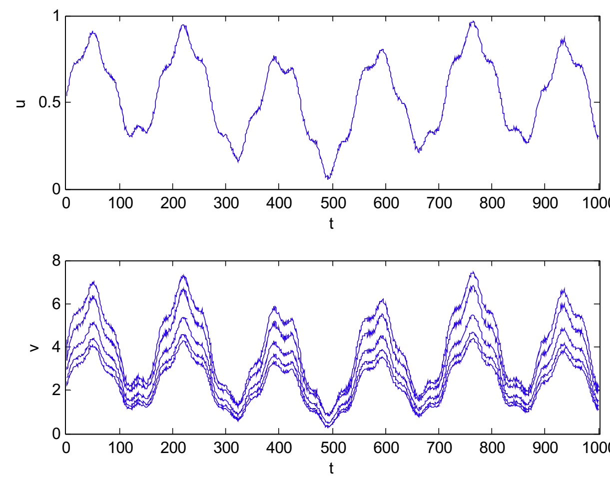

Fig. 5. Simulated dataset on the simplified car dynamics with uncertain parameters.

4. A numerical study of a simplified car dynamics

In this section, the proposed BRB identification model using l_∞-norm is applied to a nonlinear time-invariant system with uncertain physical parameters, which are given as intervals between minimal value and maximal value, e.g., $\tilde{a} \in [\underline{a}, \bar{a}]$. The observed nonlinear system is a simplified car dynamics with uncertain parameters of the engine force. The following mathematical model is used to describe the dynamics [36].

$$\dot{v} = \frac{f_e(u, v) - f_d}{m} \tag{23-a}$$

$$f_e(u, v) = \widetilde{K}_e(1 + a_1 u) \times (1 + \arctan(\tilde{a}_2 u^2 + a_3 v + a_4)) \tag{23-b}$$

where v denotes the velocity of the car in m/s, u is the position of the throttle in the interval [0, 1], the mass of the car m is equal to 1000 kg, $f_e(u,v)$ represents the force of the engine and the resistance force f_d is approximated to be 1000 N. Certain engine parameters include $a_1 = 3$, $a_3 = -0.35$ and $a_4 = 1.2$. Due to the operating conditions, some of the engine parameters are uncertain and vary within certain ranges, and they are $\widetilde{K}_e \in [600, 900]$ N and $\tilde{a}_2 \in [4.2, 7.8]$. It is obvious from Eqs. (23-a) and (23-b) that the characteristic of the engine is highly nonlinear.

To obtain a data set for the identification of the BRB model, five different responses of the car dynamics are simulated with the same input signal. In the five simulated responses, the sets of uncertain parameters include a nominal one with $\widetilde{K}_e = 700$ and $\tilde{a}_2 = 6$ and four combinations of the maximal and minimal values of both interval parameters. In this example, a linear combination of sinusoids with random noises is used to generate the input signal.

$$u(t) = 0.5 + 0.05\sin(0.15t) + 0.3\sin(0.035t) + 0.1\sin(0.01t) + \text{noise}(0, 0.025) \tag{24}$$

where noise(0, 0.025) denotes a random noise between 0 and 0.025. Fig. 5 shows that the simulated input signal and the uncertain responses of the car simplified dynamics with uncertain parameters.

It is evident from Fig. 5 that the simulated time responses of the car dynamics exhibit highly nonlinear and uncertain patterns. For identifying this simplified car dynamics using a BRB model, 5 initial referential values for the input variable u are uniformly selected from the interval [0, 1], and they are {0, 0.25, 0.5, 0.75, 1}. Based on the simulated dataset, the following referential utilities are defined for measuring the consequent elements of belief rules.

Table 6
Initial belief rule base for the simplified car dynamics.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{0, 2, 4, 6, 8\}$, $\delta_1 = 1$
1	1	0	$\{(D_1, 0.6), (D_2, 0.2), (D_3, 0), (D_4, 0), (D_5, 0), (\boldsymbol{D}, 0.2)\}$
2	1	0.25	$\{(D_1, 0.2), (D_2, 0.7), (D_3, 0), (D_4, 0), (D_5, 0), (\boldsymbol{D}, 0.1)\}$
3	1	0.5	$\{(D_1, 0), (D_2, 0.4), (D_3, 0.4), (D_4, 0), (D_5, 0), (\boldsymbol{D}, 0.2)\}$
4	1	0.75	$\{(D_1, 0), (D_2, 0.3), (D_3, 0.3), (D_4, 0.3), (D_5, 0), (\boldsymbol{D}, 0.1)\}$
5	1	1	$\{(D_1, 0), (D_2, 0), (D_3, 0.2), (D_4, 0.5), (D_5, 0.2), (\boldsymbol{D}, 0.1)\}$

Table 7
Trained belief rule base for the simplified car dynamics.

R_k	θ_k	$x(A_j)$	Consequents $\{D_1, D_2, D_3, D_4, D_5\} = \{0, 1.7548, 4.1155, 6.3797, 8\}$, $\delta_1 = 0.6975$
1	0.9993	0.0000	$\{(D_1, 0.7558), (D_2, 0.0082), (D_3, 0.0049), (D_4, 0.0007), (D_5, 0.0012), (\boldsymbol{D}, 0.2292)\}$
2	0.9815	0.4231	$\{(D_1, 0.4492), (D_2, 0.0592), (D_3, 0.0907), (D_4, 0.0321), (D_5, 0.0688), (\boldsymbol{D}, 0.3000)\}$
3	0.9776	0.5581	$\{(D_1, 0.0825), (D_2, 0.3103), (D_3, 0.1731), (D_4, 0.0596), (D_5, 0.0745), (\boldsymbol{D}, 0.3000)\}$
4	0.9760	0.6084	$\{(D_1, 0.2415), (D_2, 0.0963), (D_3, 0.1133), (D_4, 0.1392), (D_5, 0.0003), (\boldsymbol{D}, 0.4096)\}$
5	1.0000	1.0000	$\{(D_1, 0.0174), (D_2, 0.0004), (D_3, 0.0214), (D_4, 0.0331), (D_5, 0.5312), (\boldsymbol{D}, 0.3965)\}$

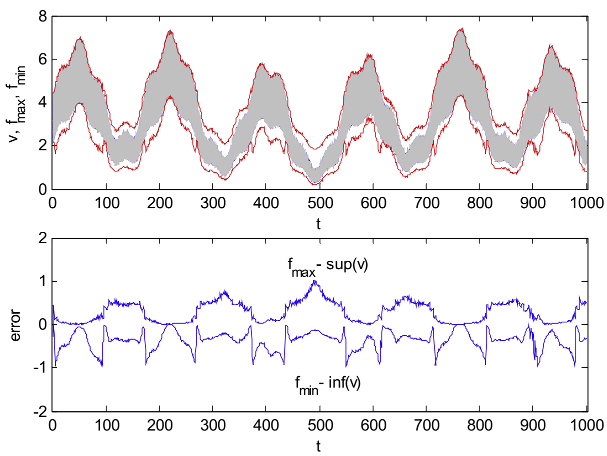

Fig. 6. Lower, upper bounds and identification errors of the prediction interval.

$$\{D_1, D_2, D_3, D_4, D_5\} = \{0, 2, 4, 6, 8\}$$

Given the above definition of the referential values for the input variable and the referential utilities for consequents, a set of initial belief rules can be constructed as listed in Table 6. Here, the remaining belief degree $\beta_{\boldsymbol{D},k}$ in each rule is initialized in an empirical way.

As discussed in Section 2.4, the initial parameters need to be trained in order to minimise the identification errors. Table 7 lists the updated parameters in the trained belief rules for this simplified car dynamics.

In the trained BRB identification model, the weight of the input variable u is updated to be $\delta = 0.6975$. One can see from Table 7 that the belief degrees on the whole set of consequent elements $\boldsymbol{D}$ is large. This also implies that the identification system is highly uncertain as discussed above. Fig. 6 shows the lower and upper bounds generated by the trained BRB identification model and corresponding identification errors.

In the upper subplot of Fig. 6, the grey area represents the domain of the uncertain outputs of the simplified car dynamics, and the lines show the lower and upper bounds identified by the trained BRB identification model. It is evident that the trained BRB identification model can successfully track and enclose the uncertain outputs. The bottom subplot of Fig. 6 shows the identification errors between the approximation bounds of the trained BRB model and the real bounds of the simulated dataset. Note that $\hat{y}_{\max} - \sup(v)$ is always positive and $\hat{y}_{\min} - \inf(v)$ is always negative, which means that the measured dataset actually lies within the lower and upper bounds identified by the BRB model.

5. Conclusion

In this paper, one novel BRB identification model along with two existing ones are constructed for the identification of uncertain nonlinear systems, through exploiting the inherent strengths of the BRB methodology dealing with various types of uncertainty. In the identification models, the MSE, l_∞-norm and MUBR are used to measure the optimality of the identification performance on a finite set of measured data respectively. Numerical studies and theoretical analyses have demonstrated that especially the BRB identification model using l_∞-norm and minimising MUBR have remarkable capabilities of capturing the lower and upper bounds of the interval outputs of uncertain nonlinear systems simultaneously. The identification models can be established without requiring a prior knowledge of the underlying dynamics of the uncertain nonlinear system, including the bound of noises and the bounds of uncertain parameters. Trade-off analysis between identification accuracy and interval credibility is briefly discussed. Finally, a numerical study of a simplified car dynamics has been conducted to demonstrate the capability and effectiveness of the BRB identification models for the modelling and identification of uncertain nonlinear systems.

Acknowledgements

The research is partially supported by the UK EPSRC under Grant No.: EP/F024606/1, the European Commission FP7 Marie Curie IRSES – REFERENCE project, and the Natural Science Foundation of China under Grant No.: 61203178, 71471151. The authors would like to thank the editor and the anonymous referees for the comments and suggestions which have substantially improved the quality of the paper.

Appendix A

The remaining belief degree $\beta_D(\boldsymbol{x})$ can be calculated by Eq. (6) in Section 2.2. To calculate the first-order partial derivative $\partial\beta_D(\boldsymbol{x})/\partial\beta_{n,k}$, let us first define some intermediate variables [6].

$$X_j = \prod_{k=1}^{K}\left(w_k(\boldsymbol{x})\beta_{j,k} + 1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right)$$

$$Y = \prod_{k=1}^{K}\left(1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}\right)$$

$$Z = \prod_{k=1}^{K}(1 - w_k(\boldsymbol{x}))$$

Thus the Eq. (6) can be re-written as

$$\beta_D(\boldsymbol{x}) = \frac{Y - Z}{\sum_{j=1}^{N} X_j - (N-1)Y - Z}$$

We first calculate the first-order partial derivatives of the intermediate variables with respect to the basic belief degree $\beta_{n,k}$.

$$\frac{\partial X_j}{\partial \beta_{n,k}} = \begin{cases} 0 = a_n X_j; \ j = n \\ \frac{-w_k(\boldsymbol{x})}{w_k(\boldsymbol{x})\beta_{j,k} + 1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}} X_j = a_j X_j; \ j \neq n \end{cases}$$

$$\frac{\partial Y}{\partial \beta_{n,k}} = \frac{-w_k(\boldsymbol{x})}{1 - w_k(\boldsymbol{x})\sum_{i=1}^{N}\beta_{i,k}} Y = bY$$

$$\frac{\partial Z}{\partial \beta_{n,k}} = 0$$

The first-order partial derivative $\partial\beta_D(\boldsymbol{x})/\partial\ \beta_{n,k}$ can be calculated as follows

$$\frac{\partial\beta_D(\boldsymbol{x})}{\partial\beta_{n,k}}=\frac{1}{\mu^2}\left[\frac{\partial(Y-Z)}{\partial\beta_{n,k}}\left(\sum_{j=1}^{N}X_j-(N-1)Y-Z\right)-(Y-Z)\frac{\partial\left(\sum_{j=1}^{N}X_j-(N-1)Y-Z\right)}{\partial\beta_{n,k}}\right]$$
$$=\frac{1}{\mu^2}\left[\frac{\partial Y}{\partial\beta_{n,k}}\sum_{j=1}^{N}X_j-\frac{\partial Y}{\partial\beta_{n,k}}(N-1)Y-\frac{\partial Y}{\partial\beta_{n,k}}Z-Y\sum_{j=1}^{N}\frac{\partial X_j}{\partial\beta_{n,k}}+Y(N-1)\frac{\partial Y}{\partial\beta_{n,k}}\right.$$
$$\left.+Z\sum_{j=1}^{N}\frac{\partial X_j}{\partial\beta_{n,k}}-Z(N-1)\frac{\partial Y}{\partial\beta_{n,k}}\right]$$
$$=\frac{1}{\mu^2}\left[bY\sum_{j=1}^{N}X_j-Y\sum_{j=1}^{N}a_jX_j+Z\sum_{j=1}^{N}a_jX_j-ZNbY\right]$$
$$=\frac{1}{\mu^2}\left[Y\sum_{j=1}^{N}(b-a_j)X_j+Z\sum_{j=1}^{N}(a_jX_j-bY)\right]$$
$$=\frac{1}{\mu^2}\left[Y\sum_{j=1,j\neq n}^{N}(b-a_j)X_j+Z\sum_{j=1,j\neq n}^{N}(a_jX_j-bY)+bY(X_n-Z)\right]$$

According to the mathematical definition of parameters, it is evident that $X_j \geqslant 0; j = 1, \ldots, N$, $Y \geqslant 0$ and $Z \geqslant 0$. We can further derive

$$b-a_j\leqslant 0; j=1,\ldots,N$$
$$a_jX_j-bY\leqslant 0; j=1,\ldots,N \text{ and } j\neq n$$
$$bY(X_n-Z)\leqslant 0$$

Thus we have $\partial\beta_D(\boldsymbol{x})/\partial\beta_{n,k} \leqslant 0$.

References

[1] J. Ackermann, Robust Control: Systems with Uncertain Physical Parameters, Springer-Verlag, Berlin, Heidelberg, 1993.
[2] J. Armengol, J. Vehí, M.Á. Sainz, P. Herrero, E.R. Gelso, SQualTrack: a tool for robust fault detection, IEEE Trans. Syst. Man Cybernet.—Part B: Cybernet. 39 (2) (2009) 475–488.
[3] J.C. Augusto, J. Liu, P.J. McCullagh, H. Wang, J.B. Yang, Management of uncertainty and spatio-temporal aspects for monitoring and diagnosis in a Smart Home, Int. J. Comput. Intell. Syst. 1 (4) (2008) 361–378.
[4] S.I. Biagiola, J.L. Figueroa, Wiener and Hammerstein uncertain models identification, Math. Comput. Simul. 79 (2009) 3296–3313.
[5] M.C. Campi, E. Weyer, Guaranteed non-asymptotic confidence regions in system identification, Automatica 41 (2005) 1751–1764.
[6] Y.W. Chen, J.B. Yang, D.L. Xu, Z.J. Zhou, D.W. Tang, Inference analysis and adaptive training for belief rule based systems, Expert Syst. Appl. 36 (10) (2011) 12845–12860.
[7] Y.W. Chen, J.B. Yang, D.L. Xu, S.L. Yang, On the inference and approximation properties of belief rule based systems, Inform. Sci. 234 (2013) 121–135.
[8] Y.W. Chen, J.B. Yang, D.L. Xu, Uncertain Nonlinear System Modeling and Identification Using Belief Rule-Based Systems, Lecture Notes in Computer Science (in Proceedings of IUKM 2013, Beijing, China, July 12–14, 2013), vol. 8032, 2013, pp. 209–218.
[9] L. Chang, Y. Zhou, J. Jiang, M. Li, X. Zhang, Structure learning for belief rule base expert system: a comparative study, Knowl.-Based Syst. 39 (2013) 159–172.
[10] C.C. Chuang, Z.J. Lee, Hybrid robust support vector machines for regression with outliers, Appl. Soft Comput. 11 (2011) 64–67.
[11] E.W. Cheney, Introduction to Approximation Theory, AMS Chelsea Publishing, 1998.
[12] H.H. Choi, Adaptive controller design for uncertain fuzzy systems using variable structure control approach, Automatica 45 (2009) 2646–2650.
[13] R. Davis, Knowledge-based systems, Science 231 (4741) (1986) 957–963.
[14] A.P. Dempster, A generalization of Bayesian inference, J. Roy. Stat. Soc., Ser. B 30 (2) (1968) 205–247.
[15] T. Denoeux, Z. Younes, F. Abdallah, Representing uncertainty on set-valued variables using belief functions, Artif. Intell. 174 (2010) 479–499.
[16] P. Giovanni, I. Lurdes, Decision Theory: Principles and Approaches, John Wiley & Sons, 2009.
[17] M.A. González-Olvera, Y. Tang, Black-box identification of a class of nonlinear systems by a recurrent neurofuzzy network, IEEE Trans. Neural Netw. 21 (4) (2010) 672–679.
[18] F. Hayes-Roth, Rule-based systems, Commun. ACM 28 (9) (1985) 921–932.
[19] V.J. Hodge, J. Austin, A survey of outlier detection methodologies, Artif. Intell. Rev. 22 (2004) 85–126.
[20] R. Isermann, M. Münchhof, Identification of Dynamical Systems: An Introduction with Applications, Springer-Verlag, Berlin, Heidelberg, 2011.
[21] A. Khosravi, S. Nahavandi, D. Creighton, A.F. Atiya, Lower upper bound estimation method for construction of neural network-based prediction intervals, IEEE Trans. Neural Netw. 22 (3) (2011) 337–346.
[22] J. Lawry, Y.C. Tang, Uncertainty modelling for vague concepts: a prototype theory approach, Artif. Intell. 173 (2009) 1539–1558.
[23] J.A. Leonard, M.A. Kramer, L.H. Ungar, A neural network architecture that computes its own reliability, Comput. Chem. Eng. 16 (9) (1992) 819–835.
[24] A. Ligęza, Logical Foundations for Rule-Based Systems, Springer-Verlag, Berlin, Heidelberg, 2006.
[25] J. Liu, L. Martinez, A. Calzada, H. Wang, A novel belief rule base representation, generation and its inference methodology, Knowl.-Based Syst. 53 (2013) 129–141.
[26] L. Ljung, System Identification—Theory for the User, second ed., Prentice-Hall, Englewood-Cliffs, NJ, 1999.
[27] J. Ma, J. Lu, G. Zhang, Information inconsistencies detection using a rule-map technique, Expert Syst. Appl. 36 (10) (2009) 12510–12519.
[28] E. Mazloumi, G. Rose, G. Currie, S. Moridpour, Prediction intervals to account for uncertainties in neural network predictions: methodology and application in bus travel time prediction, Eng. Appl. Artif. Intell. 24 (2011) 534–542.
[29] O. Nelles, Nonlinear System Identification. From Classical Approaches to Neural Networks and Fuzzy Models, Springer-Verlag, Berlin, Heidelberg, 2001.
[30] W. Reinelt, L. Ljung, Robust control of identified models with mixed parametric and non-parametric uncertainties, Eur. J. Control 9 (4) (2003) 373–380.
[31] D. Senthilkumar, C. Mahanta, Identification of uncertain nonlinear systems for robust fuzzy control, ISA Trans. 49 (2010) 27–38.
[32] G. Shafer, Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[33] D.L. Shrestha, D.P. Solomatine, Machine learning approaches for estimation of prediction interval for the model output, Neural Netw. 19 (2006) 225–235.
[34] J. Sjöberg, Q. Zhang, L. Ljung, A. Berveniste, B. Delyon, P. Glorennec, H. Hjalmarsson, A. Juditsky, Nonlinear black-box modeling in system identification: a unified overview, Automatica 31 (1995) 1691–1724.
[35] I. Škrjanc, S. Blažič, O. Agamennoni, Interval Fuzzy Model Identification Using l_∞-Norm, IEEE Trans. Fuzzy Syst. 13 (5) (2005) 561–567.
[36] I. Škrjanc, S. Blažič, O. Agamennoni, Identification of dynamical systems with a robust interval fuzzy model, Automatica 41 (2005) 327–332.
[37] I. Škrjanc, Confidence interval of fuzzy models: an example using a waste-water treatment plant, Chemometr. Intell. Lab. Syst. 96 (2009) 182–187.
[38] M.H. Stone, The generalized Weierstrass approximation theorem, Math. Mag. 21 (1948) 167–183.
[39] R. Sun, Robust reasoning: integrating rule-based and similarity-based reasoning, Artif. Intell. 75 (1995) 241–295.
[40] C.S. Tseng, B.S. Chen, H_∞ fuzzy estimation for a class of nonlinear discrete-time dynamic systems, IEEE Trans. Signal Process. 49 (11) (2001) 2605–2619.
[41] Y.M. Wang, J.B. Yang, D.L. Xu, Environmental impact assessment using the evidential reasoning approach, Eur. J. Oper. Res. 174 (3) (2006) 1885–1913.
[42] D.L. Xu, J. Liu, J.B. Yang, G.P. Liu, J. Wang, I. Jenkinson, J. Ren, Inference and learning methodology of belief-rule-based expert system for pipeline leak detection, Expert Syst. Appl. 32 (1) (2007) 103–113.
[43] R.R. Yager, Modeling uncertainty using partial information, Inform. Sci. 121 (1999) 271–294.
[44] J.B. Yang, Rule and utility based evidential reasoning approach for multiple attribute decision analysis under uncertainty, Eur. J. Oper. Res. 131 (1) (2001) 31–61.
[45] J.B. Yang, J. Liu, J. Wang, H.S. Sii, H.W. Wang, A belief rule-base inference methodology using the evidential reasoning approach – RIMER, IEEE Trans. Syst. Man Cybernet. – Part A 36 (2) (2006) 266–285.
[46] J.B. Yang, J. Liu, D.L. Xu, J. Wang, H.W. Wang, Optimization models for training belief rule based systems, IEEE Trans. Syst. Man Cybernet. – Part A 37 (4) (2007) 569–585.
[47] K. Yoon, C.L. Hwang, Multiple Attribute Decision Making: An Introduction, Sage Publications, Thousand Oaks, CA, 1995.
[48] B.C. Zhang, X.X. Han, Z.J. Zhou, L. Zhang, X.J. Yin, Y.W. Chen, Construction of a new BRB based for time series forecasting, Appl. Soft Comput. 13 (12) (2013) 4548–4556.
[49] F. Zheng, Q.G. Wang, T.H. Lee, Adaptive and robust controller design for uncertain nonlinear systems via fuzzy modeling approach, IEEE Trans. Syst. Man Cybernet. – Part B: Cybernet. 34 (1) (2004) 166–178.
[50] Z.J. Zhou, C.H. Hu, D.L. Xu, J.B. Yang, D.H. Zhou, New model for system behaviour prediction based on belief-rule-based systems, Inform. Sci. 180 (2010) 4834–4864.