Applied Soft Computing 10 (2010) 439–444

Contents lists available at ScienceDirect

Applied Soft Computing

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

Fuzzy functions based ARX model and new fuzzy basis function models for nonlinear system identification

Selami Beyhan, Musa Alci *

Electrical and Electronics Engineering Department, Ege University, Izmir, Turkey

ARTICLE INFO

Article history:
Received 28 November 2008
Received in revised form 13 May 2009
Accepted 2 August 2009
Available online 19 August 2009

Keywords:
Fuzzy functions
ARX modeling
FCM clustering
Fuzzy basis functions
System identification
LSE

ABSTRACT

In this study, auto regressive with exogenous input (ARX) modeling is improved with fuzzy functions concept (FF-ARX). Fuzzy function with least squares estimation (FF-LSE) method has been recently developed and widely used with a small improvement with respect to least squares estimation method (LSE). FF-LSE is structured with only inputs and their membership values. This proposed model aims to increase the capability of the FF-LSE by widening the regression matrix with lagged input–output values. In addition, by using same idea, we proposed also two new fuzzy basis function models. In the first, basis of the fuzzy system and lagged input–output values are structured together in the regression matrix and named as "L-FBF". Secondly, instead of using basis function, the membership values of the lagged input–output values are used in the regression matrix by using Gaussian membership functions, called "M-FBF". Therefore, the power of the fuzzy basis function is also enhanced. For the corresponding models, antecedent part parameters for the input vectors are determined with fuzzy c-means (FCM) clustering algorithm. The consequent parameters of the all models are estimated with the LSE. The proposed models are utilized and compared for the identification of nonlinear benchmark problems.

1. Introduction

System identification and modeling is an essential and important subject for controlling the systems without human intervention. A mathematical model of the system or artificial intelligent model which has same input–output characteristic with model is necessary to analyze and control the system. Modeling which is based on physical laws forms a mathematical model for the system. However, for the identification, there is no need to use previous knowledge and physical structure of the system, so they are known as the black-box identification process [1]. If the behavior of the system is much complicated, it is a big difficulty for the identification. Because of wide variety of the systems parameters, there is not a universal solution to identification of systems. Due to nonlinear behavior, in the beginning of the identification task, selection of the identification method is usually the most difficult part. Appropriate methods can be chosen according to system behavior and desired goal of identification. Therefore, it is necessary to use enhanced identification methods to determine the model that approximates the system correctly.

In addition to the success of the identification method, complexity of the method has gained much attention because of its ease implementation. For that reason, linear estimation models are more consulted in the beginning. Because of their simple structures and well-understood behavior, linear models are widely used in many process modeling. However, the mapping capability of the linear models is usually failed for many noise contaminated system or nonlinear system modeling. As a result, the linear models need to be improved every time. Two most popular models of linear models are the finite impulse response (FIR) and the ARX model. Because FIR models are constructed with the lagged inputs to capture the dynamics of the process, they are not parsimonious and require a large number of model parameters. On the other hand, ARX models are more parsimonious because they represent the model output as a linear sum of both lagged inputs and outputs. Therefore, ARX model does not require so many model parameters. Because of the delayed output components, it has more approximation capability than that of FIR model. Thus ARX model is utilized numerously for identification tasks itself and in different structures [2].

Fuzzy functions are alternate representation and reasoning schemas to the fuzzy rule base approaches. The named "fuzzy functions" are structured with the scalar variables and their membership values are added to least squares regression matrix as a new variable [3]. Therefore, the fuzzy system output parameters

* Corresponding author. Tel.: +90 232 3886024; fax: +90 232 3886024.
E-mail addresses: selami.beyhan@ege.edu.tr (S. Beyhan), musa.alci@ege.edu.tr, umurcali@gmail.com (M. Alci).

1568-4946/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2009.08.015

and least squares parameters are estimated with LSE. FF-LSE has been proved to result 10% better than LSE. The aim of this study is to improve the FF-LSE system performance. Although the research based on FF system development and mathematical basics has been started since 1999 [4–6], the practical application to the system identification, pattern classification and regression has just begun after Türkşen's works [7]. In his study, the FF system and its comparison to the fuzzy rule base were demonstrated. The FF system is usually utilized for different problems such as, system identification [3,8] and pattern classification [9]. Turkşen also recently published a review on the fuzzy system model related relation to the FF concept. The recent work, up to now, the review of the fuzzy system model relation to the FF concept is explained in [10].

On the other hand, a fuzzy system was developed by Wang and Mendel using fuzzy basis functions (FBF), in a pretty construction [11]. In that work, fuzzy basis functions or normalized membership functions were proved as universal approximators by using Stone–Weierstrass theorem. The parameters were estimated with orthogonal least squares algorithm.

Organization of this paper is as follows; in the second part, previously known linear estimation models such as, LSE, ARX, FF-LSE and FBF modeling and the new proposed models such as, FF-ARX, L-FBF and M-FBF modeling are explained in detail. In third part, all methods discussed above are tested for the challenging nonlinear benchmark systems. Finally comparisons are made on computer simulations, and results of the study are represented here to show effectiveness of the proposed models.

2. Fuzzy functions based ARX model

The input–output characteristic of the nonlinear systems is changing naturally with high noise disturbance and its time varying behavior. Therefore, the input–output appearance may be linearly or nonlinearly changed. In order to extract these dynamics, we need to use combinatorial models of linear and nonlinear methods. To identify the highly nonlinear systems, strong nonlinear models should be used such as neural network, fuzzy logic or their complex recurrent models. These methods do perform highly nonlinear static mapping. However for linear or less nonlinear systems, these nonlinear models are not well suited, resulting in less accurate identification. There should be used linear–nonlinear methods together. Several researchers have already focused on this topic in detail. In Ref. [12], the linear dynamic part was constructed by Laguerre basis and the static nonlinear part was constructed by the wavelet network, which gives effective results. By the same construction, Laguerre basis and fuzzy logic combination was studied in [13]. Finally, another model [14] was designed as a combination of stable linear system and neural network models. But in those models, the methods were more parametric and they need to use fine optimization methods. In fuzzy function modeling, parameters are linearly estimated by LSE and less number of parameters is used as compared to above models. But the regressor part of the fuzzy function is the linear and nonlinear functions of the input–outputs, so that the resulting system modeling is extracting dynamics of the models as well as above more complex methods. Therefore, fuzzy function modeling is more parsimonious than previous models.

2.1. Least squares estimation modeling

The LSE modeling is the basic linear regression model based on the minimization of error squares [15]. The output values $\{y_k\}$ $k = 1,2,\ldots,N$ of the LSE model are the linear functions of the one or more input values $\{u_{i,k}\}$ $i = 1,\ldots,nd$ and $k = 1,2,\ldots,N$. Here, N is the number of samples and nd is the number of inputs.

$$\hat{y}(k) = \beta_0 + \sum_{i=1}^{\mathrm{nd}} \beta_i u_i(k) + \varepsilon = \beta_0 + \beta_1 u_1(k) + \ldots + \beta_{\mathrm{nd}} u_{\mathrm{nd}}(k) \quad (1)$$

where β_0 is the constant (intercept parameter), β_is are the regressor coefficients (slope parameters) and ε is the random errors (or residuals) of the estimation and $\hat{y}(k)$, $u_i(k)$ are the process output and input at time step "k", respectively. The aim of LSE is to find unknown parameters β_is $i = 1,2,\ldots,$ nd with minimum sum squared error. The inputs and outputs are shown in the matrix form as following.

$$Y_{N\times 1} = \varphi_{N\times(\mathrm{nd}+1)} \beta_{(\mathrm{nd}+1)\times 1} + \varepsilon_{N\times 1} \quad (2)$$

In this representation; $\varphi_{N,\mathrm{nd}+1}$ is the regression matrix, $\varphi_{1\times(\mathrm{nd}+1)}(k) = [1 u_1(k) u_2(k) \ldots u_{\mathrm{nd}}(k)]^T$ is the regression vector at time index "k". The $Y_{N\times 1} = [y_1, y_2, \ldots, y_N]$ represents the sampled output, $\beta^T_{(\mathrm{nd}+1)\times 1} = [\beta_0, \beta_1, \ldots, \beta_{\mathrm{nd}}]$ represents the parameter vector, and $\varepsilon^T_{N\times 1} = [\varepsilon_1, \varepsilon_2 \ldots \varepsilon_N]$ represents the error vector.

The aim is to minimize the objective function's sum of error squares as follows,

$$\begin{aligned} J &= \sum_{k=1}^{N} [y(k) - (\varphi(k)\beta)]^2 \\ &= \sum_{k=1}^{N} [y(k) - (\beta_0 + \beta_1 u_1(k) + \ldots + \beta_{\mathrm{nd}} u_{\mathrm{nd}}(k)]^2 \end{aligned} \quad (3)$$

To minimize the objective function, optimally, it is necessary to take the derivative with respect to the β parameters, resulting,

$$\begin{aligned} &\min J = (Y - \varphi\beta)^T (Y - \varphi\beta) = e^T e \\ &\frac{\partial J}{\partial \beta} = 2e^T \frac{\partial e}{\partial \beta} = 2(Y - \varphi\beta)^T (-\varphi) = 0 \\ &(-\varphi)^T (Y - \varphi\beta) = 0 \\ &(\varphi^T \varphi)\beta - \varphi^T Y = 0 \end{aligned}$$

where β is the optimal parameter vector to minimize the sum square errors as fallows.

$$\beta = (\varphi^T \varphi)^{-1} \varphi^T Y \quad (4)$$

2.2. ARX modeling

ARX modeling is the simplest, linear, auto regressive, equation error model and it is a base for the other advanced stochastic block-box models. ARX model parameters are also estimated with the superior least squares method [1].

At first, the block-box model input–output pairs are measured from the process and then used for the estimation of the model parameters. The inputs and the corresponding outputs are $\{u_i\}$, $\{y_i\}$, $i = 1,2,\ldots, N$ respectively and N is the number of input–output samples.

$$y(k+1) = \sum_{i=0}^{\mathrm{nd}} \beta_i u(k-i) + \sum_{j=1}^{\mathrm{np}} \alpha_j y(k-j) + \varepsilon \quad (5)$$

where $y(k)$ and $u(k)$ are the process output and input at time step "k". In Eq. (5), nd is the number of lagged inputs and np is the number of lagged outputs. To find the parameters β_is and α_js, the same procedure is followed as above. The matrices and equations as in Eqs. ((2)–(4)) are constructed and optimal parameters are calculated in the same way. The choice of the lagged vectors of the input and output terms is dependent to system dynamics.

2.3. FCM clustering

The FCM algorithm is one of the clustering and structure identification methods. It was developed and used for different purposes such as by Abonyi and Feil [16], Xiong et al. [17], and Cheng and Lin [18]. For the proposed model, FCM algorithm is necessary to determine the regression matrix vector membership values and cluster centers. In FCM, the data sample can belong to more than one cluster with different fuzzy membership values that represent the degree of membership between data and centers of clusters [19].

At each iteration l (l = 0,1, 2,..), the membership values (u_{ik}) and cluster centers (c_i) are determined as following;

$$u_{ik} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\|x_k - c_i\|}{\|x_k - c_j\|} \right)^{2/(m-1)}}, \quad 1 \le i \le C, \quad 1 \le k \le N \tag{6}$$

$$c_i = \frac{\sum_{k=1}^{N} (u_{ik})^m x_k}{\sum_{k=1}^{N} (u_{ik})^m}, \quad 1 \le i \le C \tag{7}$$

The cost function which minimizes the Euclidian distance between the cluster centers and samples can be expressed as following.

$$J_m = \sum_{k=1}^{N} \sum_{i=1}^{C} (u_{ik})^m \|x_k - c_i\|^2, \quad 1 < m \le \infty \tag{8}$$

The algorithm is terminated when $\|U^{l+1} - U^l\| < \varepsilon$, where the U matrix is constructed with membership values of data points computed from Eq. (6), ε is a small constant to stop clustering algorithm and l is iteration number. Where the N is number of samples, the C is number of clusters, and the m is weighting factor, which is selected as 2 at all computer simulations.

2.4. Fuzzy functions based LSE (FF-LSE)

In last years, fuzzy function subject has got more attention for the identification and classification because of its simple structure and higher approximation capability. In the development of fuzzy function, the LSE method and FCM method are enough for FFs. For that reason, without any knowledge of fuzzy system concept somebody can construct and use the FFs for the desired task. For the FF-LSE, the least squares regression matrix is defined as φ = [1 b X], where X is the input vectors and b is the normalized membership values of these input vectors. The FF-LSE construction can be done in the following steps [3].

(1) Perform the FCM algorithm and define the (local) optimum membership values of the inputs (u_{ik}, computed from Eq. (6)). Then determine an alpha-cut (α) to eliminate harmonics generated by FCM.

$$\mu_{ik} = \begin{cases} u_{ik} & \text{if } u_{ik} \ge \alpha \\ \alpha & \text{if } u_{ik} < \alpha \end{cases} \tag{9}$$

(2) Normalize the membership values $\mu_{ik}(l)$ as

$$\gamma_{ik} = \frac{\mu_{ik}(x_k)}{\sum_{i=1}^{C} \mu_{ik}(x_k)} \tag{10}$$

(3) Finally construct the new least squares matrix with inputs and these normalized membership values. The $b_i = \{\gamma_{ik} | i = 1, \ldots, C; k = 1, \ldots, N\}$ has the normalized membership functions values. The b_i has the membership values of this vector with respect to all centers for given input. The estimated output values are written below.

$$Y_{N\times 1} = \varphi_{N\times(\text{np}+\text{nd}+1)} \beta_{(\text{nd}+\text{nd}+1)\times 1} + \varepsilon_{N\times 1} \tag{11}$$

Then by using Eq. (4) we can find the parameters by LSE method. Here different function of the membership values is also defined and tested in [3,8] to obtain better results such as b^2, b^m, $\exp(b)$.

2.5. Proposed Model 1: Fuzzy Functions in ARX model (FF-ARX)

The higher approximation capability of the ARX model with respect to LSE for identification is embedded into fuzzy function. In fact the FF is constructed with linear and nonlinear parts. In other words, we combined the fuzzy functions with ARX structure using the LSE. The resulting model is basically a linear estimation but nonlinear in behave since membership values are determined in nonlinear sense. For the FF-ARX model, different delays are applied for the model and so there is seen an enhancement in the approximation capability. The proposed model has importance for nonlinear systems. So the corresponding task is performed by using linear and nonlinear parts of the fuzzy function together. The consequent FF-ARX model is shown in Fig. 1.

One advantage of the FF-ARX model is the one-step modeling. There is no need of an optimization method and hence local minimum problem does not occur. First part of the work is depending on the FCM clustering algorithm. Initially, inputs are clustered at once and the resulting membership values are determined by the FCM for the FF. Then the FF and the ARX structures or regressors are combined to construct the FF-ARX structure. Finally, the LSE is enough to identify complicated model estimation.

2.6. Fuzzy Logic Systems (FLS) and Fuzzy Basis Functions (FBF)

A fuzzy system is a universal approximation tool that consists of a set of IF-THEN type rule bases, each of which has an antecedent and consequent parts [19]. These rule bases have rules to indicate the system input–output relations. Rules are constructed with linguistic propositions, which have linguistic variables and linguistic values. In fuzzy systems, these linguistic values are represented with membership functions. Therefore, they are the

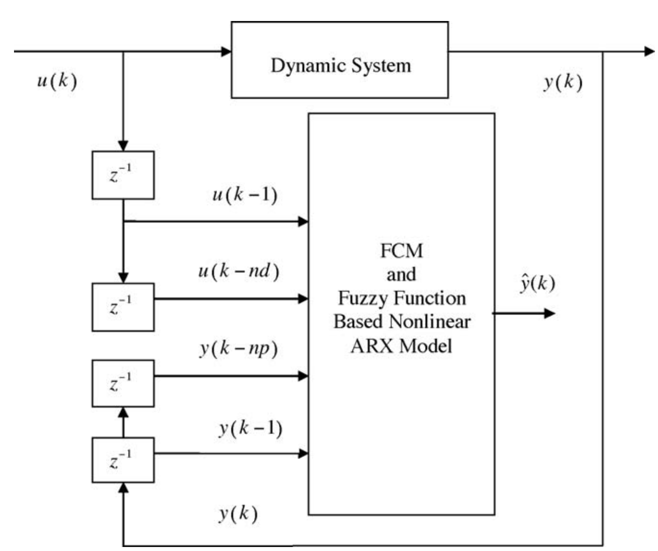

Fig. 1. FF-ARX model.

basic and essential elements of the rule base. A fuzzy system is not linear with respect to membership functions. A fuzzy rule-base consists of M rules is represented by Jang et al. [15], and Wang [19] as follows;

$$R_l : IF\, x_1 \text{ is } A_1^l \text{ and } x_2 \text{ is } A_2^l \text{ and} \ldots \text{and } x_n \text{ is } A_n^l, \text{ THEN } y^l \text{ is } B^l \tag{12}$$

where $l = 1,2, \ldots, M$, $x_i(i = 1,2,\ldots, n)$, inputs to fuzzy system, y is output variable, A_i^l, B_i^l are linguistic values which are represented by membership functions.

FLS: A fuzzy logic system can be represented as a non-orthogonal expansion using normalized input membership functions [19],

$$f(x) = \sum_{l=1}^{M} p_l(x) y_l \tag{13}$$

where $y_l \in R$ are constants or output weights. The normalized input membership functions $p_l(x)$ are defined as follows;

$$p_l(x) = \frac{\prod_{i=1}^{n} \mu_{A_i^l}(x_i)}{\sum_{l=1}^{M} \prod_{i=1}^{n} \mu_{A_i^l}(x_i)} \tag{14}$$

where $l = 1,2, \ldots, M$ rule index, and $\mu_{A_i^l}(x_i)$ are the Gaussian membership functions. The fuzzy model in Eqs. (13) and (14) is formed using singleton fuzzifier, product inference engine, and center of average defuzzifier [19]. The Gaussian membership functions are defined as fallows.

$$\mu_{A_i^l}(x_i) = \exp\left[-\left(\frac{x_i - c_i^l}{b_i^l}\right)^2\right] \tag{15}$$

Where the c_i^l, b_i^l, and y_l parameters are centers and variances of the Gaussian membership functions and output weights, respectively. In the training part of the identification, these variables are updated with optimization algorithms such as gradient descent or Genetic Algorithms [15,19].

FBF: A fuzzy system (Eq. (13)) can be represented by series expansions of the basis functions (Eq. (14)) [11]. It is a linear combination of basis that brings us to use linear estimation methods to use parameter estimation. Differences of the FBF system with respect to FLS, the membership function parameters are such as centers or standard deviations in Eq. (15) are determined by the FCM clustering algorithm or optimization algorithms. However, the output weight parameters are estimated by linear estimation methods. The kth sample regression vector is defined as $\varphi(k) = [p_1(k)\, p_2(k) \ldots p_M(k)]^T$, where M is the number of rules.

The FF model in Section 2.4 and FLS model in Section 2.6 are different methods. Due to parameter estimation, the FF concept is more similar to the FBF modeling. However, by looking the regression matrix construction, the FF modeling and FBF modeling are also different methods. The small difference between them brings the more or less degree of nonlinearity mapping. From the simulation results, the comparison is explained between methods.

2.7. Proposed fuzzy basis function models

The fuzzy basis function modeling brings us to see that the fuzzy system output parameters are linearly dependent to fuzzy system. Therefore, the output parameters can be determined in linear sense. Here, for the sake of simplicity, all model output parameters of FBF models are estimated with LSE.

2.7.1. Proposed Model 2: Lagged Terms Based Fuzzy Basis Functions (L-FBF)

The first proposed FBF model (L-FBF) regression matrix is constructed with lagged input–output vectors and the fuzzy basis. The regression vector $\varphi = [1\ u\ y\ p(u)\ p(y)]^T$ is given by,

$$\varphi(k) = [1 \quad u(k-1)\ldots u(k-\text{np}) \quad y(k-1)\ldots y(k-\text{nd}) \quad \mu_1(u(k-1))\ldots\mu_M(u(k-\text{np})) \quad \mu_1(y(k-1))\ldots\mu_M(y(k-\text{nd}))]^T \tag{16}$$

where the np is number of input delay, the nd is the number of output delay. To implement this method for identification, the following steps can be carried out.

- Define how many input–output lagged terms are used in regression vector. It can be changed by considering the model complexity.
- Perform FCM clustering algorithm to determine the centers of the regressors. The Gaussian membership functions are used with 0.5 width.
- Do normalization for the corresponding part of the regressor and construct the regression matrix.
- Calculate the weight parameters of the model by using the LSE method.

2.7.2. Proposed Model 3: Membership Functions Based Fuzzy Basis Functions (M-FBF)

In the second proposed FBF model (M-FBF) regression matrix is constructed with the lagged input–output vectors and the membership values of these vectors. The membership functions are chosen as Gaussians functions with 0.5 width and the centers are determined by the FCM method. The regression vector is given by,

$$\varphi = [1 \quad u \quad y \quad \mu(u) \quad \mu(y)]^T \tag{17}$$

and it is expressed more clearly as following,

$$\varphi(k) = [1 \quad u(k-1)\ldots u(k-np) \quad y(k-1)\ldots y(k-nd) \quad \mu_1(u(k-1))\ldots\mu_M(u(k-\text{np})) \quad \mu_1(y(k-1))\ldots\mu_M(y(k-\text{nd}))]^T \tag{18}$$

This structure does not need normalization; exact membership value of the corresponding data is located in regression matrix. In order to implement this method for identification, the following steps can be carried out.

- Define how many input–output lagged terms are used in regression vector. It can be changed by considering the model complexity.
- Perform FCM clustering algorithm to determine the centers of the regressors. The membership functions are chosen as Gaussians functions with 0.5 standard deviation. Construct regressor matrix as Eq. (17).
- Finally, calculate the weight parameters of the model by using the LSE method.

3. Simulation results

In simulations, firstly, a simple dynamic function approximation is identified by the three proposed models and then, two benchmark problems are employed to identify and compare the three proposed models with other methods. In the simulations for the faith comparison, same numbers of delays are used for inputs and outputs. In simulation results; the ARX, the FBF, the proposed FF-ARX, L-FBF, and M-FBF models are compared. In the end, there is seen that the proposed models, which are the FF-ARX and the L-FBF, have better resulting MSE values than others do. In simulations, the alpha-cut is selected as 0.6 and ε is used as 0.001 in FCM clustering.

3.1. Simple nonlinear function approximation

The aim of this testing is to see the function approximation capabilities of proposed models before the highly dynamic nonlinear systems identification. For training 150 samples of input–output data is generated by the following functions [20],

$$yd(n) = \frac{yd(n-1)}{1+yd(n-1)^2} + u(n) \quad \text{where the input}$$

$$u(n) = \sin\left(\frac{2\pi n}{25}\right).$$

To see the effectiveness of the proposed models, the frequency of the input and the one of function parameters are changed and also normally distributed white noise is added to the input in test phase. For testing 150 samples of input–output data generated by the following functions,

$$yd(n) = \frac{1.5 \times yd(n-1)}{1+1.5 \times yd(n-1)^2} + u(n) \quad \text{where the input}$$

$$u(n) = \sin\left(\frac{2\pi n}{50}\right) + \upsilon, \upsilon \sim N(0, 0.22)$$

In this simulation, the input u_{k-1} and the output y_{k-1} vectors are clustered in two groups. As a result, the regression matrix is constructed by seven vectors for the FF-ARX, the L-FBF and the M-FBF models. It means that seven parameters are used to approximate the function. The used regression vector is

$$\varphi = [1 \quad u_{k-1} \quad y_{k-1} \quad g(u_{k-1}, uc_{c1}) \quad g(u_{k-1}, uc_{c2}) \quad g(y_{k-1}, yc_{c1}) \quad g(y_{k-1}, yc_{c2})]$$

where *ucs* and *ycs* are the cluster centers of the corresponding inputs and the $g(.)$ function is changing for three proposed models. It is the normalized membership values of the inputs for the FF-ARX model, it is the basis function of the corresponding inputs for the L-FBF model, and finally, for the M-FBF model, it is the membership values of the inputs and outputs. The Gaussian membership functions are used and their centers are obtained by the FCM clustering. The resulting identification performances of the methods are represented in Fig. 2.

3.2. Box-Jenkins furnace identification

In this part of simulation study, Box-Jenkins furnace [21] identification is used to compare model capabilities. This data is frequently used in performance evolution of system identification methods [21–23]. First, the data is scaled with subtracting mean of data and dividing by standard deviation of data. First 200 input–output samples are used for the modeling the system and last 90 input–output samples are used for the testing and the resulting MSE values are shown in Table 1. The test output errors of compared models are shown in Fig. 3. The instantaneous model output $\hat{y}(k)$ is formed by using six regressors, i.e. regression vector

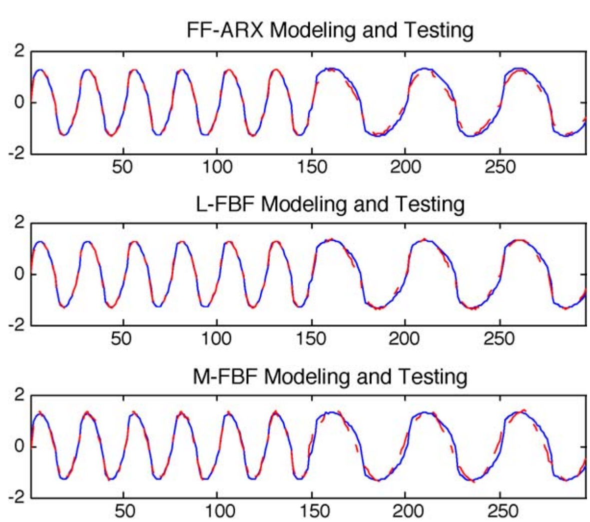

Fig. 2. Function approximation.

Table 1
Comparison for Box-Jenkins data.

Model	MSE (modeling)	MSE (testing)
ARX Model	45.1e-3	0.305
FBF Model	419e-3	0.653
FF-ARX Model	7.5e-3	0.075
L-FBF Model	5.6e-3	0.021
M-FBF Model	7.8e-3	0.056

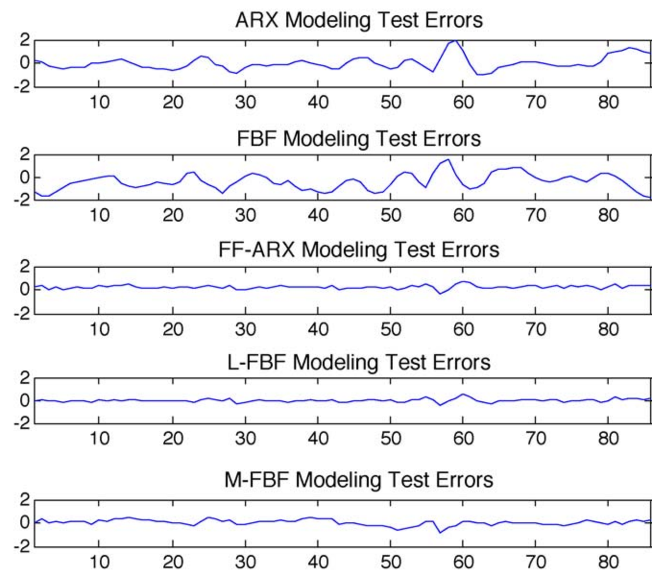

Fig. 3. Box-Jenkins furnace identification testing errors.

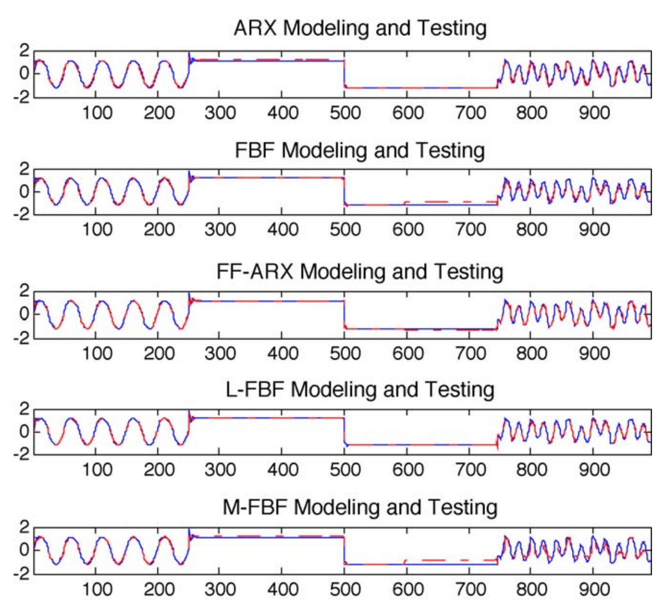

Fig. 4. Training and test performances for nonlinear dynamic model identification example (training for $t = 1{:}600$, and testing for $t = 601{:}1000$.) The system output is represented with line, and the approximations are represented with dashed dot line.

Table 2
Comparison for nonlinear dynamic model.

Model	MSE (modeling)	MSE (testing)
ARX Model	18e-3	22.3e-3
FBF Model	5.7e-3	51.7e-3
FF-ARX Model	2.5e-3	10.7e-3
L-FBF Model	2.4e-3	7.5e-3
M-FBF Model	2.9e-3	38.4e-3

φ include $y(k-1)$, $y(k-2)$, $y(k-3)$, $y(k-4)$, $u(k-1)$, $u(k-2)$, and $u(k-3)$. When modeling the system there is used different number of centers to get minimum MSE values. For this benchmark problem the optimal number of centre is selected as 2.

3.3. Dynamic nonlinear model identification

The other data is generated by another well-known benchmark problem. The identification of this dynamical system is most frequently used in the literature for comparing different learning algorithms and model types [24,25].

$$u(k)=\begin{cases}\sin(2\pi k/25) & k=1\ldots250\\ 1 \quad \text{if} \quad k=250\ldots500\\ -1 \quad \text{if} \quad k=500\ldots750\\ 0.3\sin(k\pi/25)+0.1\sin(k\pi/32)+0.6\sin(k\pi/10) & k=750\ldots1000\end{cases} \tag{19}$$

$$y(k)=\frac{[y(k-1)y(k-2)y(k-3)u(k-2)\times((y(k-3)-1)+0.5)+u(k)]}{[1+y(k-2)^2+y(k-3)^2]} \tag{20}$$

In this simulation totally 1000 samples are used. The first 600 samples for modeling the system and the last 400 samples are for the testing. The estimated outputs for all compared models are constructed with following functional form.

$$\hat{y}(k)=f(u(k-1),u(k-2),u(k-3),u(k-4),y(k-1),y(k-2),y(k-3),y(k-4)) \tag{21}$$

After some trials, the number of centers is selected as 4 and optimal alpha-cut is selected as 0.6 for this benchmark problem. The resulting MSE values for different models are shown in Table 2 and simulation results are shown in Fig. 4.

This identified system is highly nonlinear and dynamic thus it is difficult to identify by linear estimation models. In the benchmark data, the test input–output signals differ than the modeling signals. Even so it is modeled and tested with small MSE values.

4. Conclusion

In this paper, we have proposed the FF-ARX, the L-FBF, the M-FBF models and compared their estimation capabilities. The FF-ARX model and the L-FBF model have better function approximation capability with respect to the other known LSE models. The power of FF and ARX models are combined in the proposed FF-ARX model to improve system identification parameters. On the other hand, in the L-FBF model, the power of the nonlinear structure of the FBF and linear structure of the other basis are combined using the LSE. The FBF models are nonlinear with respect to inputs. The whole LSE based models are simple to construct and use for identification of nonlinear dynamic systems. The proposed models are simple and feasible compared to existing nonlinear system identification methods. Our results indicate that the fuzzy function subject has improved in the proposed ARX model structure and in the proposed L-FBF model structure.

References

[1] L. Ljung, System Identification Theory for the User, Prentice Hall PTR, 1999.
[2] M.N. Nounou, H.N. Nounou, Improving the prediction and parsimony of ARX models using multiscale estimation, Applied Soft Computing 7 (2007) 711–721.
[3] I.B. Türkşen, Fuzzy functions with LSE, Applied Soft Computing 8 (2008) 1178–1188.
[4] M. Demirci, Fuzzy functions and their fundamental properties, Fuzzy Sets and Systems 106 (1999) 239–246.
[5] M. Demirci, Fuzzy functions and their applications, Journal of Mathematical Analysis and Applications 252 (2000) 495–517.
[6] A. Wielgus, Algorithms of multilevel synthesis of fuzzy functions, IEEE International Conference on Fuzzy Systems, vol. 2, 25–29 July 2004, pp. 929–934, Budapest, Hungary.
[7] I.B. Türkşen, A. Celikyilmaz, Comparison of fuzzy functions with fuzzy rule base approaches, International Journal of Fuzzy Systems 8 (September (3)) (2006).
[8] I.B. Türksen, Mixed fuzzy functions, Annual Meeting of the North American Fuzzy Information Processing Society, NAFIPS 2008, 19–22 May 2008, pp. 1–5, New York, USA.
[9] A. Çelikyılmaz, I.B. Türkşen, R. Aktaş, M.M. Doğanay, N.B. Ceylan, A new classifier design with fuzzy functions, Lecture Notes in Computer Science, vol. 4482/2007, pp. 136–143, 2007.
[10] I.B. Türkşen, Review of fuzzy system models with an emphasis on fuzzy functions, Transactions of the Institute of Measurement and Control 31 (1) (2009) 7–31.
[11] L.-X. Wang, J.M. Mendel, Fuzzy basis functions, universal approximation, and orthogonal least squares learning, IEEE Transactions on Neural Networks 3 (5) (1992).
[12] P. Aadaleesan, N. Miglan, R. Sharma, P. Saha, Nonlinear system identification using Wiener type Laguerre–Wavelet network model, Elseiver Chemical Engineering Science 63 (2008) 3932–3941.
[13] M. Alci, Fuzzy rule-base driven orthogonal approximation, Neural Comput. Appl. 17 (2008) 501–507 (Springer).
[14] F. Abdollahi, H.A.T. Rajnikant, V. Patel, Stable identification of nonlinear systems using neural networks: theory and experiments, IEEE/ASME Transactions on Mechatronics 11 (August (4)) (2006).
[15] J.-S.R. Jang, C.-T. Sun, E. Mizutani, Neuro Fuzzy and Soft Computing, 4, Prentice Hall, USA, 1997, pp. 73–84.
[16] J. Abonyi, B. Feil, Cluster Analysis for Data Mining and System Identification, Birkhauser Verlag AG, Basel, Boston, Berlin, 2007, ISBN 978-3-7643-7987-2.
[17] S.-W. Xiong, H.-B. Liu, X.-X. Niu, Fuzzy support vector machines based on FCM clustering, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, vol. 5, pp. 18–21, 18–21 August 2005, Guangzhou, China.
[18] S.-T.L.-C. Cheng, Su-Yu Lin, A FCM-based deterministic forecasting model for fuzzy time series, Computers and Mathematics with Applications 56 (December (12)) (2008) 3052–3063 (Elsevier).
[19] L.-X. Wang, A Course in Fuzzy Systems and Control, Prentice-Hall Inc., USA, 1997, ISBN 0-13-593005-7.
[20] E. Righeto, L.H.M. Grassi, J.A. Pereira, Nonlinear plant identification by wavelets, Mechatronics Symposium of the 17th COBEM - International Congress of Mechanical Engineering, vol. 1, pp. 392–398, November 10–14, 2004, São Paulo, Brazil.
[21] W. Chung, E. Kim, A new two-phase approach to fuzzy modeling for nonlinear function approximation, IEICE Trans. Inf. Syst. E89–D (September (9)) (2006).
[22] M. Alci, M.H. Asyali, S. Beyhan, Fuzzy-ARMAX modeling for nonlinear system identification, in: Proceedings of First International Conference on Soft Computing Technologies in Economy, 34–40, November, Baku, Azerbaijan, (2007), pp. 19–21.
[23] S. Purwar, I.N. Kar, A.N. Jha, On-line system identification of complex systems using Chebyshev neural networks, Appl. Soft Comput. 7 (2007) 364–372.
[24] M.A. Gonzalez-Olivera, Y. Tang, A new recurrent neurofuzzy network for identification of dynamic systems, Fuzzy Sets Syst. 158 (2007) 1023–1035 (Science Direct).
[25] P.A. Mastorocostas, J.B. Theocharis, A stable learning algorithm for block-diagonal recurrent neural networks: application to the analysis of lung sounds, 2006 © IEEE 1083–4419.