REVIEW

Online system identification using fractional-order Hammerstein model with noise cancellation

Mohammad Jahani Moghaddam

Abstract Slow convergence and low accuracy are two main drawbacks in nonlinear system identification methods. It becomes more complicated when time delay and noises are considered. In this paper, considering a fractional-order Hammerstein model, an online identification method is proposed. A combination of an evolutionary optimization method and recursive least square algorithm is used to estimate the system parameters and orders in the presence of unknown noises. Finally, simulation results are taken to prove the effectiveness of the proposed algorithm.

Keywords Online system identification · Nonlinear systems · Hammerstein model · Modified radial basis function

1 Introduction

Previously, standard systems as well as standard control systems, regardless of reality, were all considered to be integer-order systems. Recent researches, however, have shown that some of real-world systems are fractional order. In recent years, an increasing number of real-time physical systems have been better

M. Jahani Moghaddam (✉)
Department of Electrical Engineering, Langarud Branch,
Islamic Azad University, Langarud, Iran
e-mail: Dr.jahani@iau.ac.ir

described by fractional-order differential equations (FODEs) than classical integer-order models. So, researchers and engineers are increasingly using fractional-order dynamical models to model real physical systems.

Identifying of the mathematical model of the system is very important in analyzing the properties of the system as well as designing a suitable controller for the system. System identification is a standard tool for modeling unknown systems whose main purpose is to determine the structure and parameters of the mathematical model of the system to reproduce the dynamic behavior of the system. This process becomes more difficult when physical systems are described by FODE instead of integer-order models. In this regard, the high complexity and lack of sufficient mathematical tools had led to little attention to fractional-order (FO) dynamical systems in theory and practice [1, 2]. But today, however, with the growth of computers and their ability to compute complex integrals and FO derivatives, this problem has been somewhat solved and fractional calculations have become an attractive research topic in the scientific and industrial communities. In recent decades, due to this growth, along with the theoretical research of fractional integrals and derivatives [3, 4], the use of fractional operators in various fields has been significantly developed.

Fractional calculus has been introduced in various fields of science and engineering [5, 6], including

identification of thermal systems [7], identification of biological tissues [8], image processing [9], signal processing [10], path planning [11], path tracking [12], robotics [13], mechanical damping [14], battery [15], control theory and its applications [16, 17], mechanics [18], diffusion [19]. Also, systems with long-memory transient characteristics and independent frequency domain [14, 20, 21], transmission and distribution lines, electromechanical processes, dielectric polarization, viscoelastic materials such as polymers and rubbers, relaxation phenomena of organic dielectric materials, flexible structures, traffic in information networks and biological systems [7, 19, 22–29], colored noise, chaos [30], controllers [31], etc., can be modeled more appropriately by FO models than integer-order models. This confirms the fact that a significant number of real systems are generally fractional. Although for many of them, the degree of fractionality is very low. Thus, the application of fractional calculus has become the focus of international academic research, and the identification of such fractional-order systems has aroused growing interest in the scientific community. Although FO models are more suitable for describing dynamic systems than integer-order models, they require suitable methods for analytical or numerical calculations of FODEs [3, 4].

The aim of this paper is presenting a new online method for identifying nonlinear systems which offers two features simultaneously: increasing accuracy and decreasing computations. For this purpose, a nonlinear system identification based on FO Hammerstein model has been considered. The proposed method uses recursive identification algorithm with the ability of online identification.

The reason for using the Hammerstein model (Fig. 1) is that this model does not require a lot of basic information from the system, which makes identifying using process data relatively easier. The motivation for using fractional calculations in system identification is to preserve the features and phenomena ignored by integer-order models by FO models, as well as the fact that the dynamic behavior of an increasing number of real processes can be more accurately expressed using fractional models.

In the FO Hammerstein model identification using input/output data, all unknown system parameters include:

1. The coefficients of numerators and denominators of the transfer function and fractional degrees in the linear dynamics part.
2. Bezir–Bernstein polynomial (BBP) coefficients or radial basis function neural network (RBFNN) parameters including centers, widths and connection weights in the nonlinear static part are estimated. The recursive method in identification is based on updating unknown parameters by adding new input/output data using a recursive optimization algorithm. In this paper, recursive least square (RLS) algorithm is used for this purpose.

The fractional Hammerstein model is developed to identify multi-input–single-output (MISO) nonlinear systems with the following structure (Fig. 2). It is clear that the identification of MISO systems has many practical applications and it can be generalized to multi-input–multi-output (MIMO) systems.

Modeling the behavior of the nonlinear static part is a major challenge in the Hammerstein model. In this paper, two methods are used to represent this part:

1. Bezier–Bernstein polynomials: From the point of numerical analysis view, although different types of polynomial functions can be used to estimate the function, in [32] it is shown that Bernstein basic functions are the best and most stable basic functions against other polynomial basic functions.
2. Artificial neural network: An important advantage of using artificial neural network is overcoming limitations such as slow convergence and complexity of the structure [33].

Weaknesses and limitations of existing methods of identifying the Hammerstein model, which has become a motivation to present a new method, can be mentioned as follows:
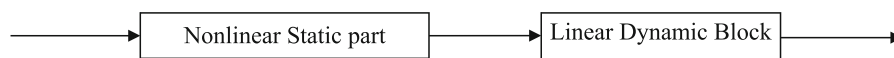
Fig. 1 Single-input–single-output Hammerstein model structure

Fig. 2 MISO Hammerstein model structure
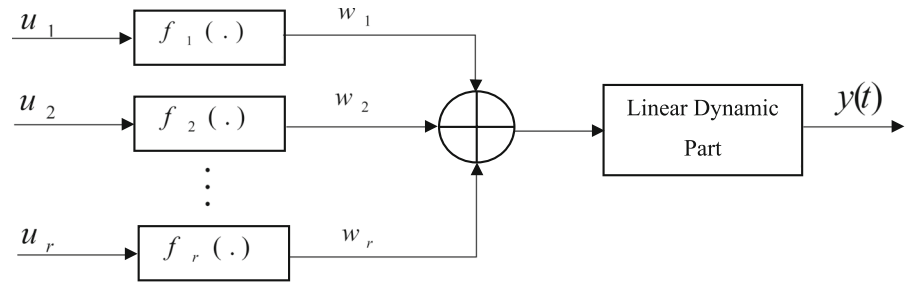
1. Limited to the proportional fractional orders [34–41]
2. Assuming the fractional orders as known [34–36, 39, 40]
3. Disqualification for online identification due to computational complexity [34–41]
4. High estimation errors [34–41]
5. Not to consider time delays especially in online mode [34–47]
6. Lack of generality of the presented models. For example, most methods are only applicable to nonlinear systems that have quasi-linear properties [37, 38];
7. Inability in system identification in the presence of noises [43–48].

In this paper, a recursive method that is generalized to online identification will be used and the Hammerstein model will be considered with the following features:

1. The linear part transfer function is considered as a fractional order, which, in addition to more accurate identification of the system, allows the reduction of the number of parameters as a feature of FO systems.
2. Due to the BBP properties and RBFNN in more accurate modeling of nonlinear dynamics, these two functions are used to represent the nonlinear part of the Hammerstein model.
3. The FO Hammerstein model is developed in MISO mode.
4. According to the use of the recursive method and other mentioned features, the proposed identification method is generalized to the online mode.

2 Mathematical background

2.1 Fractional-order models

A continuous-time FO dynamical system can be expressed by a FODE as follows [27, 49]:

$$H(D^{\alpha_0\alpha_1\cdots\alpha_n})y(t) = G(D^{\beta_0\beta_1\cdots\beta_n})u(t),$$
$$H(D^{\alpha_0\alpha_1\cdots\alpha_n}) = \sum_{k=0}^{n} a_k D^{\alpha_k},$$
$$G(D^{\beta_0\beta_1\cdots\beta_n}) = \sum_{k=0}^{m} b_k D^{\beta_k}$$

(1)

where $a_k, b_k \in \mathbb{R}$. In explicit form:

$$a_n D^{\alpha_n}y(t) + a_{n-1}D^{\alpha_{n-1}}y(t) + \cdots + a_0 D^{\alpha_0}y(t)$$
$$= b_m D^{\beta_m}u(t) + b_{m-1}D^{\beta_{m-1}}u(t) + \cdots + b_0 D^{\beta_0}u(t)$$

(2)

By applying the Laplace transform with zero initial conditions, the input–output representation of the FO system can be obtained in the form of a transfer function:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^{\beta_m} + b_{m-1}s^{\beta_{m-1}} + \cdots + b_0 s^{\beta_0}}{a_n s^{\alpha_n} + a_{n-1}s^{\alpha_{n-1}} + \cdots + a_0 s^{\alpha_0}}$$

(3)

2.2 The structure of Hammerstein model

As mentioned, the MISO Hammerstein model consists of a linear static subsystem and a linear dynamic part. In general, the system is modeled as follows:

$$y(t) + \sum_{i=1}^{n} a_i D^{\alpha_i}y(t) = \sum_{j=0}^{m} b_j D^{\beta_j}(w_1(t) + \ldots + w_r(t))$$
$$w_k(t) = f(u_k(t))$$

(4)

where $y(t)$ is the output of the system and $u_k(t), k = 1, 2, \ldots, r$ specifies the inputs, and $w_k(t), k = 1, 2, \ldots, r$ are the outputs of the nonlinear part and the input to the dynamic one, and n, m are the input and output delays for the linear subsystem.

In this paper, for the linear dynamics part, the FO transfer function is used as follows:

$$G(s) = \frac{B(s)}{A(s)} = \frac{b_m s^{\beta_m} + \cdots + b_1 s^{\beta_1} + b_0 s^{\beta_0}}{1 + a_n s^{\alpha_n} + \cdots + a_1 s^{\alpha_1}} \qquad (5)$$

where $a_1, \ldots, a_n, b_0, \ldots, b_m \in R$ are coefficients and $\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_m \in R$ are fractional orders. The corresponding FODE is shown as follows:

$$a_n D^{\alpha_n} y(t) + a_{n-1} D^{\alpha_{n-1}} y(t) + \cdots + a_0 D^{\alpha_0} y(t)$$
$$= b_m D^{\beta_m} u(t) + b_{m-1} D^{\beta_{m-1}} u(t) + \cdots + b_0 D^{\beta_0} u(t) \qquad (6)$$

In the considered fractional-order transfer function, there is no proportional order constraint.

In order to use the output estimation method, Eq. (6) must be written in the regression form:

$$y(t) = -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j} w(t) \qquad (7)$$

In this paper, three different structures are considered for the nonlinear static part:

2.2.1 Bezier–Bernstein polynomials

Reference [32] has shown that among the various types of polynomial functions for estimating a function, Bernstein's basic functions are the best and most stable of the basic functions against other polynomial basic functions. Therefore, the first method considered for modeling the nonlinear static part is the use of BBP functions. In 1912, S.N. Bernstein introduced the following polynomials for a function defined on the interval [1, 0] [50]:

$$B_n f(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1 - t)^{n-i} f\left(\frac{i}{n}\right) \ n = 1, 2, \ldots \qquad (8)$$

Bernstein's polynomials can be defined on an interval [a, b] by the following equation [50]:

$$B_{i,n}(t) = \binom{n}{i} \frac{(t - a)^i (b - t)^{n-i}}{(b - a)^n} \ i = 0, 1, 2, \ldots, n \qquad (9)$$

These polynomials can be used to estimate any continuous function in the interval [a, b] and have the following properties [50]:

$$\sum_{i=0}^{n} B_{i,n}(t) = 1 \qquad (10)$$

A general Bezier curve of degree n, defined by $n + 1$ vertices, can be expressed as follows [50]:

$$B(t) = \sum_{i=0}^{n} a_i B_{i,n}(t) \ t \in [a, b] \qquad (11)$$

where a_i defines the ith vertex and provides information about the shape of the B-curve. Bezier extended the idea of estimating a function to the estimation of a polygon in which $n + 1$ vertices of a polygon are estimated by Bernstein's basis. As a result, it was called the Bezier–Bernstein polynomial curve.

The Bezier–Bernstein polynomial used for the nonlinear static part in this paper is considered in the following form:

$$f(u(t)) = \sum_{j=0}^{d} \delta_j B_{j,d}(x(u(t))) \qquad (12)$$

where j, d are non-negative integers, δ_j are weights that must be specified, $x(u(t))$ converts the input change interval to the interval [1, 0] ($x \in [0, 1]$), and $B_{j,d}(.)$ are the BBPs with the following definition:

$$B_{j,d}(x) = \binom{d}{j} x^j (1 - x)^{d-j} \qquad (13)$$

The number of Bernstein univariate polynomials of degree d is d + 1.

In [40], a formula for Bernstein polynomials defined in interval [a, b] is given as follows:

$$B_{j,d}(x) = \binom{d}{j} \frac{(x - a)^j (b - x)^{d-j}}{(b - a)^d}$$
$$j = 0, 1, \ldots, d \quad a \leq x \leq b$$
$$, a = \min(u(t)), b = \max(u(t)) \qquad (14)$$

With this definition, the relation (12) can be written as follows:

$$f(u(t)) = \sum_{j=0}^{d} \delta_j B_{j,d}(u(t))$$ (15)

Using BBPs to represent the nonlinear part, we have in relation (4):

$$w_k(t) = f_k(u_k(t)) = \sum_{i=0}^{d} \delta_{ki} B_{k\,i,d}(u_k(t))$$ (16)

$B_{k\,i,d}(u_k(t)), i = 0, \ldots, d, k = 1, \ldots, r$ are BBPs related to kth input and $\delta_{k\,i}, i = 0, 1, \ldots, d, k = 1, \ldots, r$ are related weights to be determined.

2.2.2 Radial basis function neural network (RBFNN)

In this case, the identification of the fractional-order Hammerstein model with RBFNN in the nonlinear static part and the fractional-order transfer function in the linear dynamic part is investigated.

The two advantages of using a RBFNN for the structure of a nonlinear gain function are the ability to quickly modify modeling during changes in process dynamics and to overcome constraints such as slow convergence and structural complexity [33]. In fact, the ability of Hammerstein models in nonlinear dynamics modeling with relatively simpler models is combined with the ability of more accurately and simpler system estimation using fractional-order transfer functions. As a result, the problems of inadequacy for online identification and high estimation error in existing methods are eliminated. Also, due to RBFNN capabilities for estimating functions, its use in this paper, by increasing the accuracy of nonlinear part estimation, affects the reduction of estimation error. In this case, the output of the nonlinear part and the input to the linear part are obtained from the following relation:

$$w_k(t) = f_k(u_k(t)) = \omega_{0k} + \sum_{h=1}^{M} \omega_{hk} \phi_{hk}(u_k(t))$$ (17)

$\phi_{hk}(.), h = 1, \ldots, M, k = 1, \ldots, r$ are the Gaussian functions of hth hidden node related to kth input and $\omega_{hk}(.), h = 1, \ldots, M, k = 1, \ldots, r$ are the connection weights from hth hidden node to the output related to kth input which should be identified. M is the number of radial basis neural network functions.

$$\phi_{hk}(u_k(t)) = \exp\left\{-\left(\|u_k(t) - c_{hk}\|^2 / d_{hk}^2\right)\right\}$$ (18)

$c_{hk}, d_{hk}, h = 1, \ldots, M, k = 1, \ldots, r$ are the centers and widths of the hth RBF hidden unit associated with the kth input and $\|.\|$ defines the Euclidean norm.

2.2.3 Modified radial basis function neural network (MRBFNN)

RBFNNs consist of only one layer of activation functions that are radially symmetric. In the standard form, the number RBFNN parameters increases exponentially with increasing the number of inputs. In the modified RBFNN, the centers and widths of Gaussian functions are concentrated in a single adjustable point instead of different points, and only different weights are used. Also, in RBFNN, the number of hidden nodes is equal to the number of sampled training data. But in the MRFNN, the number of hidden nodes is limited and selectable. These two features significantly reduce the number of unknown parameters in the identification process. In this case, the Gaussian functions associated with kth input are no longer dependent on the hidden nodes and have a constant center and width for each input. They are as follows:

$$\phi_k(u_k(t)) = \exp\left\{-\left(\|u_k(t) - c_k\|^2 / d_k^2\right)\right\}$$ (19)

That is, $\phi_{hk}(.)$ Gaussian functions have the same centers and widths and are simplified to $\phi_k(.), k = 1, \ldots, r$.

3 System identification algorithms

In the proposed identification method, the modified genetic algorithm (MGA) [51] is applied to identify fractional orders and time delays in the dynamic part and to identify the centers and widths of RBF units (Eqs. 17–19) in the nonlinear part as well as the production of initial estimations for the coefficients of the fractional-order transfer function, BBP weights (Eq. 16) and NN weights (Eqs. 17–19).

GA with an innovative strategy is called as modified genetic algorithm (MGA). Comparing with the classic GA in which the best solutions or chromosomes are selected and transferred to the next generation, the best characteristics or properties are selected and transferred to the next generation in MGA. It is inspired from artificial genetic operation in some

agronomic products that the best properties of different types by genetic manipulation are transferred to the new product. In this algorithm, the crossover is performed exchanging the best genes between the chromosomes [51].

The crossover step is the main difference between the classic GA and MGA. In MGA, after selection the best chromosomes, the genes are changed between them. The aim of continuing this process is finding the best genes instead of the best chromosomes. The artificial parents are generated with the best genes and putting them together. Better solutions, escaping from local optima and faster convergence are some of the advantages of MGA which are very vital in the online identification process. These advantages will be illustrated by examples [51].

For the identification process, a part of the input/output data is used to obtain all the unknown parameters using MGA. In the next step, using these initial estimations, the recursive least squares (RLS) algorithm is used to update and optimize them using the rest of data. The effective combination of these two algorithms provides the ability to track the nonlinear time-variable behavior of the system. The proposed online system identification structure is shown in Fig. 3.

3.1 MISO Hammerstein model using BBPs

Considering Eq. (15), Eq. (7) becomes:

$$y(t) = -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j} \left(\sum_{l=0}^{d} \delta_l B_{l,d}(u(t)) \right) \tag{20}$$

Equation (20) for MISO mode can be written in the following form:

$$y(t) = -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) +$$
$$+ \sum_{j1=0}^{m_1} b_{j1} D^{\beta_{j1}} \left(\sum_{l=0}^{d} \delta_l B_{l,d}(u_1(t)) \right) + ... + \sum_{jr=0}^{m_r} b_{jr} D^{\beta_{jr}} \left(\sum_{l=0}^{d} \delta_l B_{l,d}(u_r(t)) \right) \tag{21}$$

According to Eq. (9), in the general case, for BBP of d degree, we have:

$$y(t) = -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j1=1}^{m_1} b_{j1} D^{\beta_{j1}} \sum_{l1=0}^{d} h_{l1} u^{l1} + ...$$
$$+ \sum_{jr=1}^{m_r} b_{jr} D^{\beta_{jr}} \sum_{lr=0}^{d} h_{lr} u^{lr} \tag{22}$$

Using Grunwald–Letnikov estimation:

$$y(k+1) = -\sum_{i=1}^{n} a'_i Y_i(k)$$
$$+ \sum_{j=0}^{m} \left(\sum_{l=0}^{d} b1'_{jl}(F1_{jl}(k)) + ... + \sum_{l=0}^{d} br'_{jl}(Fr_{jl}(k)) \right) \tag{23}$$

where

$$a'_i = \frac{\frac{a_i}{h^{\alpha_i}}}{1 + \sum_{k=1}^{n} \frac{a_k}{h^{\alpha_k}}}, \ bv'_{jl}$$
$$= \frac{\frac{b_j h_{lv}}{h^{\beta_j}}}{1 + \sum_{k=1}^{n} \frac{a_k}{h^{\alpha_k}}} \quad \begin{matrix} 1 \le i \le n, 1 \le j \le m, 0 \le l \le d, \\ 1 \le v \le r \end{matrix} \tag{24}$$

$$Y_i(k) = \sum_{j=1}^{N} (-1)^j \binom{\alpha_i}{j} y(k+1-j),$$
$$\text{Fv}_{jl}(k) = \sum_{i=0}^{N} (-1)^i \binom{\beta_i}{jv} (u_v(k+1-i))^l \ F1 \le \text{Fv} \le Fr \tag{25}$$
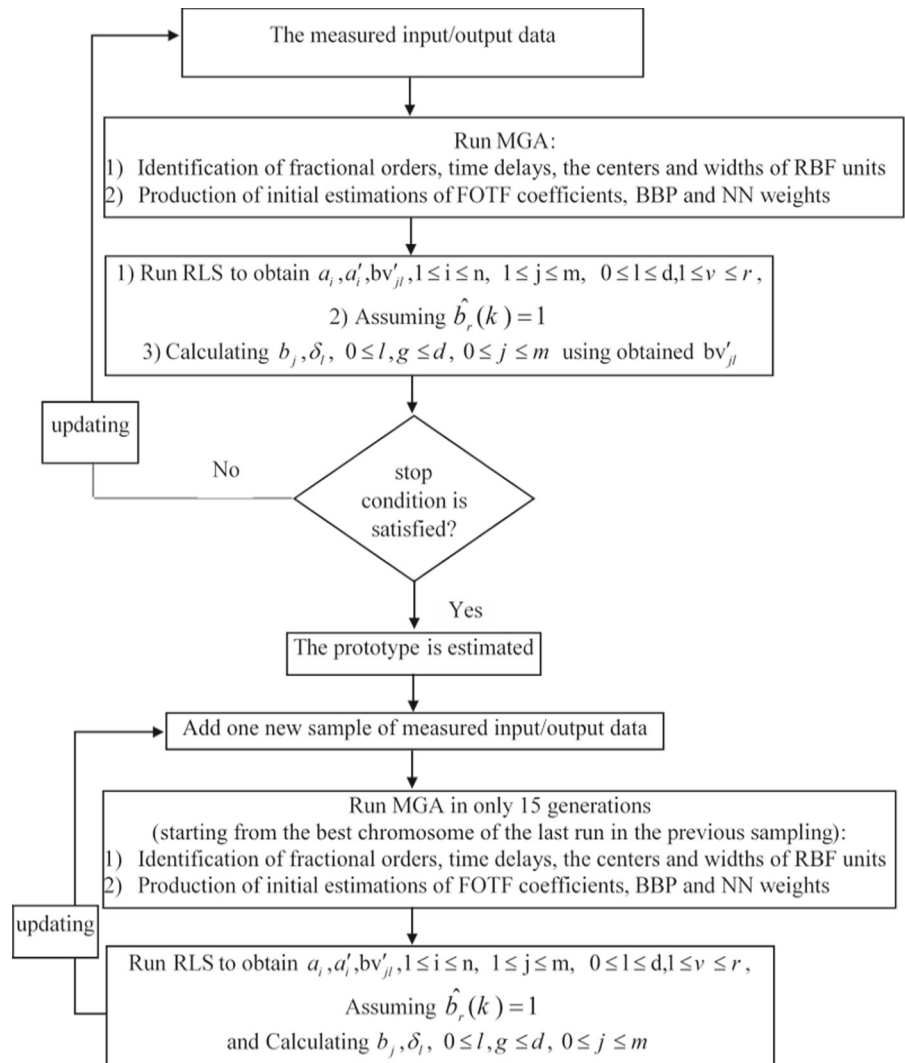
If we consider the measured data $u(t)$ and $y^*(t) = y(t) + p(t)$ where $p(t)$ is the disturbance signal, Eqs. (3–43) can be rewritten as follows:

$$y^*(k+1) = -\sum_{i=1}^{n} a'_i Y_i^*(k) +$$
$$+ \sum_{j=0}^{m_1} \sum_{l=0}^{d} b1'_{jl}(F1_{jl}(k)) + ... + \sum_{j=0}^{m_r} \sum_{l=0}^{d} br'_{jl}(Fr_{jl}(k)) + e(k+1) \tag{26}$$

while

$$Y_i^*(k) = \sum_{j=1}^{N} (-1)^j \binom{\alpha_i}{j} y^*(k+1-j) \tag{27}$$

Fig. 3 Online system identification block diagram

$$e(k+1) = p(k+1)$$
$$+ \sum_{i=1}^{n} a_i' \sum_{j=1}^{N} (-1)^j \binom{\alpha_i}{j} p(k+1-j)$$

$$(28)$$

Equation (26) is linear with respect to coefficients and can be expressed as follows:

$$y(k+1) = \theta^T \phi(k) + e(k+1) \qquad (29)$$

which:

$$\theta = \left[a_1', \ldots, a_n', b1_{00}', \ldots, b1_{0d}', \ldots, b1_{m_10}', \ldots, b1_{md}', \ldots, br_{00}', \ldots, \right.$$
$$\left. br_{0d}', \ldots, br_{m0}', \ldots, br_{md}' \right]^T$$

$$(30)$$

$$\phi(k) = \left[-Y_1^*(k), \ldots, -Y_n^*(k), F1_{00}(k), \ldots, F1_{0d}(k), \ldots, F1_{m0}(k), \ldots, \right.$$
$$\left. F1_{md}(k), \ldots, Fr_{00}(k), \ldots, Fr_{0d}(k), \ldots, Fr_{m0}(k), \ldots, Fr_{md}(k) \right]^T$$

$$(31)$$

The estimation vector $\hat{\theta}_k$ is obtained by minimizing the following quadratic least squares criterion:

$$\hat{\theta}_k = \arg\min_{\theta} \frac{1}{k} \sum_{i=1}^{k} \left[y(i) - \hat{y}(i, \theta) \right]^2 \qquad (32)$$

The solution to this problem can be obtained by the least squares estimation:

$$\hat{\theta}_k = \left[\sum_{i=1}^{k} \phi(i-1)\phi^T(i-1) \right]^{-1} \sum_{i=1}^{k} \phi(i-1)y(i) \tag{33}$$

provided that there is the inverse $\left[\sum_{i=1}^{k} \phi(i-1)\phi^T(i-1) \right]^{-1}$.

In order to provide the online identification possibility, the recursive algorithms are required. For this purpose, a recursive version of Eq. (33) is proposed, which can be written as follows:

$$\begin{cases} \hat{\theta}_{k+1} = \hat{\theta}_k + G_k\phi(k)\varepsilon(k+1) \\ G_{k+1} = G_k - \dfrac{G_k\phi(k)\phi^T(k)G_k}{1 + \phi^T(k)G_k\phi(k)} \\ \varepsilon(k+1) = \dfrac{y(k+1) - \hat{\theta}_k^T\phi(k)}{1 + \phi^T(k)G_k\phi(k)} \end{cases} \tag{34}$$

where the initial value of the adaptation gain matrix G_k is generally selected as:

$$G_0 = \frac{1}{\gamma}I; 0 < \gamma \ll 1 \tag{35}$$

For static systems with slow-varying parameters, a forgetting factor $\lambda, 0 < \lambda < 1$ can be considered. Therefore, the RLS algorithm is converted as:

$$\begin{cases} \hat{\theta}_{k+1} = \hat{\theta}_k + G_k\phi(k)\varepsilon(k+1) \\ G_{k+1} = \dfrac{1}{\lambda}\left[G_k - \dfrac{G_k\phi(k)\phi^T(k)G_k}{\lambda + \phi^T(k)G_k\phi(k)} \right] \\ \varepsilon(k+1) = \dfrac{y(k+1) - \hat{\theta}_k^T\phi(k)}{1 + \phi^T(k)G_k\phi(k)} \end{cases} \tag{36}$$

Using this recursive algorithm, the coefficients a_i', bv_{jl}' are determined; then, the values of a_i, b_j can be obtained using Eq. (24). Finally, Eq. (23) is used to obtain the estimated output values. The convergence of RLS algorithm in the reference [51] has been proven.

3.2 MISO Hammerstein model using RBFNN

As mentioned, in order to increase the accuracy of system dynamics investigation, time-delay information must be considered in practical processes or systems identification. In this case, the system is considered with time delays in the inputs. It is modeled as follows:

$$y(t) + \sum_{i=1}^{n} a_i D^{\alpha_i} y(t) = \sum_{j=0}^{m} b_j D^{\beta_j}(w_1(t-\gamma_1) + \ldots + w_r(t-\gamma_r))$$
$$w_k(t-\gamma_k) = f(u_k(t-\gamma_k)) \tag{37}$$

where $y(t)$ is system output, $w_k(t), k = 1, 2, \ldots, r$ are RBFNN outputs and the inputs to the linear dynamics section. $\gamma_k, k = 1, 2, \ldots, r$ are time delays and n, m are the input and output delays for the linear subsystem. Nonlinear static functions in this case are considered as follows:

$$w_k(t-\gamma_k) = f_k(u_k(t-\gamma_k))$$
$$= \omega_{0k} + \sum_{h=1}^{M} \omega_{hk}\phi_{hk}(u_k(t-\gamma_k)) \tag{38}$$

which $u_1, u_2, \ldots, u_r$ define inputs. As shown in the following equation, $\phi_{hk}(.), h = 1, \ldots, M, k = 1, \ldots, r$ are Gaussian functions related to kth input. Also, $\omega_{hk}(.), h = 1, \ldots, M, k = 1, \ldots, r$ are the connection weights from hth hidden node to the output and related to the kth input that must be determined. M is the number of RBFs.

$$\phi_{hk}(u_k(t)) = \exp\left\{ -\left(\|u_k(t) - c_{hk}\|^2 \Big/ d_{hk}^2 \right) \right\} \tag{39}$$

$c_{hk}, d_{hk}, h = 1, \ldots, M, k = 1, \ldots, r$ are the centers and widths of the hth RBF hidden unit associated with kth input. $\|.\|$ defines the Euclidean norm.

Also, the linear subsystem is considered as follows:

$$G(s) = \frac{B(s)}{A(s)} = \frac{b_m s^{\beta_m} + \cdots + b_1 s^{\beta_1} + b_0 s^{\beta_0}}{1 + a_n s^{\alpha_n} + \cdots + a_1 s^{\alpha_1}} \qquad (40)$$

which $a_1, \ldots, a_n, b_0, \ldots, b_m \in R$ are coefficients and $\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_m \in R$ are the fractional orders.

Considering Eqs. (38) to (40), the input–output relationship is equal to:

$$
\begin{aligned}
y(t) &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j} \left(\sum_{k=1}^{r} \sum_{h=0}^{M} (\omega_{hk} \phi_{hk}(u_k(t-\gamma))) \right) = \\
&= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j} (\omega_{01} \phi_{01}(u_1(t-\gamma)) + \ldots \\
&\quad + \omega_{M1} \phi_{M1}(u_1(t-\gamma)) + \ldots + \omega_{0r} \phi_{0r}(u_r(t-\gamma)) \\
&\quad + \ldots + \omega_{Mr} \phi_{Mr}(u_r(t-\gamma)))
\end{aligned}
\qquad (41)
$$

If the measured inputs and outputs are $u(t)$ and $y^*(t) = y(t) + q(t)$, respectively, which $q(t)$ is a random Gaussian noise with zero mean and σ^2 variance, Eq. (41) is rewritten as follows:

$$
\begin{aligned}
y^*(t) &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y^*(t) + \sum_{j=0}^{m} b_j D^{\beta_j} \left(\sum_{k=1}^{r} \sum_{h=0}^{M} (\omega_{hk} \phi_{hk}(u_k(t-\gamma_k))) \right) = \\
&= -\sum_{i=1}^{n} a_i D^{\alpha_i} y^*(t) + \sum_{j=0}^{m} b_j D^{\beta_j} (\omega_{01} \phi_{01}(u_1(t-\gamma_1)) + \ldots + \\
&\quad + \omega_{M1} \phi_{M1}(u_1(t-\gamma_1)) + \ldots + \omega_{0r} \phi_{0r}(u_r(t-\gamma_r)) + \ldots + \omega_{Mr} \phi_{Mr}(u_r(t-\gamma_r))) + q(t)
\end{aligned}
\qquad (42)
$$

Unknown parameters are divided into two subcategories: The first subset includes time delays $\gamma_k, k = 1, 2, \ldots, r$, fractional orders $\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_m$, and centers and widths of RBF units $d_{hk}, c_{hk}, h = 0, \ldots, M, k = 1, \ldots, r$. And the second category includes the coefficients of the transfer function $a_1, \ldots, a_n, b_0, \ldots, b_m$ and the connection weights $\omega_{hk}(\cdot), h = 1, \ldots, M, k = 1, \ldots, r$ from the jth hidden node to the output. In this paper, input/output data are divided into two parts. The first part is used to obtain both sets of unknown parameters using MGA. Then, using these estimations, RLS algorithm uses the I/O data second part to update and optimize the second subcategory of unknown parameters.

The standard RLS algorithm needs to set initial values for unknown parameters. In the standard method, these initial values are calculated using the batch least squares algorithm from multiple prototypes. In this method, the regressors dimension of the known parameters determines the number of samples required to determine the unique solution. In this paper, MGA is used for this task.

In order to use the output error estimation method, the input–output relationship (41) must be written in regression form:

$$y(t) = z^T(t)\theta + v(t) \qquad (43)$$

where z is known parameters including input–output data:

$$z(t) = \left[z_a^T(t), z_{b_0}^T(t), z_{b_1}^T(t), \ldots, z_{b_r}^T(t), 1 \right]^T \qquad (44)$$

$$z_a(t) = [-D^{\alpha_1} y^*(t), -D^{\alpha_2} y^*(t), \ldots, -D^{\alpha_n} y^*(t)]^T \qquad (45)$$

$$
z_{b_j}(t) = \left[D^{\beta_j} \phi_1(u_1(t-\gamma)), D^{\beta_j} \phi_2(u_2(t-\gamma)), \ldots, D^{\beta_j} \phi_r(u_r(t-\gamma)) \right] \qquad (46)
$$

That:

$$
\begin{aligned}
\phi_k(u_k(t-\gamma)) &= [\phi_{0k}(u_k(t-\gamma)), \ldots, \phi_{Mk}(u_k(t-\gamma))], k \\
&= 1, \ldots, r
\end{aligned}
\qquad (47)
$$

And θ shows unknown parameters:

$$\theta = \left[\theta_a^T, \theta_{b_0}^T, \theta_{b_1}^T, \ldots, \theta_{b_r}^T, \theta_{\omega_0} \right]^T \qquad (48)$$

Including:

$$\theta_a = [a_1, a_2, \ldots, a_n]^T \qquad (49)$$

$$\theta_{b_j} = \left[\theta_{b_{j,1}}, \theta_{b_{j,2}}, \ldots, \theta_{b_{j,M}}\right]^T = \left[b_j\omega_1, b_j\omega_2, \ldots, b_j\omega_M\right]^T \tag{50}$$

$$\theta_{\omega_0} = [b_r\omega_0] \tag{51}$$

Which in (50):

$$\omega_h = [\omega_{h1}, \omega_{h2}, \ldots, \omega_{hr}], h = 1, \ldots, M \tag{52}$$

Provided that $\left[\sum_{i=1}^{k} z(i-1)z^T(i-1)\right]^{-1}$ has a definable value, solving this problem using the LS estimation is:

$$\hat{\theta}_k = \left[\sum_{i=1}^{k} z(i-1)z^T(i-1)\right]^{-1} \sum_{i=1}^{k} z(i-1)y(i) \tag{53}$$

When the parameters matrix z is singular or poorly conditioned, its inverse calculation will be very difficult. To avoid the problem of inverse calculation in online identification, the recursive version of Eq. (53) is written as follows [51].

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + L(k)\varepsilon(k) \\ \varepsilon(k) = y(k) - z^T(k)\hat{\theta}(k-1) \\ P(k) = P(k-1) - \dfrac{P(k-1)z(k)z^T(k)P(k-1)}{1 + z^T(k)P(k-1)z(k)} \\ L(k) = \dfrac{P(k-1)z(k)}{1 + z^T(k)P(k-1)z(k)} \end{cases} \tag{54}$$

In general, the initial value of the adaptation gain matrix P is selected as follows [51]:

$$P_0 = \frac{1}{\eta}I; 0 < \eta \ll 1 \tag{55}$$

Without losing the generality, we assume that $\hat{b}_r(k) = 1$:

$$\begin{cases} \hat{\theta}_{b_r}(k) = [\hat{\omega}_1(k), \hat{\omega}_2(k), ..., \hat{\omega}_M(k)]^T \\ \hat{\theta}_{\omega_0}(k) = \hat{\omega}_0(k) \end{cases} \tag{56}$$

The unknown coefficients $a_i, i = 1, ..., n$ are obtained directly from $\hat{\theta}_a$. The $b_j, 0 \le j \le m$ values

are calculated using linear least squares, assuming $\hat{b}_r(k) = 1$ as the following:

$$\hat{b}_j(k) = \sum_{h=1}^{M} \hat{\theta}_{b_{r,h}}(k)\hat{\theta}_{b_{j,h}}(k) \bigg/ \sum_{h=1}^{M} \hat{\theta}_{b_{r,h}}^2(k), j = 0, 1, \ldots, m-1 \tag{57}$$

3.3 MISO Hammerstein model using MRBFNN

In this case, Eq. (41) is modified as:

$$\begin{aligned} y(t) &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j}\left(\sum_{k=1}^{r}\sum_{h=0}^{M} (\omega_{hk}\phi_k(u_k(t-\gamma_k)))\right) \\ &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j}(\omega_{01}\phi_1(u_1(t-\gamma_1)) + \ldots \\ &\quad + \omega_{M1}\phi_1(u_1(t-\gamma_1)) + \ldots + \omega_{0r}\phi_r(u_r(t-\gamma_r)) + \ldots + \omega_{Mr}\phi_r(u_r(t-\gamma_r))) \\ &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j}((\omega_{01} + \ldots + \omega_{M1})\phi_1(u_1(t-\gamma_1)) + \ldots \\ &\quad + (\omega_{0r} + \ldots + \omega_{Mr})\phi_r(u_r(t-\gamma_r))) \\ &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y(t) + \sum_{j=0}^{m} b_j D^{\beta_j}(\Omega_1\phi_1(u_1(t-\gamma_1)) + \ldots + \Omega_r\phi_r(u_r(t-\gamma_r))) \end{aligned} \tag{58}$$

If the measured inputs and outputs are $u(t)$ and $y^*(t) = y(t) + q(t)$, respectively, which $q(t)$ is a random Gaussian noise with zero mean and σ^2 variance, Eq. (58) is rewritten as follows:

$$\begin{aligned} y^*(t) &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y^*(t) + \sum_{j=0}^{m} b_j D^{\beta_j}\left(\sum_{k=1}^{r}\sum_{h=0}^{M} (\omega_{hk}\phi_{hk}(u_k(t-\gamma_k)))\right) \\ &= -\sum_{i=1}^{n} a_i D^{\alpha_i} y^*(t) + \sum_{j=0}^{m} b_j D^{\beta_j}(\omega_{01}\phi_{01}(u_1(t-\gamma_1)) + \ldots \\ &\quad + \omega_{M1}\phi_{M1}(u_1(t-\gamma_1)) + \ldots + \omega_{0r}\phi_{0r}(u_r(t-\gamma_r)) + \ldots + \omega_{Mr}\phi_{Mr}(u_r(t-\gamma_r))) + q(t) \end{aligned} \tag{59}$$

And the relation (46) is modified as:

$$\begin{aligned} z_{b_j}(t) &= \left[D^{\beta_j}\phi_1(u_1(t-\gamma)), D^{\beta_j}\phi_2(u_2(t-\gamma)), \right. \\ &\quad \left. \ldots, D^{\beta_j}\phi_r(u_r(t-\gamma))\right]^T \end{aligned} \tag{60}$$

In this case, time delays $\gamma_k, k = 1, 2, \ldots, r$, fractional orders $\alpha_1, \ldots, \alpha_n, \beta_0, \ldots, \beta_m$, and centers and widths of RBF units $d_{hk}, c_{hk}, h = 0, \ldots, M, k = 1, \ldots, r$ are identified using MGA. And an initial estimation for the coefficients of the transfer function $a_1, \ldots, a_n, b_0, \ldots, b_m$ and the connection weights from

the jth hidden node to the output $\omega_{hk}(\cdot), h = 1, \ldots, M, k = 1, \ldots, r$ are obtained using this evolutionary algorithm. Then, these estimations are updated and optimized using RLS algorithm.

In comparison with classic RBFNN which the number of unknown parameters is equal to $3M + 1$ (i.e., $c_{hk}, d_{hk}, \omega_{hk}, \gamma_k$ and $h = 1, \ldots, M$ where M is the data samples number), their number in modified RBFNN is equal to 4 (i.e. $d_k, c_k, \Omega_k, \gamma_k$).

4 Simulation results

The main reason for using the recursive method is the ability to generalize to online identification. In order to demonstrate the ability of online identification, two examples have been considered:

1. Hammerstein model with piecewise nonlinear characteristic such as dead-zone characteristic [52]
2. Continuous-time linear parameter-varying (CT LPV) nonlinear benchmark system [53]

Both systems have been identified online using all three presented structures. One hundred samples were used to estimate the prototype. Then, by adding any input/output data in online mode, the model is updated using RLS. In the online mode, the MGA is executed only once at the beginning of the detection process and then only the LS algorithm is executed recursively. Therefore, all that is required in each instance is the recursive formula of the LS algorithm.

In addition, it is considered a case in which with adding each input / output data, MGA runs in a small number of generations, i.e., 15 generations. In this case, the MGA updates all the unknown parameters in each sampling starting from the best chromosome of the last run in the previous sampling (instead of starting from random chromosomes). Then, RLS goes one step further and updates the connection coefficients and weights.

4.1 Hammerstein model with piecewise nonlinear characteristic

In this Hammerstein model, the recursive relation forms the linear dynamic part:

$$y(k) = 1.6961y(k - 1) - 0.8651y(k - 2) \\ + 0.5895h(k - 1) + 0.4701h(k - 2) \quad (61)$$

And the discrete nonlinear part is described by the following equation:

$$N(u) = \begin{cases} u + 0.28 & u \le -0.28 \\ 0 & -0.28 \le u \le 0.28 \\ u - 0.28 & u \ge 0.28 \end{cases} \quad (62)$$

Actual output and estimated output in the last sample (using 200 sampled data) for the Hammerstein model with Bezier–Bernstein polynomials in Fig. 4, for the Hammerstein model with RBFNN in Fig. 5 and with MRBFNN, is shown in Fig. 6. The corresponding estimation error using the three nonlinear functions Bezier–Bernstein, RBFNN and MRBFNN is presented in Figs. 7, 8 and 9, respectively.
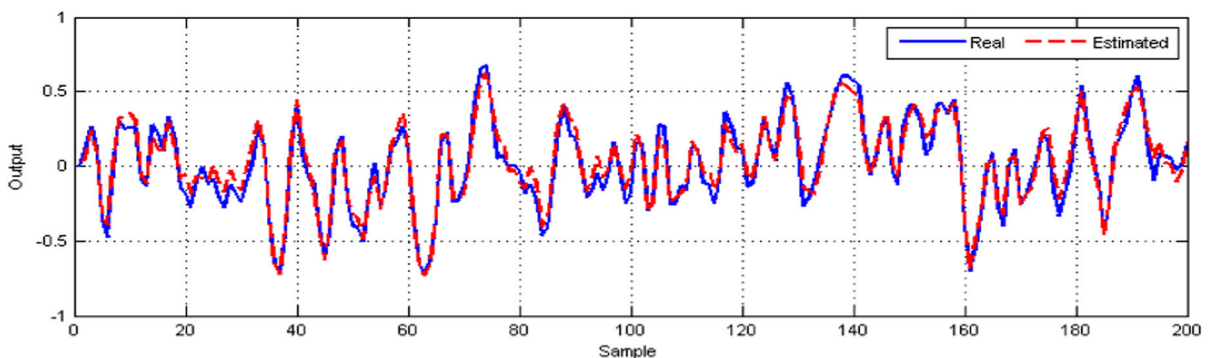
Fig. 4 The actual (blue line) and estimated (red dashed line) outputs using Bezier–Bernstein polynomial—Example 1. (Color figure online)
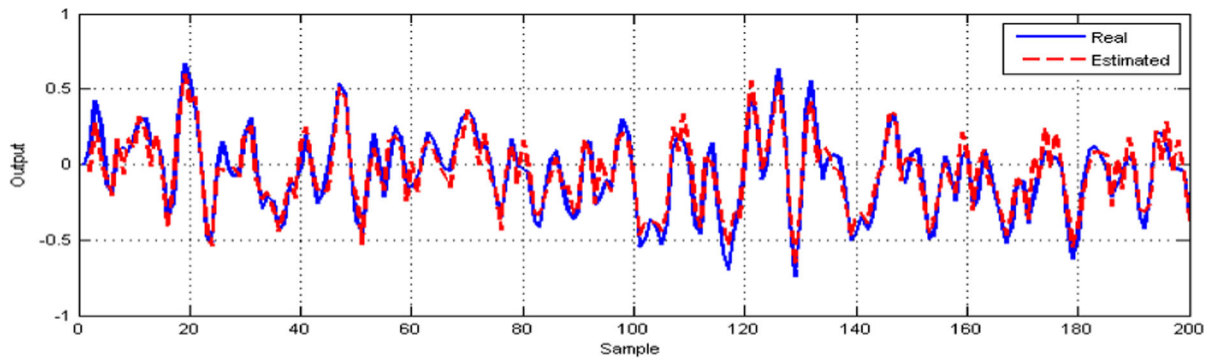
Fig. 5 The actual (blue line) and estimated (red dashed line) outputs using RBFNN—Example 1. (Color figure online)
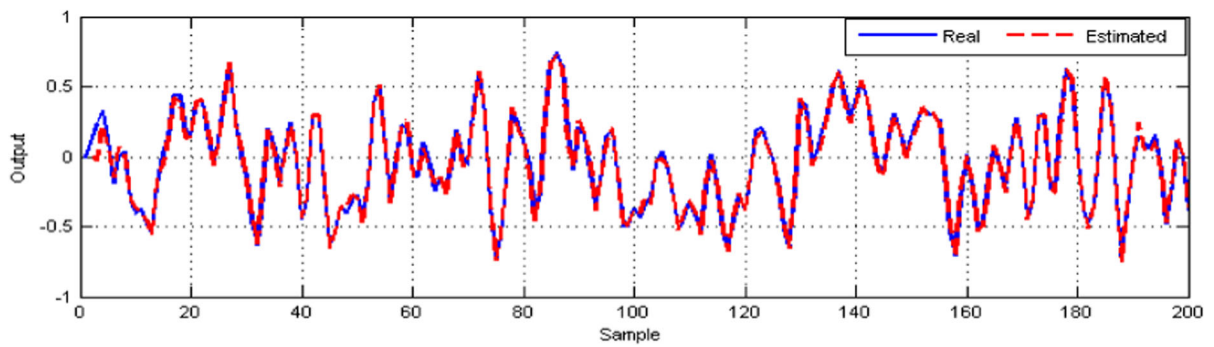
Fig. 6 The actual (blue line) and estimated (red dashed line) outputs using MRBFNN—Example 1. (Color figure online)
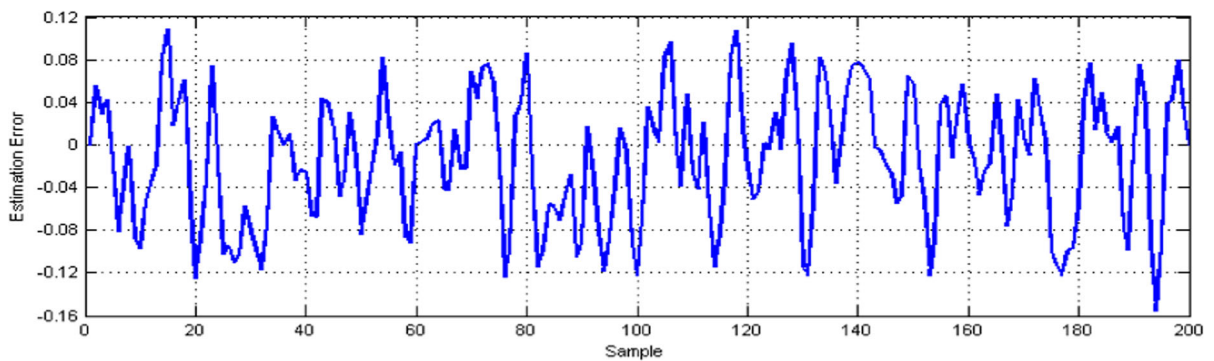
Fig. 7 The estimation error using BBP—Example 1
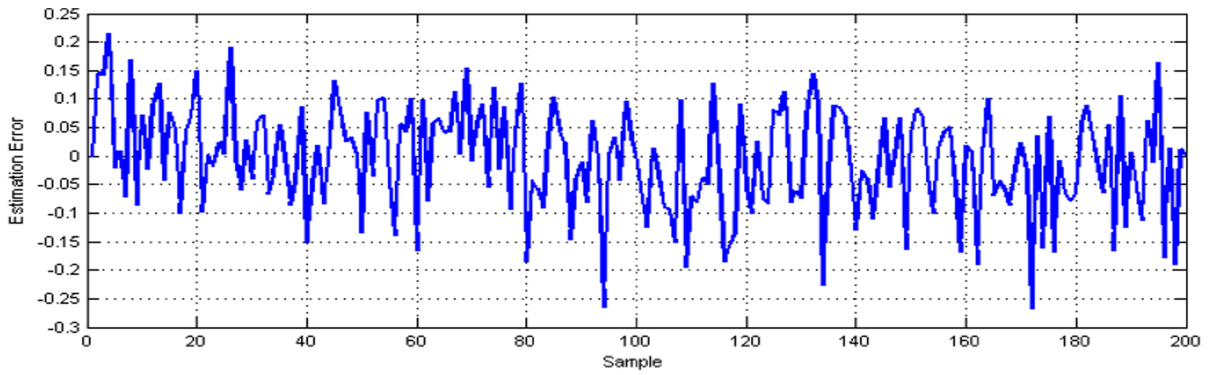
Fig. 8 The estimation error using RBFNN—Example 1
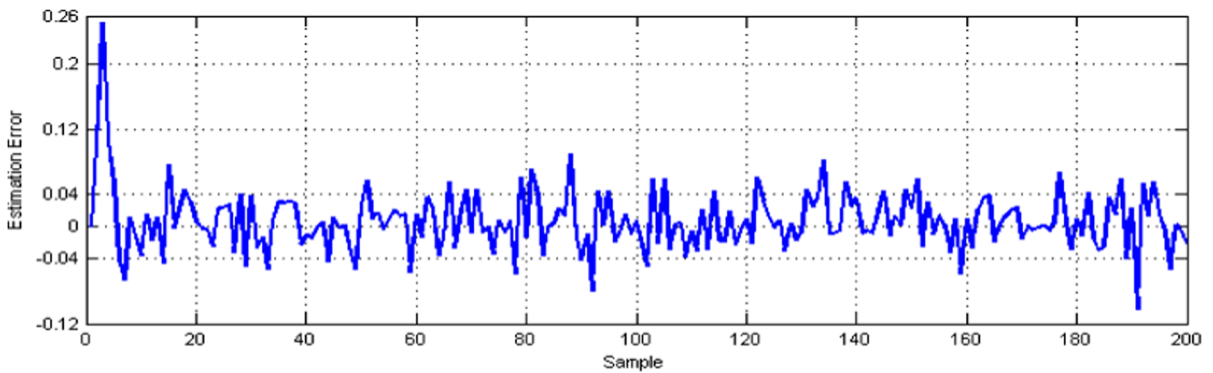
Fig. 9 The estimation error using MRBFNN—Example 1

In order to compare the estimation accuracy of three different structures, the changes of estimation error during the identification process (in different samples) in these three structures are shown in Figs. 10, 11 and 12, and the values of these errors are given in Table 1.
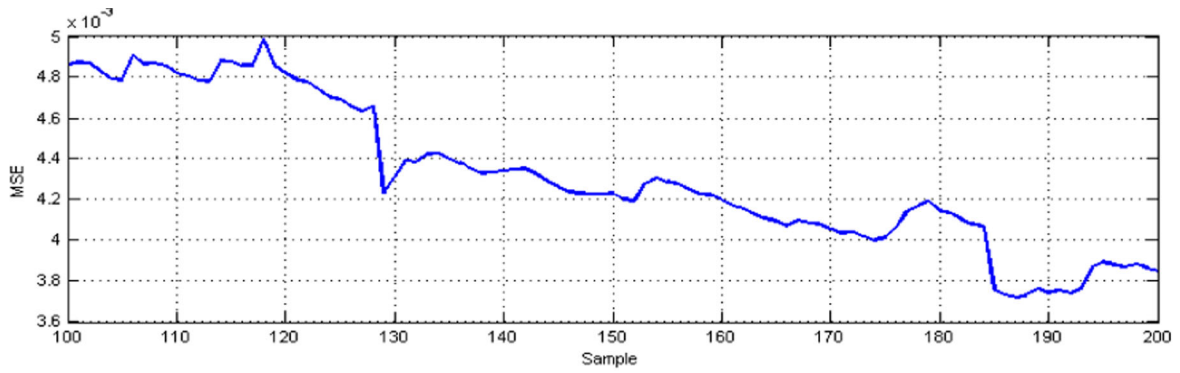
(a) BBP
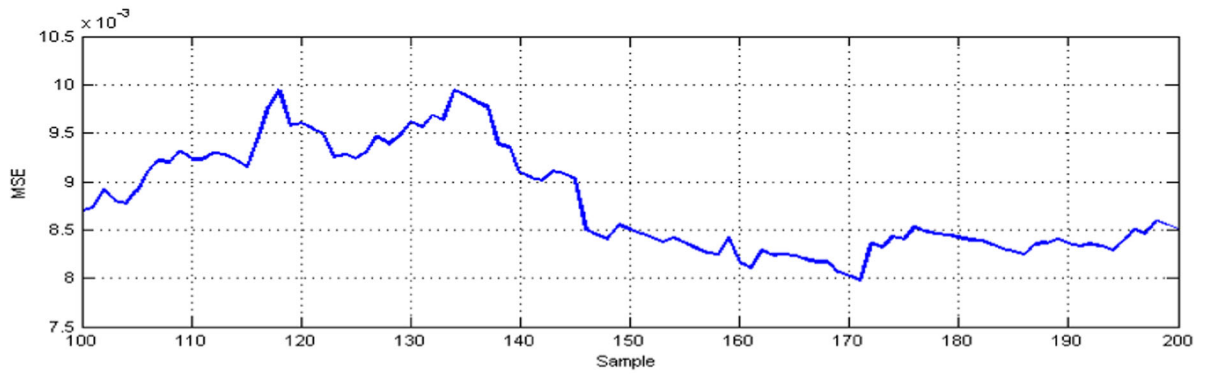
(b) RBFNN
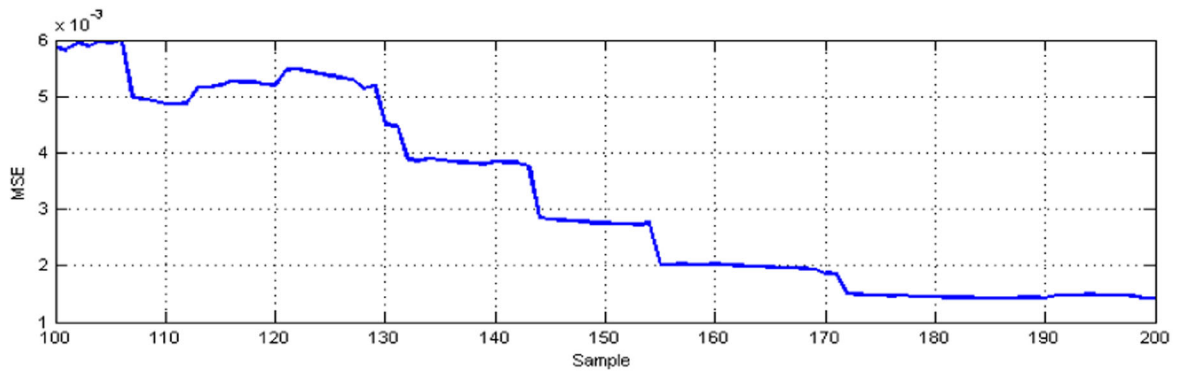
(c) MRBFNN

Fig. 10 The MSE estimation error with high noise using three different structures—Example 1
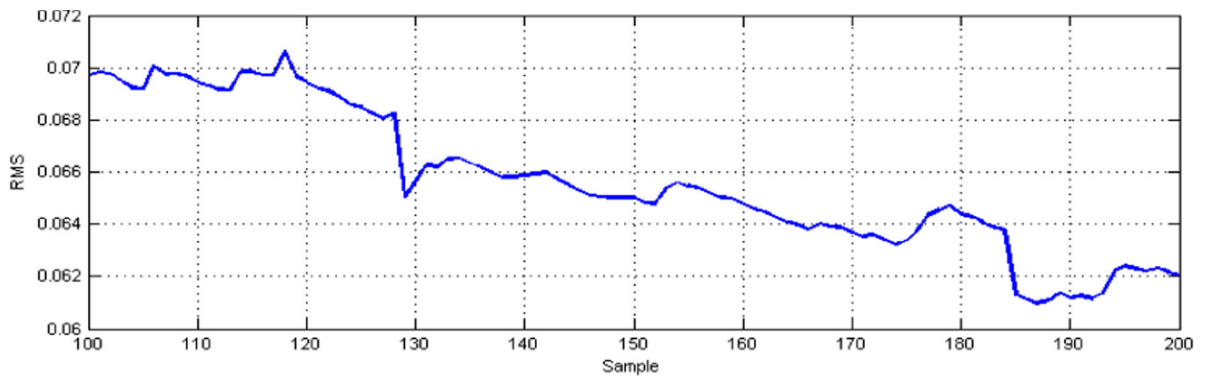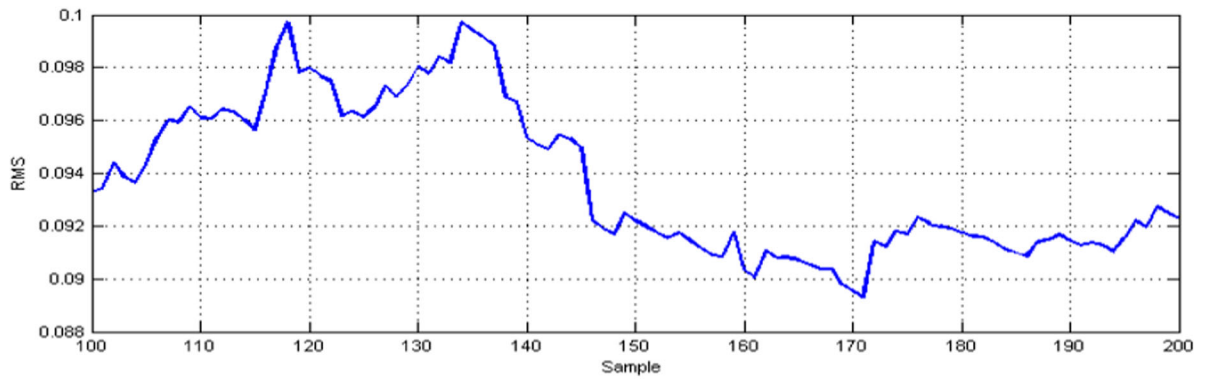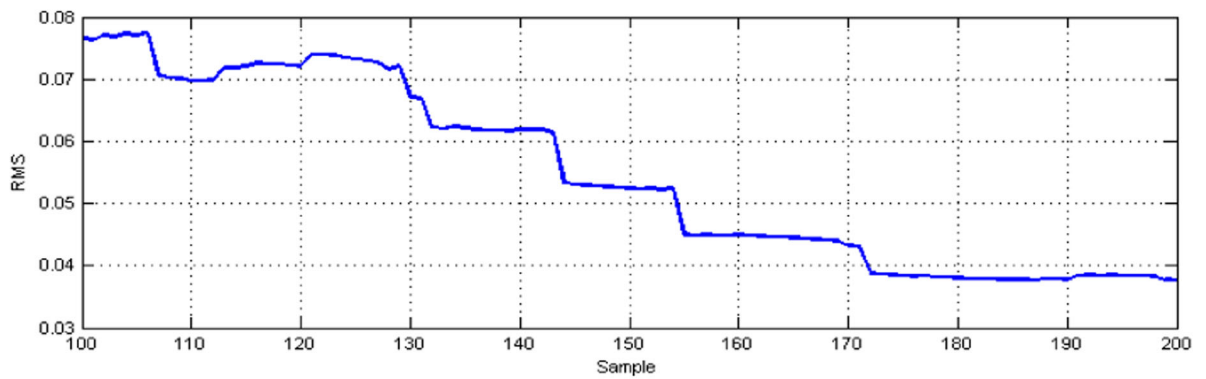
(a) BBP

(b) RBFNN

(c) MRBFNN

Fig. 11 The RMS estimation error with high noise using three different structures—Example 1
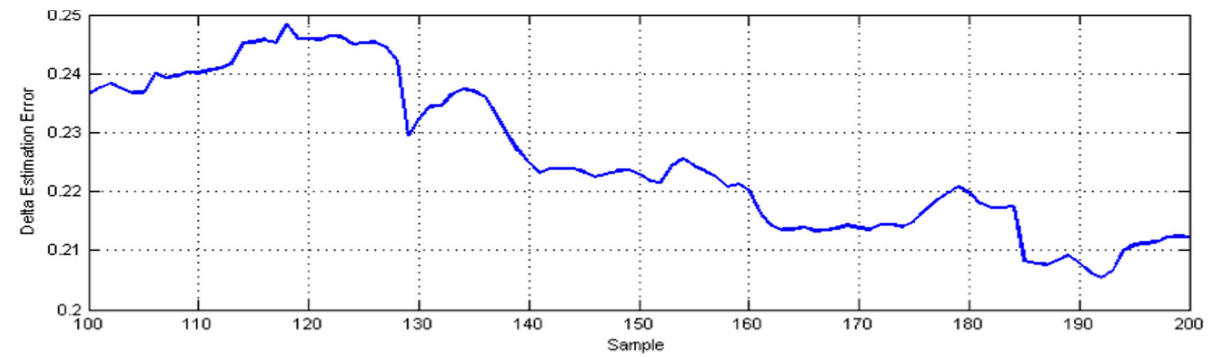
(a) BBP
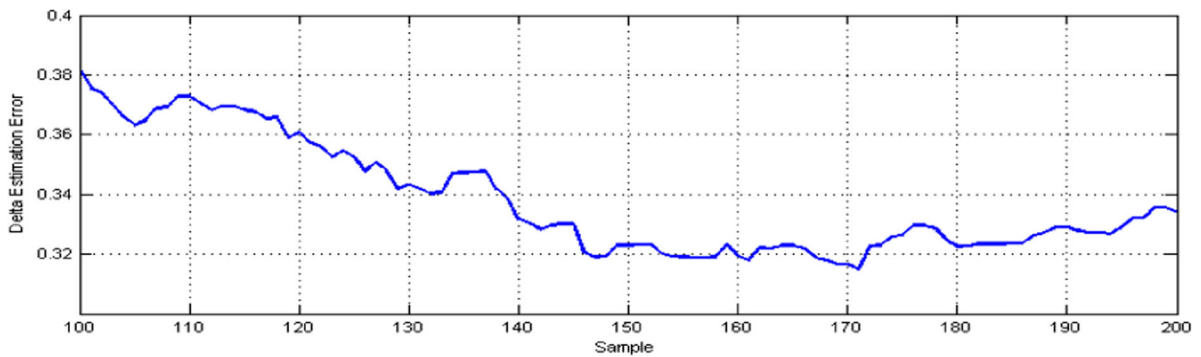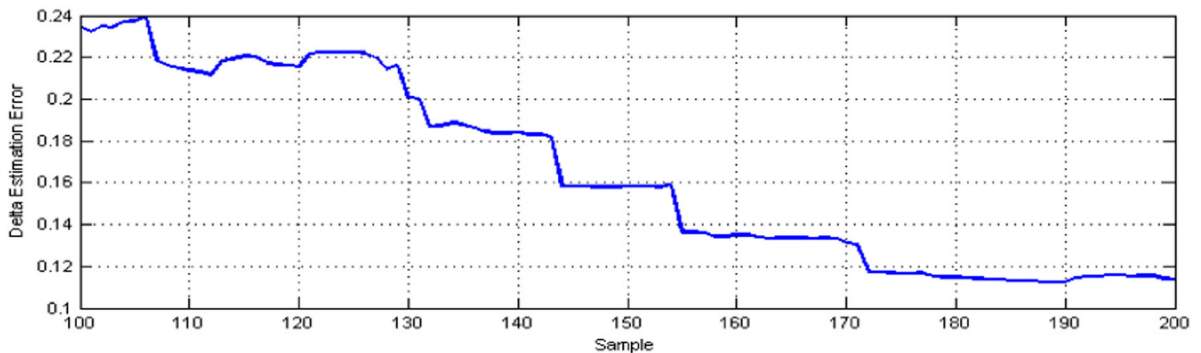
(b) RBFNN

(c) MRBFNN

Fig. 12 The δ estimation error with high noise using three different structures—Example 1

Table 1 The estimation error comparison between three presented structures—Example 1

Nonlinear static part	MSE	RMS	δ
BBP	0.0038477	0.06203	0.21222
RBFNN	0.0088902	0.094288	0.29976
MRBFNN	0.0014219	0.037708	0.11377

The comparison of the estimation errors in Table 1 and corresponding Figs. 10, 11 to 12 shows the superiority of identification accuracy using the MRBFNN in Hammerstein model over the two other structures.

Also, in order to compare the parameters convergence speed in three different structures, the variations of the identified parameters for the Hammerstein
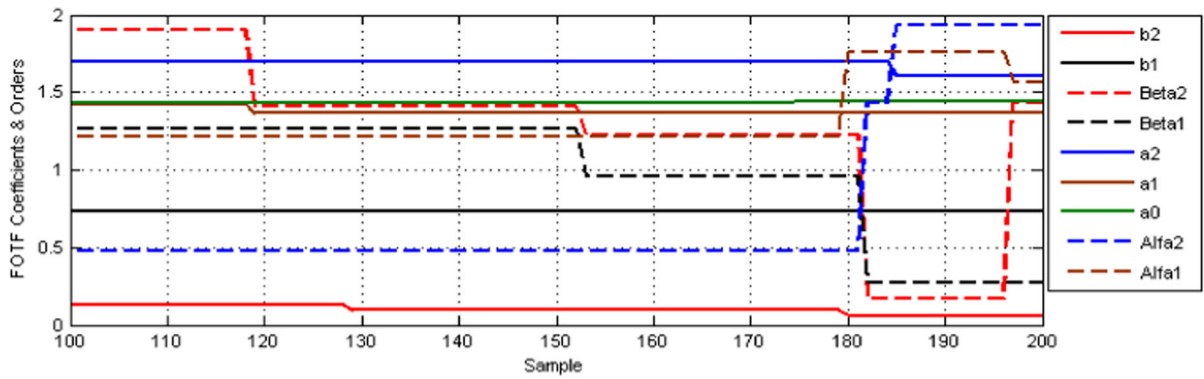
Fig. 13 The estimation of transfer function coefficients and fractional orders using BBP—Example 1
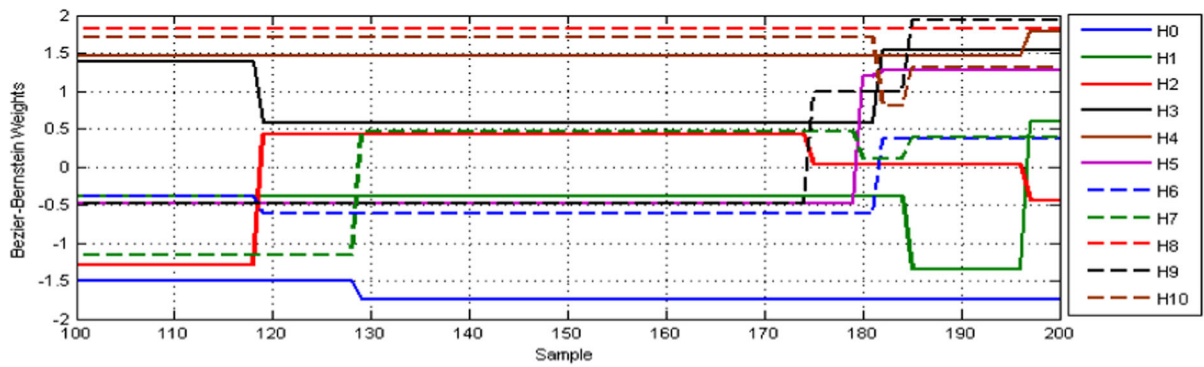
Fig. 14 The estimation of polynomial weights using BBP—Example 1

model with Bezier–Bernstein polynomials in Figs. 13 and 14, for the Hammerstein model with RBFNN in Figs. 15, 16 and 17 and for the Hammerstein model with MRBFNN are presented in Figs. 18, 19 and 20. The comparison of the figures shows that the identification process of the Hammerstein model with MRBFNN has the highest convergence speed due to the smaller number of unknown parameters identified between the three proposed structures.
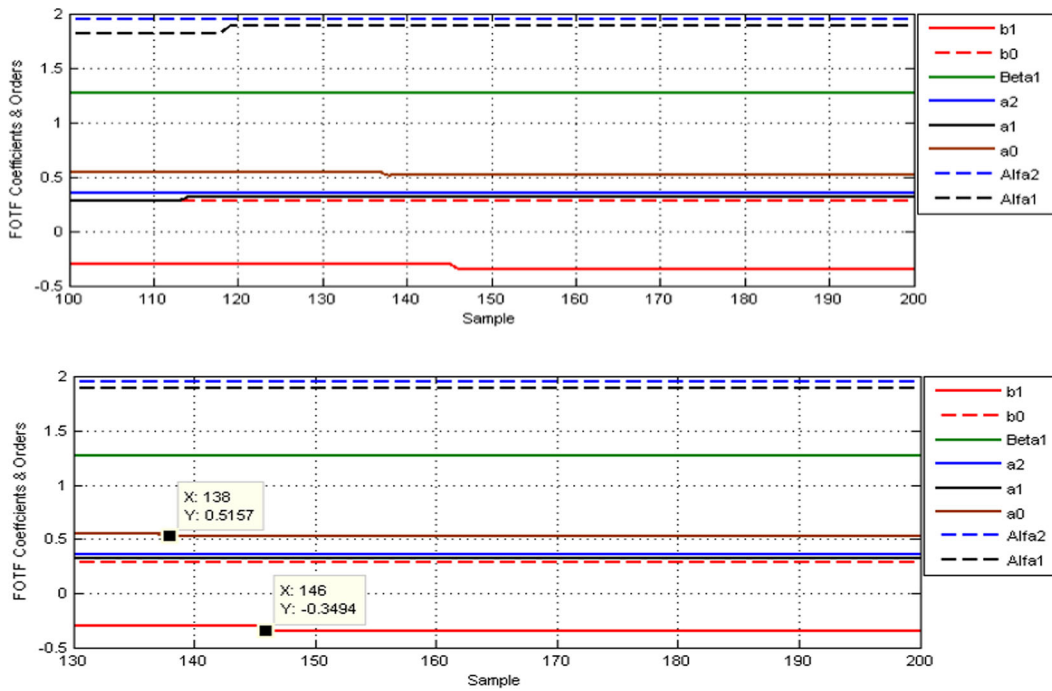
Fig. 15 The estimation of transfer function coefficients and fractional orders using RBFNN—Example 1
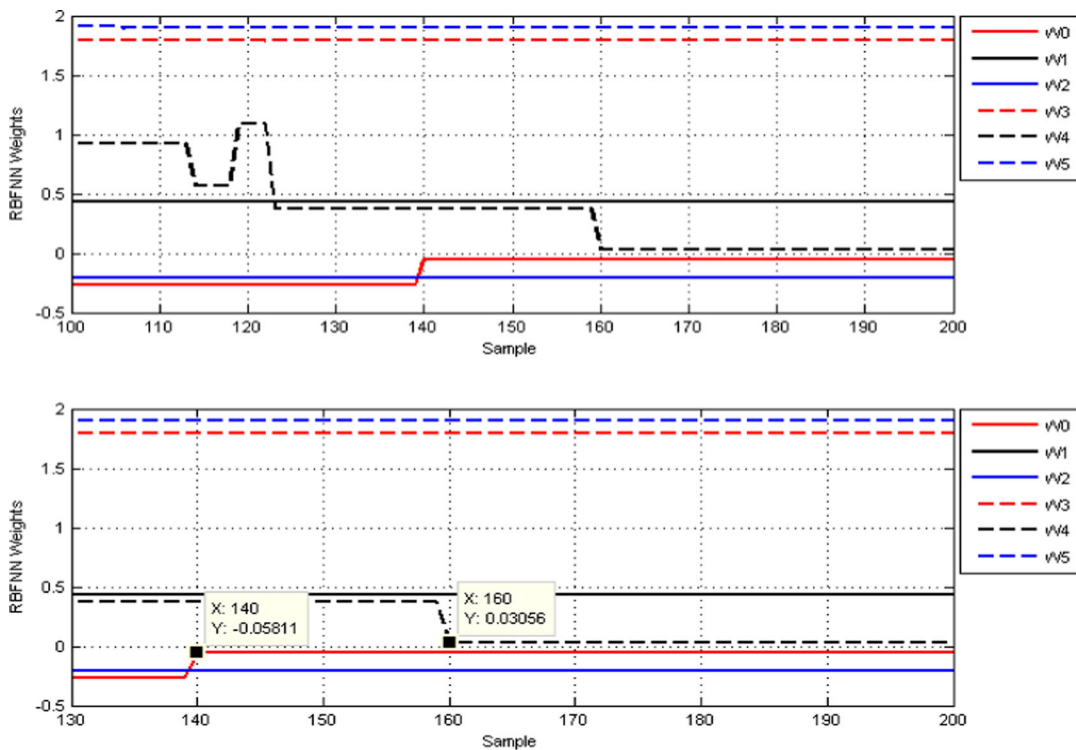
Fig. 16 The estimation of NN weights using RBFNN—Example 1
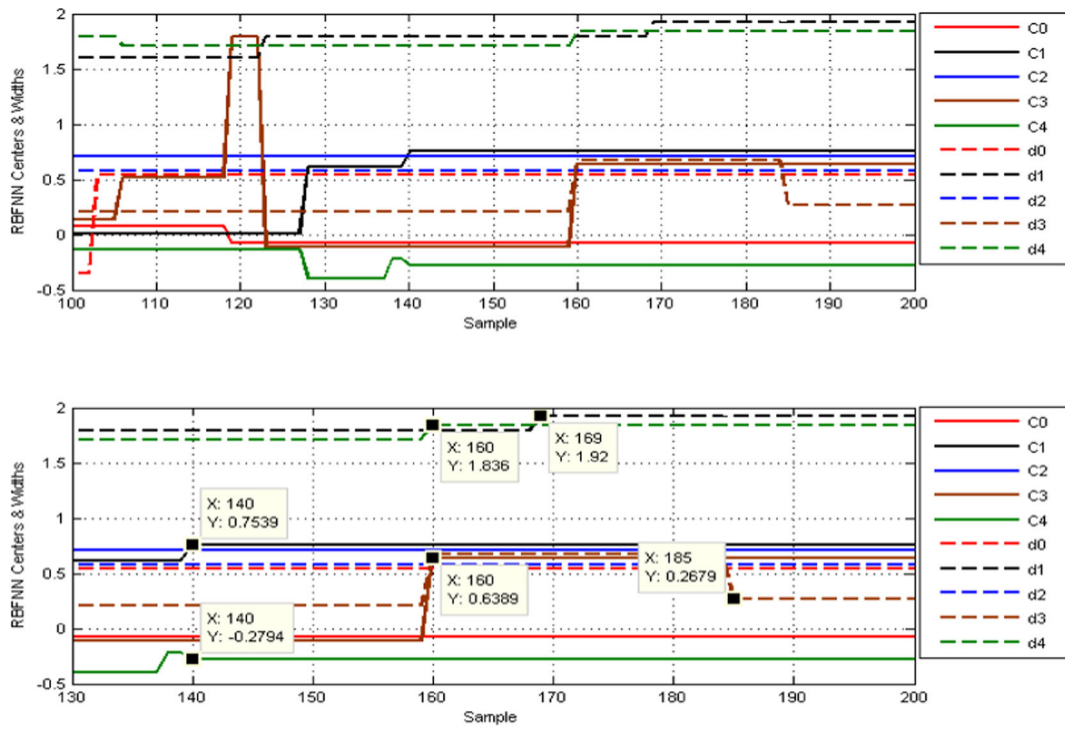
Fig. 17 The estimation of NN centers and widths using RBFNN—Example 1
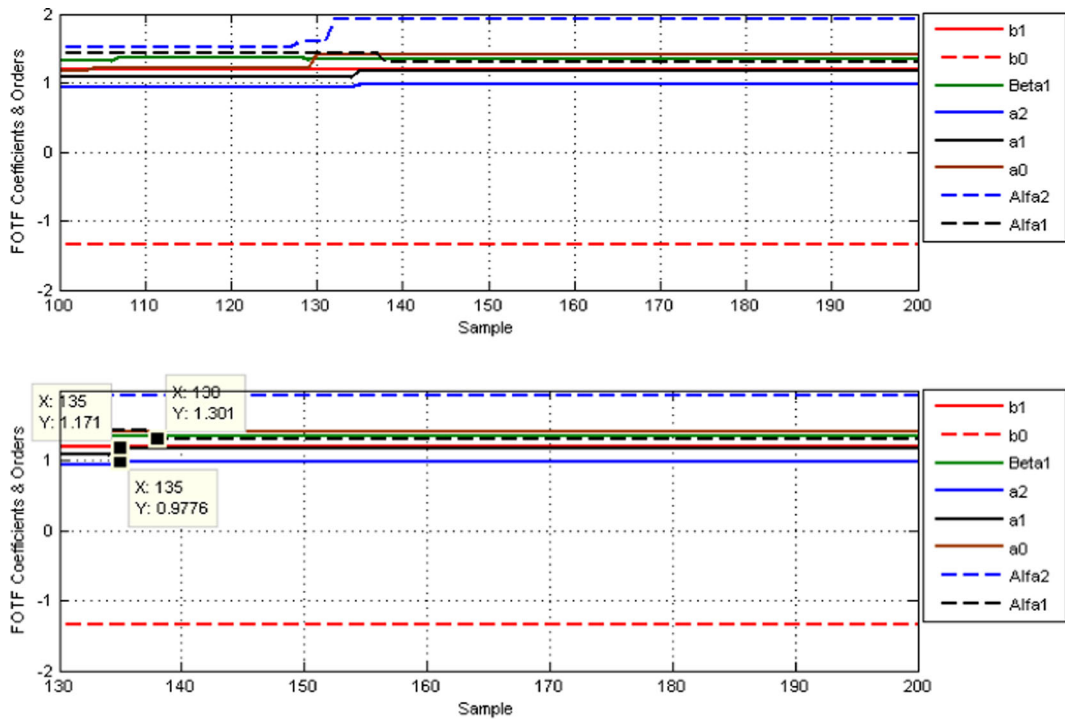
Fig. 18 The estimation of transfer function coefficients and fractional orders using MRBFNN—Example 1
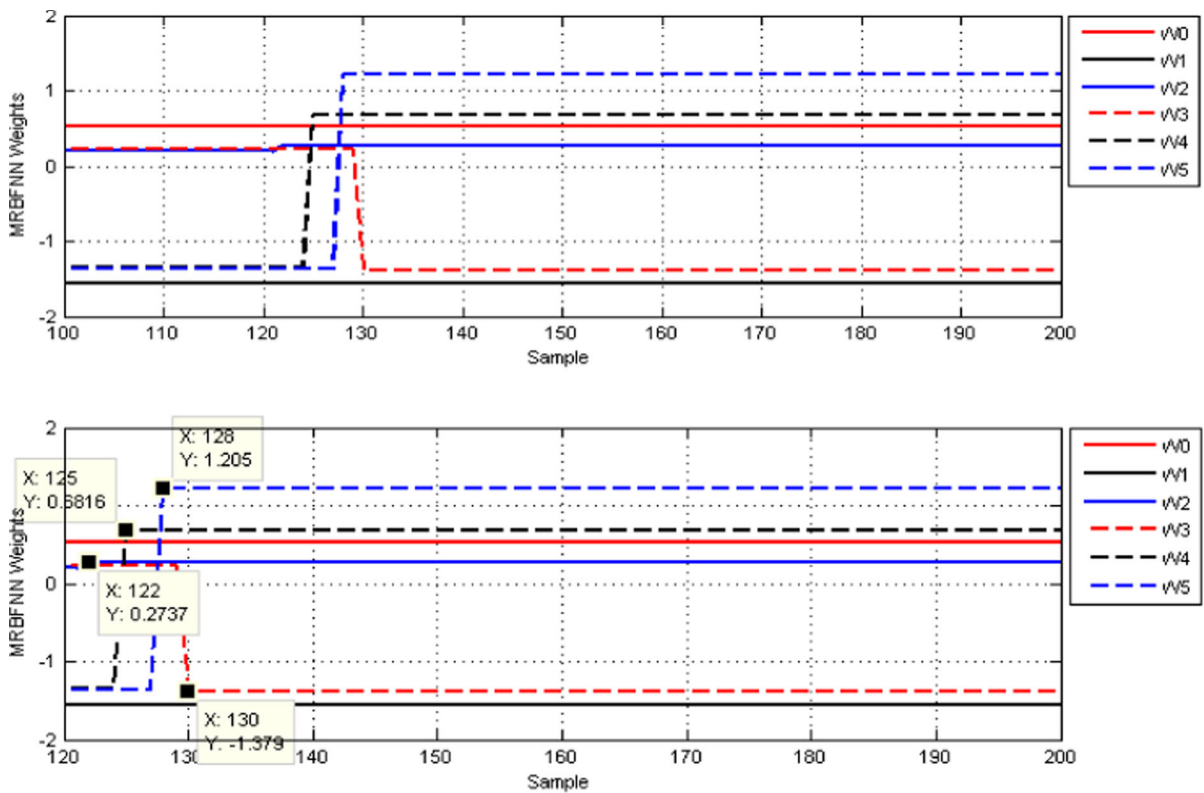
Fig. 19 The estimation of NN weights using MRBFNN—Example 1
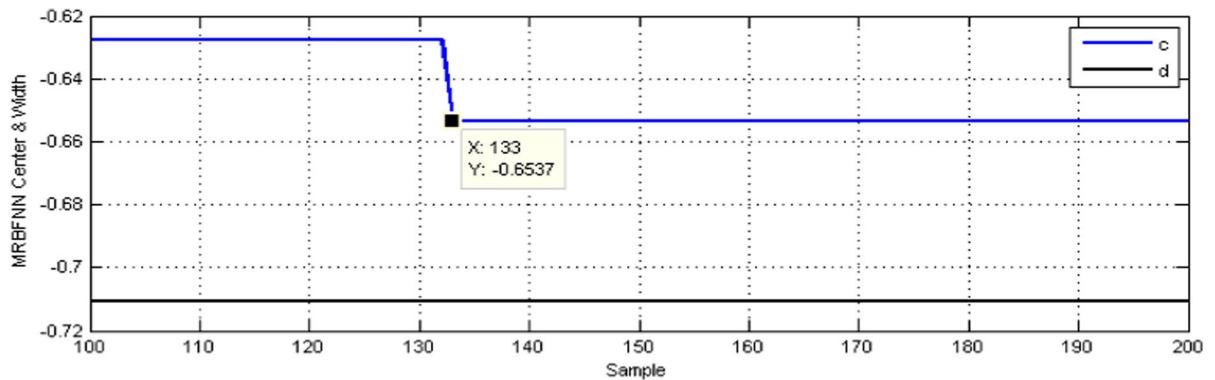
Fig. 20 The estimation of NN centers and widths using MRBFNN—Example 1

4.2 Continuous-time linear parameter-varying (CT LPV) nonlinear benchmark system

The second system is a benchmark problem proposed by Rao and Garnier in [54]. This problem is inspired by the "moving pole" parameter-varying system, a fourth-order system with non-minimum phase freezing dynamics and complex poles pair dependent to p parameter. This system is defined as follows:

$$\begin{cases} A_0(d,p)\chi_0(t) = B_0(d,p)u(t) \\ y(t) = \chi_0(t) + v_0(t) \end{cases} \quad (63)$$

where the d operator is a time derivative, p is a time-dependent programming variable, A_0 and B_0 are polynomials of d with coefficients a_i^0 and b_i^0. These coefficients are functions of p; v_0 is a process with semi-static noise with limited and unrelated spectral density as follows:

$$v_0(t) = H_0(q)e_0(t) = \frac{C_0(q^{-1})}{D_0(q^{-1})}e_0(t) \quad (64)$$

The $e_0(t)$ is a white noise process with a mean of zero, q^{-1} is the backward time shift operator, C_0 and D_0 are polynomials with constant coefficients. The values of the parameters used in this paper are as follows [53]:

$$\begin{cases} A_0(d,p) = d^4 + a_1^0(p)d^3 + a_2^0(p)d^2 + a_3^0(p)d + a_4^0(p) \\ B_0(d,p) = b_0^0(p)d + b_1^0(p) \\ H_0(q) = \dfrac{1}{1 - q^{-1} + 0.2q^{-2}} \end{cases} \quad (65)$$

For coefficients:

$$\begin{aligned} &a_1^0(p) = 5 + 0.25p & &a_2^0(p) = 408 + 3p + 0.25p^2 \\ &a_3^0(p) = 416 + 108p + p^2 & &a_4^0(p) = 1600 + 800p + 100p^2 \\ &b_0^0(p) = -6400 - 3200p - 400p^2 & &b_1^0(p) = 1600 + 800p + 100p^2 \end{aligned}$$

$$(66)$$

The input signal is a uniform distribution sequence in the interval $[-1, 1]$, p is selected as $p(t) = \sin(\pi t)$, and the sampling time is 1ms.

Actual output and estimated output in the last sample (using 500 sampled data) for the Hammerstein model with Buzzer–Bernstein polynomials are shown in Fig. 21, for the Hammerstein model with RBFNN in Fig. 22 and for the MRBFNN in Fig. 23. The corresponding estimation error using the three nonlinear functions Bezier–Bernstein, RBFNN and MRBFNN is presented in Figs. 24, 25 and 26, respectively.
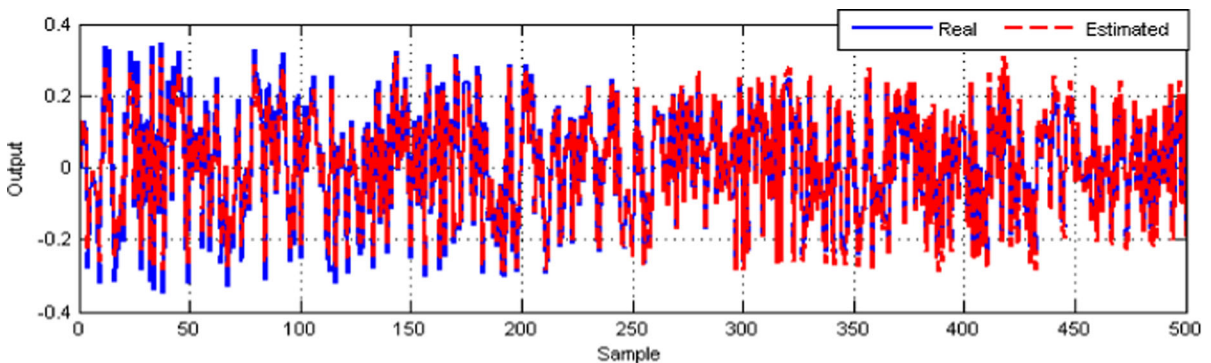
Fig. 21 The actual (blue line) and estimated (red dashed line) outputs using BBP—example 2. (Color figure online)
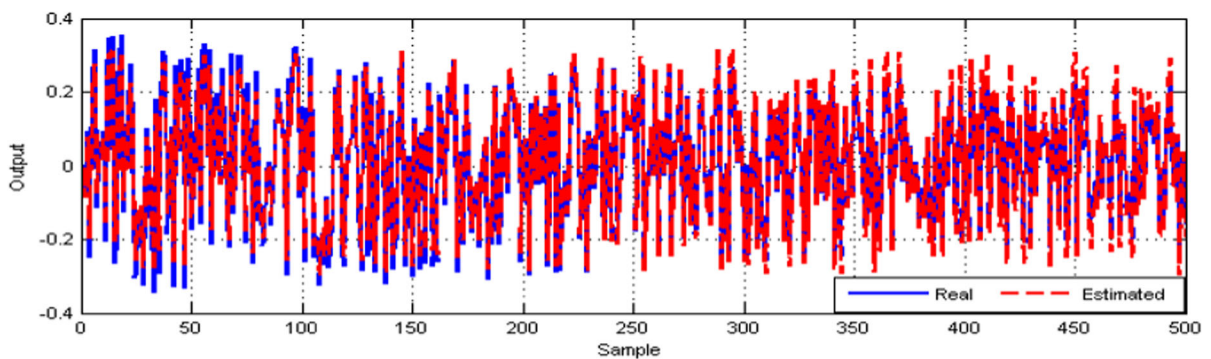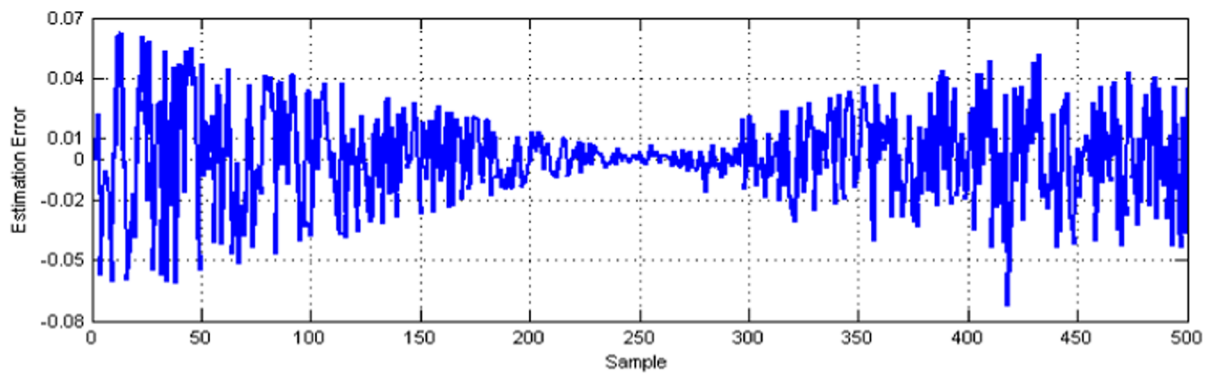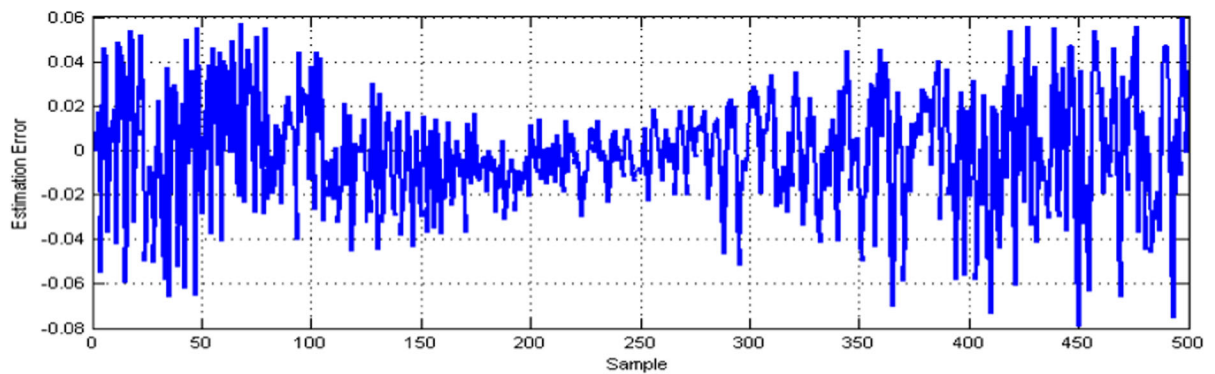

Fig. 22 The actual (blue line) and estimated (red dashed line) outputs using RBFNN—example 2. (Color figure online)
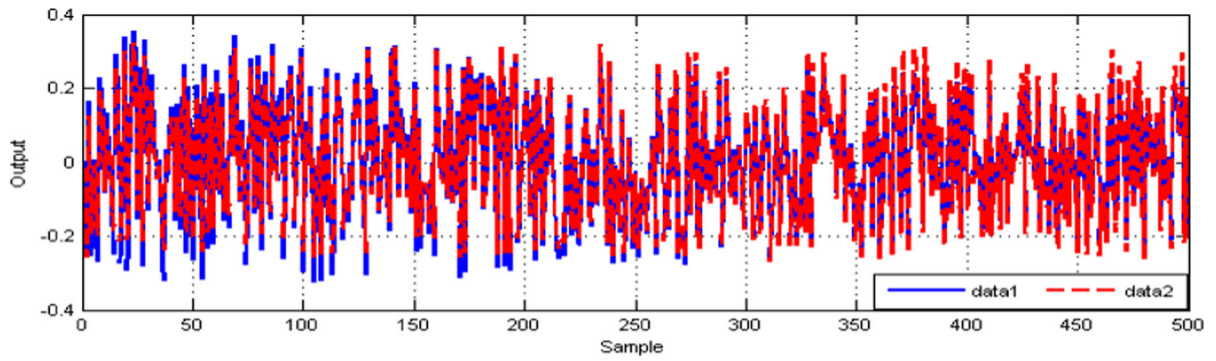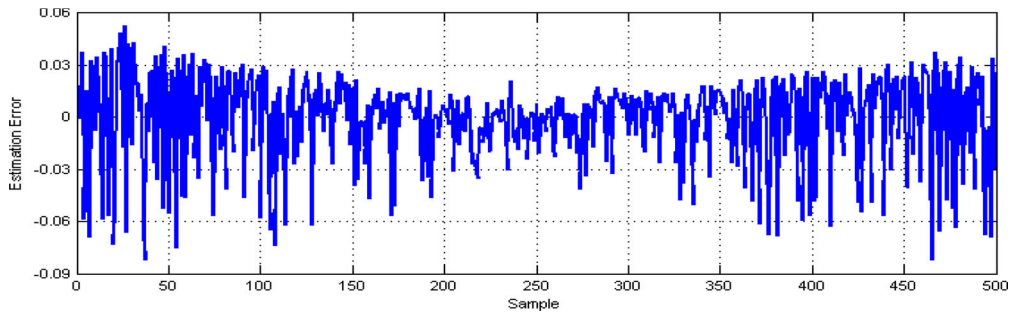
Fig. 23 The actual (blue line) and estimated (red dashed line) outputs using MRBFNN—example 2. (Color figure online)

**Fig. 24** The estimation error using BBP—Example 2

**Fig. 25** The estimation error using RBFNN—Example 2

Fig. 26 The estimation error using MRBFNN—Example 2

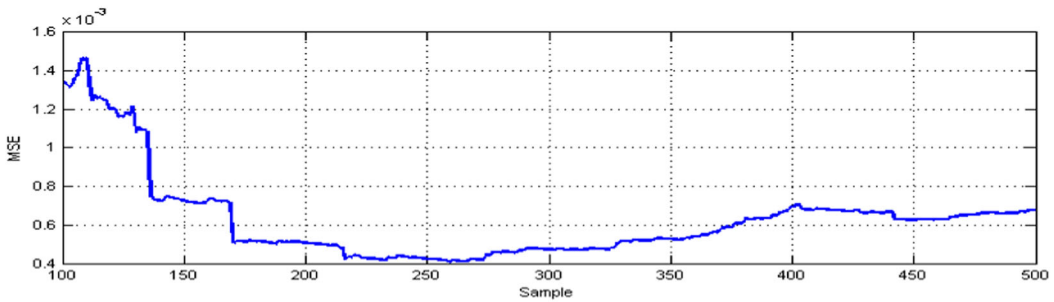In order to compare the accuracy of estimation with three different structures, the variations of estimation error during the identification process (in different samples) in these three structures are shown in Figs. 27, 28 and 29 and the values of these errors are given in Table 2.
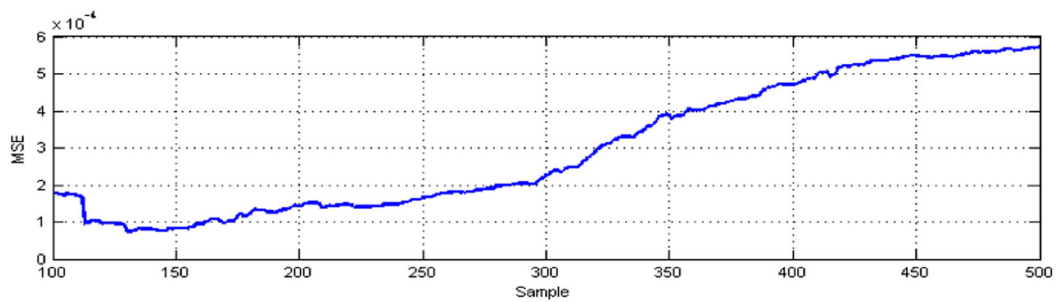
(a) BBP
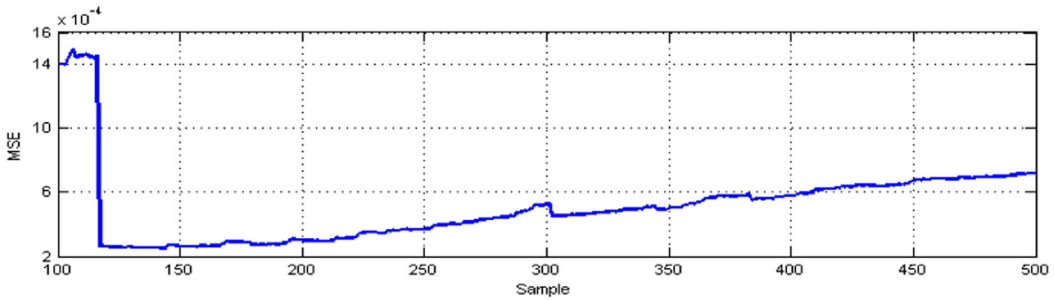
(b) RBFNN

(c) MRBFNN

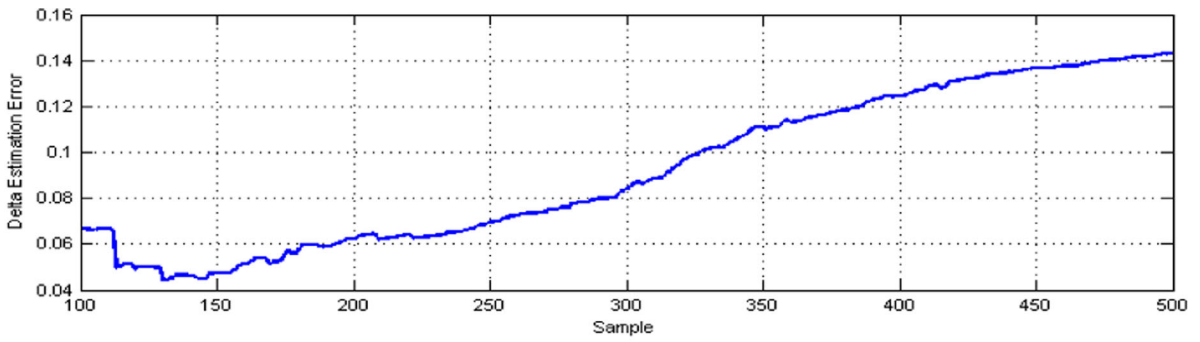**Fig. 27** The MSE estimation error with high noise using three different structures—Example 2
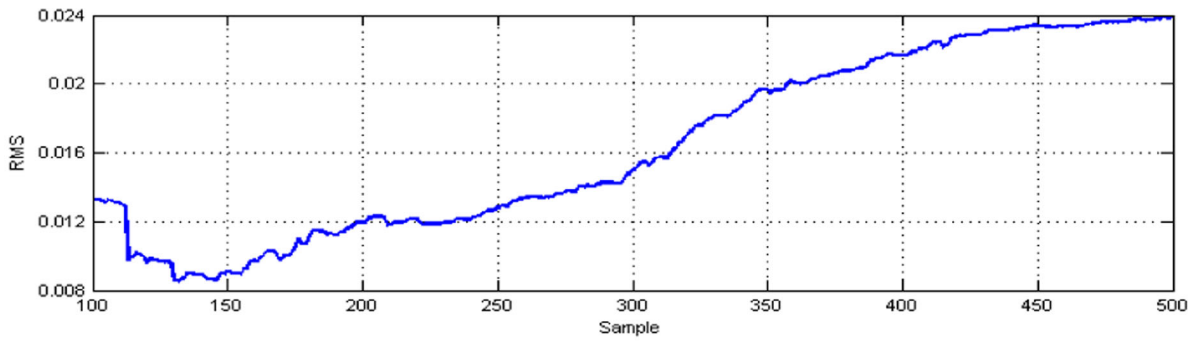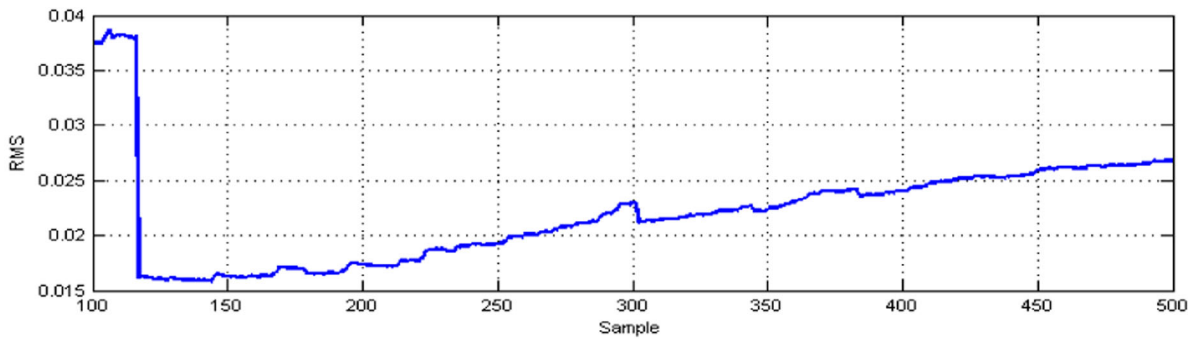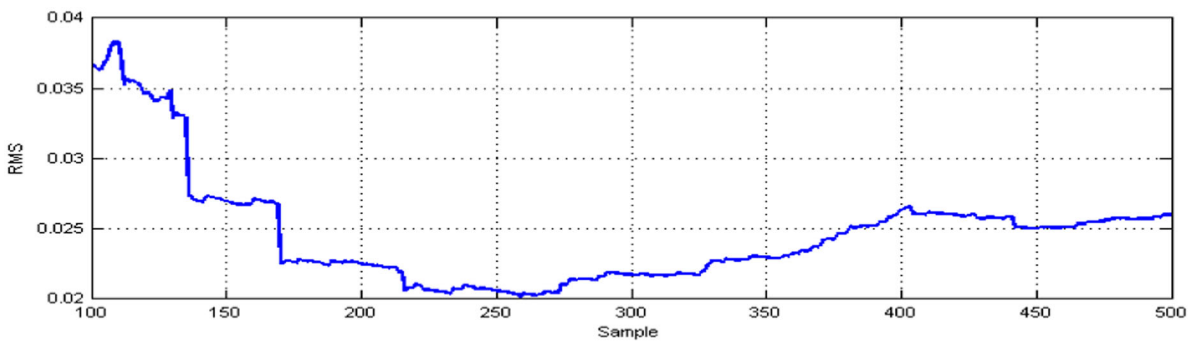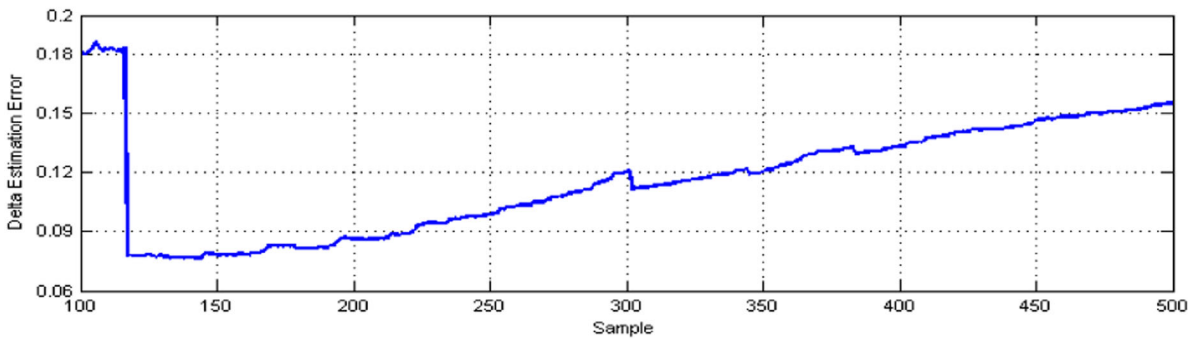
(a) BBP

(b) RBFNN

(c) MRBFNN

Fig. 28 The RMS estimation error with high noise using three different structures—Example 2
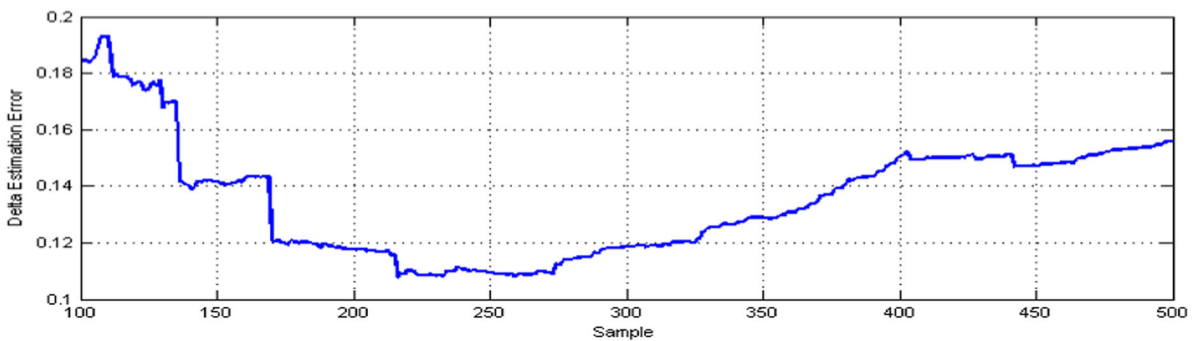
(a) BBP

(b) RBFNN

(c) MRBFNN

Fig. 29 The δ estimation error with high noise using three different structures—Example 2

**Table 2** The estimation error comparison between three presented structures—Example 2

Nonlinear static part	MSE	RMS	δ
BBP	0.00057078	0.023891	0.14321
RBFNN	0.00071952	0.026824	0.15532
MRBFNN	0.00067751	0.026029	0.15569

The comparison of the results of estimation errors in Table 2 and the corresponding Figs. 26, 27 and 28 shows the superiority of identification accuracy in the Hammerstein model using the MRBFNN compared to the Hammerstein model with the classic RBFNN.
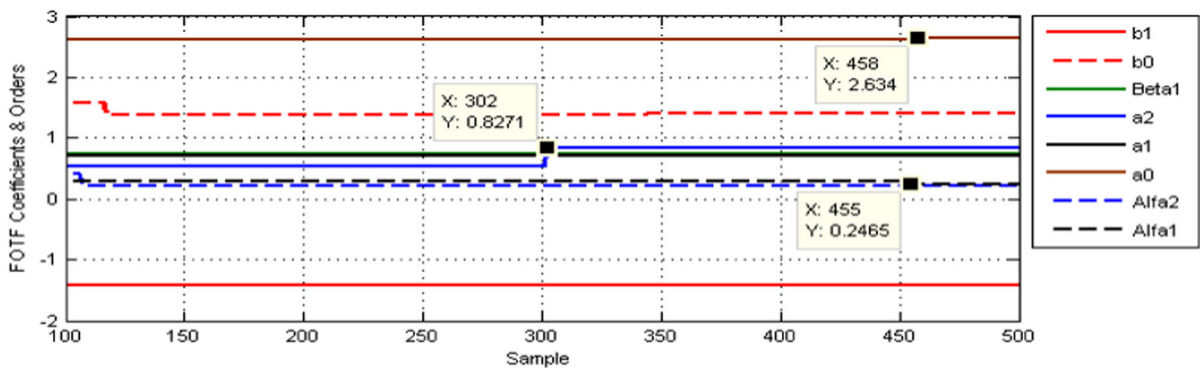
Also, in order to compare the convergence velocity of the parameters in three different structures, the
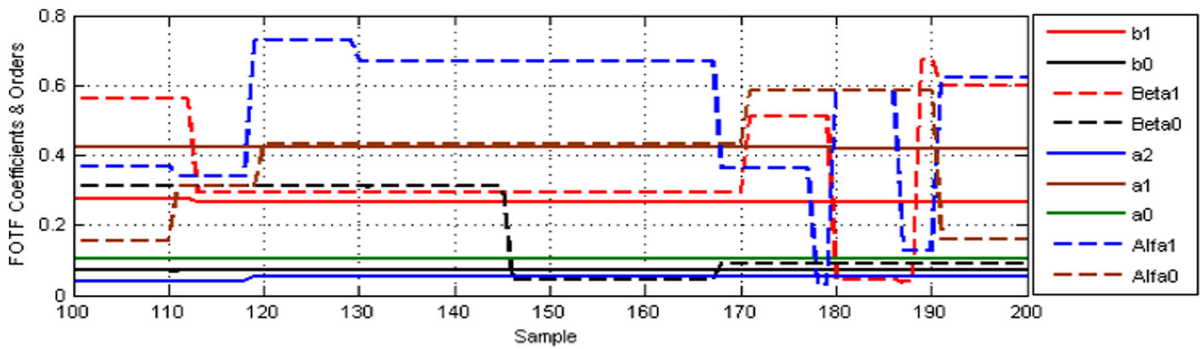
Fig. 30 The estimation of transfer function coefficients and fractional orders using BBP—Example 2
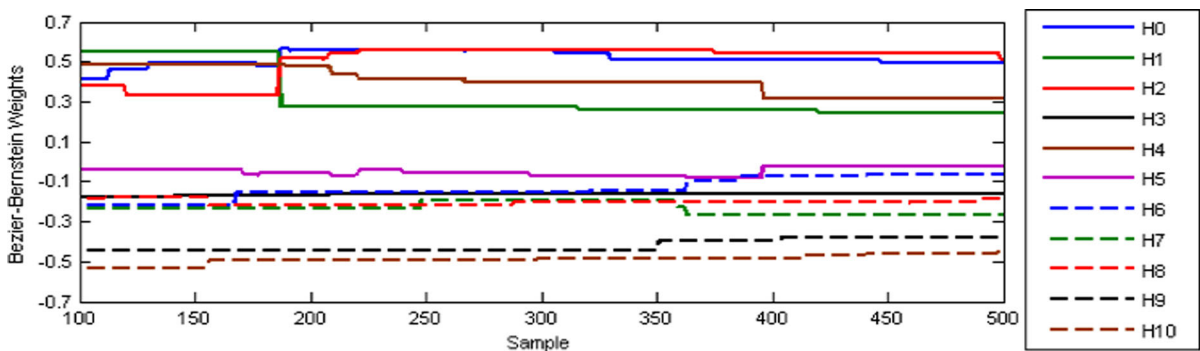
Fig. 31 The estimation of polynomial weights using BBP—Example 2

Fig. 32 The estimation of transfer function coefficients and fractional orders using RBFNN—Example 2

variations of the identified parameters for Hammerstein model with Bezier–Bernstein polynomials are presented in Figs. 30 and 31, for Hammerstein model with RBFNN in Figs. 32, 33 and 34 and for the Hammerstein model with MRBFNN in Figs. 35, 36 and 37. The comparison of the figures shows that the process of identifying the Hammerstein model with MRBFNN has the highest convergence speed between the three proposed structures, due to the smaller number of unknown parameters identified.
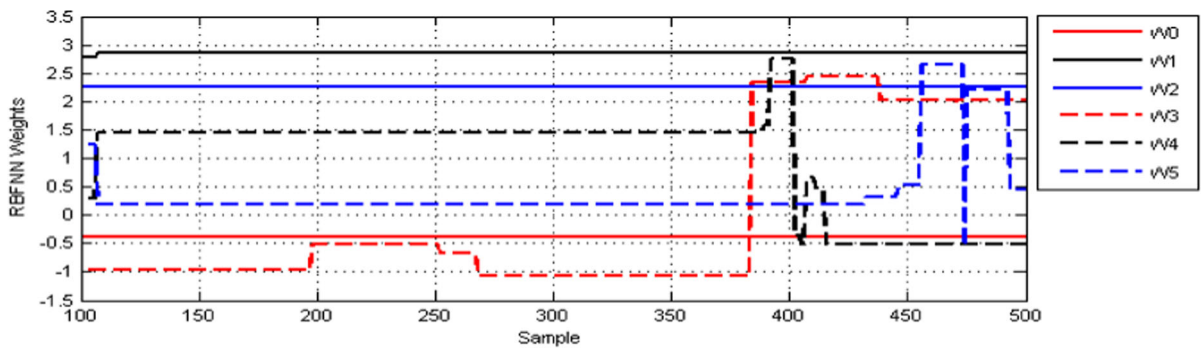
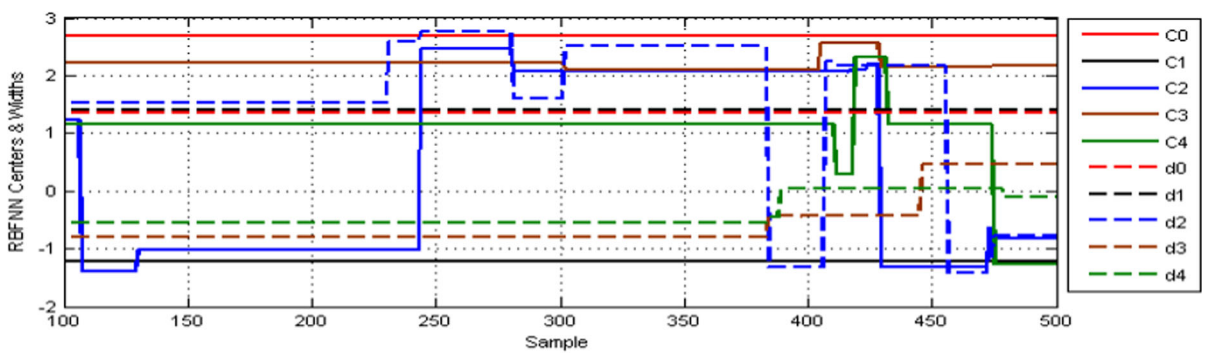**Fig. 33** The estimation of NN weights using RBFNN—Example 2

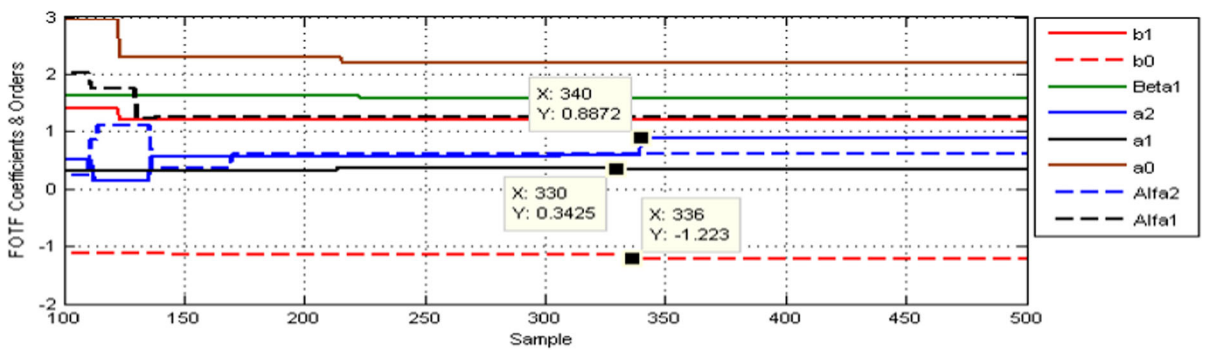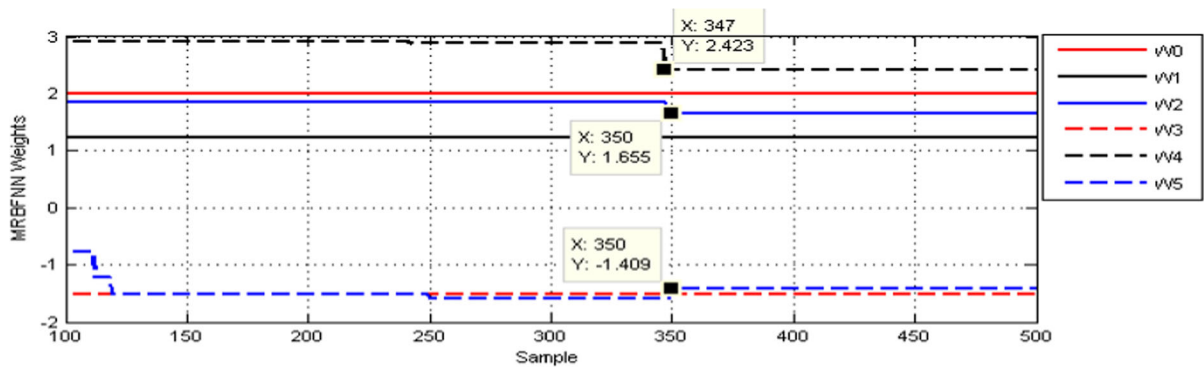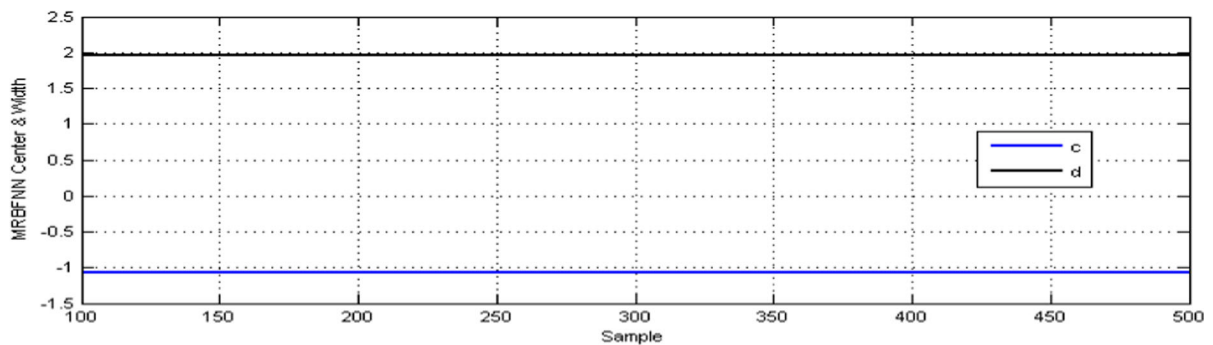**Fig. 34** The estimation of NN centers and widths using RBFNN—Example 2

Fig. 35 The estimation of transfer function coefficients and fractional orders using MRBFNN—Example 2

Fig. 36 The estimation of NN weights using MRBFNN—Example 2

Fig. 37 The estimation of NN centers and widths using RBFNN—Example 2

5 Conclusion

In this paper, a numerical example and a benchmark problem were introduced to evaluate the accuracy of the proposed online identification method. Then, each of the three proposed structures was used to identify these systems and the results were presented. These results confirm the ability of the proposed method to accurately identify the system and eliminate noise. A comparison is made between the online identification of the three proposed structures in terms of convergence speed and estimation accuracy, which shows the relative superiority of accuracy using the modified one. This comparison shows that using the Hammerstein model with the MRBFNN, the accuracy of estimation and the speed of convergence of the parameters in the online mode increase compared to the use of the classical NN. The estimation accuracy of the modified NN in the first example is 63.05% higher

than the Bezier–Bernstein polynomial and 18.7% lower in the second example. Considering the speed of convergence of the modified NN compared to the use of Bezier–Bernstein polynomials, the use of the Hammerstein structure with the modified NN is recommended.

Declarations

References

1. Oustaloup, A.: From fractality to non-integer derivation through recursivity, a property common to these two concepts: a fundamental idea from a new process control strategy. In: Proceeding of the 12th IMACS World Congress, Paris, pp. 18–22. (1988)
2. Manabe, S.: The non-integer integral and its application to control systems. Mitsubishi Denki Lab. Rep. **2**(2), 1–4 (1961)
3. Miller, K.S., Ross, B.: An Introduction to the Fractional Calculus and Fractional Differential Equations. Wiley, New York (1993)
4. Podlubny, I.: Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of their Solution and some of their Applications. Academic Press, Cambridge (1998)
5. Baleanu, D., Güvenç, Z.B., Machado, J.T. (eds.): New Trends in Nanotechnology and Fractional Calculus Applications, p. 978. Springer, New York (2010)
6. Sabatier, J., Agrawal, O.P., Machado, J.T.: Advances in Fractional Calculus. Springer, Dordrecht, The Netherlands (2007)
7. Gabano, J.D., Poinot, T., Kanoun, H.: Identification of a thermal system using continuous linear parameter-varying fractional modelling. IET Control Theory Appl. **5**(7), 889–899 (2011)
8. Ionescu, C.M., Hodrea, R., De Keyser, R.: Variable time-delay estimation for anesthesia control during intensive care. IEEE Trans. Biomed. Eng. **58**(2), 363–369 (2011)
9. Mathieu, B., Melchior, P., Oustaloup, A., Ceyral, C.: Fractional differentiation for edge detection. Signal Process. **83**(11), 2421–2432 (2003)
10. Barbosa, R.S., Machado, J.T., Silva, M.F.: Time domain design of fractional differintegrators using least-squares. Signal Process. **86**(10), 2567–2581 (2006)
11. Melchior, P., Orsoni, B., Lavialle, O., Poty, A., Oustaloup, A.: Consideration of obstacle danger level in path planning using A$_*$ and fast-marching optimisation: comparative study. Signal Process. **83**(11), 2387–2396 (2003)
12. Yousfi, N., Melchior, P., Rekik, C., Derbel, N., Oustaloup, A.: Design of centralized CRONE controller combined with MIMO-QFT approach applied to non square multivariable systems. Int. J. Comput. Appl. **45**(16), 0975–8887 (2012)
13. Silva, M.F., Machado, J.T., Lopes, A.M.: Fractional order control of a hexapod robot. Nonlinear Dyn. **38**(1–4), 417–433 (2004)
14. Cao, J., Ma, C., Xie, H., Jiang, Z.: Nonlinear dynamics of duffing system with fractional order damping. J. Comput. Nonlinear Dyn. **5**(4), 041012 (2010)
15. Cugnet, M., Sabatier, J., Laruelle, S., Grugeon, S., Sahut, B., Oustaloup, A., Tarascon, J.M.: On lead-acid-battery resistance and cranking-capability estimation. IEEE Trans. Ind. Electron. **57**(3), 909–917 (2010)
16. Monje, C.A., Vinagre, B.M., Feliu, V., Chen, Y.: Tuning and auto-tuning of fractional order controllers for industry applications. Control Eng. Pract. **16**(7), 798–812 (2008)
17. Zhao, M., Wang, J.: Outer synchronization between fractional-order complex networks: a non-fragile observer-based control scheme. Entropy **15**(4), 1357–1374 (2013)
18. Baleanu, D., Golmankhaneh, A., Nigmatullin, R., Golmankhaneh, A.: Fractional newtonian mechanics. Open Phys. **8**(1), 120–125 (2010)
19. Chen, W., Ye, L., Sun, H.: Fractional diffusion equations by the Kansa method. Comput. Math. Appl. **59**(5), 1614–1620 (2010)
20. Suchorsky, M.K., Rand, R.H.: A pair of van der Pol oscillators coupled by fractional derivatives. Nonlinear Dyn. **69**(1–2), 313–324 (2012)
21. Yang, J.H., Zhu, H.: Vibrational resonance in duffing systems with fractional-order damping. Chaos Interdiscip. J. Nonlinear Sci. **22**(1), 013112 (2012)
22. Rossikhin, Y.A., Shitikova, M.V.: Analysis of damped vibrations of thin bodies embedded into a fractional derivative viscoelastic medium. J. Mech. Behav. Mater. **21**(5–6), 155–159 (2013)
23. Bagley, R.L., Calico, R.A.: Fractional order state equations for the control of viscoelasticallydamped structures. J. Guid. Control Dyn. **14**(2), 304–311 (1991)
24. Battaglia, J.L., Le Lay, L., Batsale, J.C., Oustaloup, A., Cois, O.: Heat flux estimation through inverted non integer identification models. Int. J. Therm. Sci. **39**(3), 374–389 (2000)
25. Zaborovsky, V., & Meylanov, R.: Informational network traffic model based on fractional calculus. In: Info-tech and Info-net, 2001. Proceedings. ICII 2001-Beijing. 2001 International Conferences, vol. 1, pp.58–63. IEEE (2001)
26. Vinagre, B. M., & Feliu, V.: Modeling and control of dynamic system using fractional calculus: application to electrochemical processes and flexible structures. In: Proceeding 41st IEEE Conference Decision and Control vol. 1, pp. 214–239. (2002)
27. Sjöberg, M.M., Kari, L.: Non-linear behavior of a rubber isolator system using fractional derivatives. Veh. Syst. Dyn. **37**(3), 217–236 (2002)
28. Reyes-Melo, M. E., Martinez-Vega, J. J., Guerrero-Salazar, C. A., & Ortiz-Mendez, U.: Application of fractional calculus to modelling of relaxation phenomena of organic dielectric materials. In: Solid Dielectrics, 2004. ICSD 2004. Proceedings of the 2004 IEEE International Conference on vol. 2, pp. 530–533). IEEE (2004)
29. Ionescu, C., Desager, K., De Keyser, R.: Fractional order model parameters for the respiratory input impedance in healthy and in asthmatic children. Comput. Methods Programs Biomed. **101**(3), 315–323 (2011)
30. Li, C., Peng, G.: Chaos in Chen's system with a fractional order. Chaos Solitons Fractals **22**(2), 443–450 (2004)
31. Chen, Y., Petras, I., & Xue, D.: Fractional order control-a tutorial. In: 2009 American Control Conference, pp. 1397–1411. IEEE (2009)
32. Farouki, R., Goodman, T.: On the optimal stability of the Bernstein basis. Math. Comput. Am. Math. Soc. **65**(216), 1553–1566 (1996)

33. Pislaru, C., Shebani, A.: Identification of nonlinear systems using radial basis function neural network. Int. J. Comput. Inf. Syst. Control Eng. **8**(9), 1528–1533 (2014)

34. Hong, X., Mitchell, R.J.: Hammerstein model identification algorithm using Bezier-Bernstein approximation. IET Control Theory Appl. **1**(4), 1149–1159 (2007)

35. Aoun, M., Malti, R., Cois, O., & Oustaloup, A.: System identification using fractional Hammerstein models. In: Proceeding of the 15th IFAC World Congress, pp. T-We-M21. (2002)

36. Liao, Z., Zhu, Z., Liang, S., Peng, C., Wang, Y.: Subspace identification for fractional order Hammerstein systems based on instrumental variables. Int. J. Control Autom. Syst. **10**(5), 947–953 (2012)

37. Vanbeylen, L.: A fractional approach to identify Wiener-Hammerstein systems. Automatica **50**(3), 903–909 (2014)

38. Zhao, Y., Li, Y., & Chen, Y.: Complete parametric identification of fractional order Hammerstein systems. In: Fractional Differentiation and its Applications (ICFDA), 2014 International Conference on, pp. 1–6. IEEE (2014)

39. Li, Y., Zhai, L., Chen, Y., & Ahn, H. S.: Fractional-order iterative learning control and identification for fractional-order Hammerstein system. In: Intelligent Control and Automation (WCICA), 2014 11th World Congress on, pp. 840–845. IEEE (2014)

40. Hammar, K., Djamah, T., & Bettayeb, M.: Fractional Hammerstein system identification using polynomial nonlinear state space model. In: Control, Engineering & Information Technology (CEIT), 2015 3rd International Conference on, pp. 1–6. IEEE (2015)

41. Hammar, K., Djamah, T., & Bettayeb, M.: Fractional Hammerstein system identification using particle swarm optimization. In: 2015 7th International Conference on Modelling, Identification and Control (ICMIC), pp. 1–6. IEEE (2015)

42. Ivanov, D. V.: Identification discrete fractional order Hammerstein systems. In: Control and Communications (SIBCON), 2015 International Siberian Conference on, pp. 1–4. IEEE (2015)

43. Karimshoushtari, M., Novara, C.: Design of experiments for nonlinear system identification: a set membership approach. Automatica **119**, 109036 (2020)

44. Karagoz, R., Batselier, K.: Nonlinear system identification with regularized tensor network B-splines. Automatica **122**, 109300 (2020)

45. Kaltenbacher, B., Nguyen, T.T.N.: A model reference adaptive system approach for nonlinear online parameter identification. Inverse Probl. **37**(5), 055006 (2021)

46. Du, Y., Liu, F., Qiu, J., Buss, M.: A novel recursive approach for online identification of continuous-time switched nonlinear systems. Int. J. Robust Nonlinear Control **31**(15), 7546–7565 (2021)

47. Mania, H., Jordan, M.I., Recht, B.: Active learning for nonlinear system identification with guarantees. J. Mach. Learn. Res. **23**, 32–41 (2022)

48. Leylaz, G., Wang, S., Sun, J.Q.: Identification of nonlinear dynamical systems with time delay. Int. J. Dyn. Control **10**(1), 13–24 (2022)

49. Monje, C.A., Chen, Y., Vinagre, B.M., Xue, D., Feliu-Batlle, V.: Fractional-order Systems and Controls: Fundamentals and Applications. Springer, Cham (2010)

50. Mohammed, A.A.: Approximate solution of differential equations of fractional orders using Bernstein-Bézier polynomial. Al- Mustansiriya J. Sci. **23**(1), 65–74 (2012)

51. Jahani Moghaddam, M., Mojallali, H., Teshnehlab, M.: A multiple-input–single-output fractional-order Hammerstein model identification based on modified neural network. Math. Methods Appl. Sci. **41**(16), 6252–6271 (2018)

52. Elleuch, K., Chaari, A.: Modeling and identification of Hammerstein system by using triangular basis functions. Int. J. Electr. Comput. Energ. Electron. Commun. Eng. **1**, 1 (2011)

53. Laurain, V., Tóth, R., Gilson, M., Garnier, H.: Direct identification of continuous-time linear parameter-varying input/output models. IET Control Theory Appl. **5**(7), 878–888 (2011)

54. Rao, G. P., & Garnier, H.: Numerical illustrations of the relevance of direct continuous-time model identification. In: 15th Triennial IFAC World Congress on Automatic Control. Barcelona, Barcelona (Spain) (2002)