

Research Article

Recursive Identification for Fractional Order Hammerstein Model Based on ADELS

Qibing Jin , Youliang Ye , Wu Cai , and Zeyu Wang 

College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

Correspondence should be addressed to Youliang Ye; ylye@mail.buct.edu.cn

Received 1 December 2020; Revised 4 February 2021; Accepted 15 February 2021; Published 23 February 2021

Academic Editor: Rohit Salgotra

Copyright © 2021 Qibing Jin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper deals with the identification of the fractional order Hammerstein model by using proposed adaptive differential evolution with the Local search strategy (ADELS) algorithm with the steepest descent method and the overparameterization based auxiliary model recursive least squares (OAMRLS) algorithm. The parameters of the static nonlinear block and the dynamic linear block of the model are all unknown, including the fractional order. The initial value of the parameter is obtained by the proposed ADELS algorithm. The main innovation of ADELS is to adaptively generate the next generation based on the fitness function value within the population through scoring rules and introduce Chebyshev mapping into the newly generated population for local search. Based on the steepest descent method, the fractional order identification using initial values is derived. The remaining parameters are derived through the OAMRLS algorithm. With the initial value obtained by ADELS, the identification result of the algorithm is more accurate. The simulation results illustrate the significance of the proposed algorithm.

1. Introduction

Currently, systems in industrial processes have become more and more increasingly complex, such as chemical plants, robotic arms, etc. This requires us to provide a more accurate mathematical model. Therefore, many parameter identification methods have been studied for system modelling and identification, such as bilinear systems, linear systems, and nonlinear systems [1–4]. Chen et al. proposed a recursive identification method to solve the problem of bilinear-parameter model identification [5]. An adaptive data filtering-based gradient iterative algorithm is presented for the system parameter identification [6]. In [7], based on Bayesian optimization, a serial two-level decomposition structure is introduced into the deep hybrid model.

As a widely represented mathematical model, the nonlinear model has attracted more and more attention. The Hammerstein model is a typical nonlinear model composed of static nonlinear links and dynamic linear links. It can reflect the characteristics of process characteristics and describe a series of nonlinear processes such as neutralization processes [8,9],

lithium-ion battery thermal systems [10,11], heat dissipation systems [12], and so on.

Due to the wide application of the Hammerstein model, how to identify accurate mathematical models has become the research direction of many researchers. At present, some methods that can effectively identify the Hammerstein model are proposed, such as the least squares algorithm [13], maximum likelihood algorithm [14,15], etc. There are also some improved algorithms based on traditional algorithms, such as support vector machines [16,17] and multi-innovative stochastic gradient [18,19].

In recent years, many scholars have researched the static nonlinear module of the Hammerstein model, hoping to obtain a mathematical model with higher accuracy and wider applicability. Some methods to describe the nonlinear part are proposed, such as radial basis functions [20], neural fuzzy networks [21,22], polynomials [23], etc. However, these studies rarely focus on dynamic linear modules, which are based on the integer order. Compared with integer-order differential operators, fractional order operators have more accurate results in the expression of heredity, memory, and other phenomena. Fractional calculus extends the order of

calculus from the domain of integers to real numbers and even complex numbers. The increased degrees of freedom of fractional order operators can enable fractional order systems to describe some complex physical phenomena more accurately, such as the fuel cell physical phenomena [24], voltammetric E-tongue system [25], electric vehicle Li-ion batteries [26], etc. Experiments show that fractional order systems can more accurately model actual systems. Therefore, the research on the identification of fractional order Hammerstein model becomes very meaningful.

At present, some identification methods of fractional order systems have been studied. The enhanced response sensitivity approach can reduce the sensitivity of the identification parameter results in measurement noise [27]. Block pulse functions can identify fractional order systems with initial conditions [28]. The recursive instrumental variable algorithm has acceptable accuracy for fractional order models in the absence of noise [29]. However, there are few studies about the identification problem of fractional order Hammerstein model. In [30], parameters of the fractional Hammerstein model are identified by using particle swarm optimization algorithm, which does not consider the internal relationship between system parameters. In [31], the Levenberg–Marquardt Algorithm combined with two decomposition principles is developed to identify parameters of the fractional Hammerstein model. However, when the system has noise interference, the effect of identification is not satisfactory.

The intelligent optimization algorithm has attracted increasingly scholars' attention because of its generalization, simple parameter setting, and easy programming. It has been applied to the field of parameter optimization. In [32], the modified particle swarm optimization algorithm was presented and applied to the parameter identification of fractional order PID controllers. In [33], an innovative metaheuristic algorithm called Hybrid Gray Wolf Optimization Algorithm was proposed, and it is applied to search the best parameters of the fuel cell. However, the results of the intelligent optimization algorithm may be inaccurate because of falling into local optimization.

Based on the above background, this paper will adopt a new method to parameterize the nonlinear and linear coefficients and fractional order of the fractional Hammerstein model, which has rarely been considered before. The identification of each parameter has a corresponding mathematical derivation. In the proposed method, the ADELS algorithm adds a domain search strategy and improves the setting of algorithm parameters. Parameters of the fractional order Hammerstein model including the fractional order are identified. The identification result provides a relatively accurate initial value for the subsequent algorithms and solves the problem that most algorithms rely on the initial value. Then, the fractional order identification method is proposed by using the principle of steepest gradient descent, and the auxiliary model recursive least squares algorithm is used to estimate the coefficients of the fractional order Hammerstein model. Numerical simulations prove the effectiveness of the proposed methods.

The main contributions are to propose an adaptive differential evolution with local search strategy (ADELS) algorithm for identifying the initial parameters of the fractional order Hammerstein system and develop new recursive identification methods for identifying the fractional order and coefficients of the fractional order nonlinear system by using the auxiliary model. Compared with the classic DE, the proposed ADELS algorithm has higher estimation accuracy because of the adaptive strategy and the introduction of Chebyshev sequence. The self-adaptive operator and mutation strategy make the individual take the corresponding search strategy according to the fitness function value, and the Chebyshev sequence has the characteristics of periodicity, randomness, and ergodicity, which makes up for the defect that DE is easy to fall into local optimum. Based on the auxiliary model, the recursive algorithm of fractional order Hammerstein model is given by using the overparameterized least squares and gradient descent algorithm. Both fractional orders and coefficients of the system are considered. The proposed ADELS algorithm can be applied to many fields like robust control, multi-objective optimization, and economic dispatch problems [34–36]. The proposed recursive identification can be extended to some other systems such as linear, bilinear, and nonlinear systems [37–39].

In this paper, the mathematical background of identifying the fractional Hammerstein model is discussed in Section 2. In Section 3, an improved DE algorithm is proposed. The overparameterized auxiliary model recursive least squares algorithm and fractional order identification algorithm are discussed in Section 4. In Section 5, numerical simulations prove the significance of the proposed method. Finally, Section 6 gives some conclusions.

2. Mathematical Background

2.1. Fractional Differentiation. Fractional calculus is the general form of integer calculus. So far, there are many definitions of fractional calculus. In some cases, different definitions of fractional calculus are completely equivalent. The rationality of these definitions has been proven in the literature. Among them, the definitions of three types of fractional calculus are widely used: Grünwald–Letnikov (GL), Riemann–Liouville (RL), and Caputo definitions [40].

In this article, considering the ease of calculation, the GL definition will be used [41]. The α fractional derivative of $f(t)$ is defined as follows:

$${}_a D_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{j=0}^{[t-a/h]} (-1)^j \binom{\alpha}{j} f(t-jh), \quad (1)$$

where a and t are the limits, h is the sampling time, and $[\cdot]$ stands for the integer part. Newton's binomial coefficient can be calculated by using the Euler's Gamma function as follows:

$$\binom{\alpha}{j} = \frac{\Gamma(\alpha + 1)}{\Gamma(j + 1)\Gamma(\alpha - j + 1)}. \quad (2)$$

By using w_j^α to replace the binomial coefficient as follows:

$$w_j^\alpha = (-1)^j \binom{\alpha}{j}. \quad (3)$$

The α fractional derivative of $f(t)$ can be approximated as [42]

$$D^\alpha f(t) = \frac{1}{h^\alpha} \sum_{j=0}^{\lfloor t-a/h \rfloor} w_j^\alpha f(t - jh), \quad (4)$$

where $D^\alpha f(t)$ denotes ${}_a D_t^\alpha f(t)$ when $a = 0$.

2.2. Differential Evolution. Differential evolution (DE) algorithm has better stability and global search ability. The population of each generation of DE is composed of NP parameter vectors with dimension D. Each individual is represented as

$$x_i^G = (x_{i,1}^G, x_{i,2}^G, \dots, x_{i,D}^G), \quad i = 1, 2, \dots, \text{NP}, \quad (5)$$

where i means the sequence of individuals in the population, G denotes the number of evolutionary generations, and NP is the size of the population.

In DE, random initial populations are generated by a uniform probability distribution. After initialization, the population evolves by using three steps: mutation, crossover, and selection.

2.2.1. Mutation. For each target vector x_i^G at generation G , the mutation vector is generated as follows:

$$v_i^G = x_{r1}^G + F \cdot (x_{r2}^G - x_{r3}^G), \quad r1 \neq r2 \neq r3 \neq i \in [1, \text{NP}], \quad (6)$$

where $r1, r2$, and $r3$ are different individuals randomly selected from the population except the individual i and F is

the mutation operator that defines the amplitude of the difference vector. When the generated mutation vector exceeds the boundary value, it will be replaced by the newly generated vector according to the boundary update rule. Some other mutation strategies have also been widely used in the DE algorithm [43–45].

2.2.2. Crossover. To increase the diversity of trial vectors, crossover operations are introduced as follows:

$$u_{j,i}^G = \begin{cases} v_{j,i}^G, & \text{if } \text{rand}_j \leq \text{CR} \text{ or } j = r_k, \\ x_{j,i}^G, & \text{otherwise.} \end{cases} \quad (7)$$

CR is the crossover operator that controls the probability of accepting the mutation in each dimension, rand_j is a random number generated by a uniform distribution in $[0,1]$, j represents the dimension of the individual, and $r_k \in [1, 2, \dots, D]$ is a uniformly distributed random integer.

2.2.3. Selection. The trial vector $u_{i,G}$ and the target vector x_i^G are compared according to the greedy criterion, and the vector with the best fitness value will appear in the next generation population. The selection operation to minimize the fitness function can be expressed as

$$x_i^{G+1} = \begin{cases} u_i^G, & \text{if } f(u_i^G) \leq f(x_i^G) \\ x_i^G, & x \geq 0. \end{cases} \quad (8)$$

Repeat the three steps above until the termination condition is met, such as the total number of iterations or the accuracy requirement of the algorithm.

2.3. Fractional Order Hammerstein Model. The fractional order Hammerstein model is shown in Figure 1.

The system inputs and outputs are $u(t)$ and $y(t)$, $v(t)$ is the Gaussian white noise, and $y'(t)$ is the measured output. It consists of a static nonlinear part and a dynamic linear part. The fractional order differential equation of the linear part can be expressed as

$$a_n D^{n\alpha} y(t) + a_{n-1} D^{(n-1)\alpha} y(t) + \dots + a_0 y(t) = b_m D^{m\alpha} u'(t) + b_{m-1} D^{(m-1)\alpha} u'(t) + \dots + b_0 u'(t). \quad (9)$$

The fractional order transfer function of the linear part is defined as follows:

$$G(s) = \frac{b_m s^{m\alpha} + b_{m-1} s^{(m-1)\alpha} + \dots + b_0}{a_n s^{n\alpha} + a_{n-1} s^{(n-1)\alpha} + \dots + a_0}, \quad (10)$$

where $a_1, \dots, a_n, b_0, b_1, \dots, b_m$ are the coefficients and α is the fractional order. They are unknown and need to be identified. Without loss of generality, assume $a_0 = 1$. The input $u'(t)$ of the linear link is generated by the system input $u(t)$ through the nonlinear block, which can be expressed by the following formula:

$$u'(t) = \sum_{k=1}^{n_c} c_k \cdot f_k(u(t)), \quad (11)$$

where c_1, \dots, c_{n_c} are unknown coefficients to be identified and $f_1(\cdot), \dots, f_{n_c}(\cdot)$ are a series of basis functions.

In summary, the fractional Hammerstein system discussed in this article can be expressed as follows:

$$\begin{cases} y(t) + \sum_{i=1}^{n_a} a_i D^{i\alpha} y(t) = \sum_{j=0}^{n_b} b_j D^{j\alpha} u'(t), \\ y'(t) = y(t) + v(t). \end{cases} \quad (12)$$

3. Adaptive Differential Evolution with Local Search Strategy

The DE algorithm is an intelligent search algorithm worked by cooperation and competition among individuals in the population. It still remains the population-based global search strategy in the evolutionary algorithm. As an efficient parallel search algorithm, the DE algorithm has strong global convergence and robustness, which is worthy of theoretical and applied research.

The adaptive differential evolution with the local search strategy (ADELS) algorithm proposed in this paper has been innovatively improved in three places. Mutation and crossover operators in the DE algorithm control the enlargement ratio of the deviation vector and the probability that the trial vector comes from the mutation vector, which plays a pivotal role. Therefore, the operators in the proposed algorithm adopt the strategy of adaptive change according to the individual fitness value, which improves the optimization efficiency of the algorithm. Different mutation strategies also affect the accuracy of parameter identification. In this paper, different mutation strategies are selected according to the fitness value, which further improves the efficiency of the algorithm. Then, the local search ability of the algorithm is improved by using the Chebyshev chaotic sequence, which makes up for the shortcoming that the DE algorithm is easy to fall into the local optimum.

3.1. Parameters Adaption. Mutation operator and crossover operator affect the optimization ability and the convergence speed of the algorithm. Generally, these operators are fixed constants. In this paper, the individual score is obtained based on the fitness value of each individual in each generation by introducing a nonlinear scoring method, and the score ranges from 0 to 1. The operator will be determined based on the score of each individual, which makes the operator more adaptable. The scoring rules based on the minimized objective function are as follows:

$$\text{score}_i^G = \begin{cases} 0.5 + 0.5 \cdot \sqrt{\frac{(f_{\text{avg}} - f_i^G)}{(f_{\text{avg}} - f_{\text{min}})}}, & f_i^G < f_{\text{avg}}, \\ 0.5 \cdot \sqrt{\frac{f_{\text{max}} - f_i^G}{(f_{\text{max}} - f_{\text{avg}})}}, & \text{otherwise,} \end{cases} \quad (13)$$

where f_{min} , f_{max} , and f_{avg} mean the minimum, maximum, and average fitness function values in the current generation population, respectively, where f_i^G means the value of the i individual fitness function in the G generation. It can be seen from the above formula that individuals with fitness function values below the average get higher scores. Then, the formulas of mutation operator and crossover operator can be expressed as follows:

$$F_i^G = F_{\text{max}} - \text{score}_i^G \cdot (F_{\text{max}} - F_{\text{min}}), \quad (14)$$

$$\text{CR}_i^G = \text{CR}_{\text{max}} - \text{score}_i^G \cdot (\text{CR}_{\text{max}} - \text{CR}_{\text{min}}), \quad (15)$$

where F_i^G and CR_i^G are the i -th mutation operator and the i -th crossover operator at generation G ; F_{max} and F_{min} represent the limits of the mutation operator; and CR_{max} and CR_{min} are the limits of the crossover operator.

It can be seen from the above formulas that the value of the operator changes with the individual's fitness value score. When the individual's fitness value is low, the score will be high according to equation (13), and the value of F_i^G and CR_i^G will decrease. When the individual's fitness value is higher than the average, the score is low, and the value of F_i^G and CR_i^G will increase accordingly. The individual is close to the optimal solution when the individual's fitness value is low, and the required mutation range and crossover probability are reduced. Then, the values of the mutation operator and crossover operator need to be reduced accordingly to find the optimal value. On the contrary, when the individual's fitness value is high, the value of the operator needs to be increased to search for the optimal solution in a larger range.

3.2. Mutation Strategy. In each iteration, individuals are sorted according to the fitness value, and the population will be divided into two subpopulations with different individual fitness function. Two subpopulations adopt different mutation strategies as follows:

$$\begin{cases} v_i^G = x_{\text{Best}}^G + F_i^G \cdot (x_{r1}^G - x_{r2}^G), & f_i^G < f_{\text{avg}}, \\ v_i^G = x_{r1}^G + F_i^G \cdot (x_{r2}^G - x_{r3}^G), & \text{otherwise,} \end{cases} \quad (16)$$

where x_{r1}^G , x_{r2}^G , and x_{r3}^G are different individuals randomly selected from the population except for the individual x_i^G and x_{Best}^G represents the best individual of the previous generation population. From equation (16), it is worth noticing that at least four individuals in the population are needed.

3.3. Local Search Strategy. To avoid the algorithm falling into local optimum, a Chebyshev chaotic sequence is introduced to perform a local search for the optimal individual. Chaotic systems have the characteristics of periodicity, randomness, and ergodicity, which can increase the diversity of individual populations. The family of Chebyshev polynomial can be expressed as

$$\rho_m(h) = \cos(m \cdot \cos^{-1}(h)), \quad m = 0, 1, 2, \dots \quad (17)$$

Then, the Chebyshev map can be expressed as

$$h_{n+1} = \rho_m(h_n), \quad h_n \in [-1, 1], \quad (18)$$

where m is the order of the polynomial. In this paper, to get better ergodicity of the mapping, the parameter of the Chebyshev map is set to $m = 4$ and $h_1 = 0.7$.

Then, a dual search strategy is introduced to search for the best gene for each dimension of the optimal individual. Due to the randomness and ergodicity of chaos, chaotic mapping has a better performance in improving the diversity of the population compared to a uniform probability distribution. The local search strategy is as follows:

$$\begin{cases} x_{i,j}' = x_{\text{Best},j} + (x_{j\text{max}} - x_{j\text{min}}) \cdot \frac{h_{n+1}}{G^2}, \\ x_{i,j}'' = x_{\text{Best},j} + (x_{j\text{max}} - x_{j\text{min}}) \cdot \frac{h_{n+1}}{G^{\sqrt{G}}}, \end{cases} \quad (19)$$

where x_i' and x_i'' are obtained by the local search, and they will be compared with the current optimal individual based on the fitness value. If a new optimal individual is obtained, it will be randomly copied to ten individuals of the next generation population.

Because of adaptive operators, different mutation strategies, and local search based on the Chebyshev chaotic map, the algorithm's optimization ability and convergence speed have been significantly improved according to the following numerical simulations. The effectiveness of the ADELS algorithm will be proved in the numerical simulation in Section 4. The proposed ADELS Algorithm 1 is shown as follows.

4. The Identification Algorithm of Fractional Order Hammerstein Model

In the fractional order Hammerstein model, the polynomial coefficients of the static nonlinear block, the coefficients of the dynamic linear block, and the fractional order are all needed to be determined, which is rarely concerned before. In this paper, a set of input and output data will be first used to obtain the initial values of coefficients and the fractional order through the proposed ADELS algorithm. Then, the initial value will be used to get the final parameter identification result of the fractional order Hammerstein model through an over-parameterized based auxiliary model recursive least squares (OAMRLS) algorithm and the steepest descent method.

4.1. *Parameter Estimation Using Overparameterized Auxiliary Model Recursive Least Squares (OAMRLS) Algorithm.* The OAMRLS algorithm is required to get the initial values of the parameters, which have been calculated in the ADELS algorithm. According to equations (1) and (2), the fractional Hammerstein system can be expressed as follows:

$$y'(t) = \frac{1}{(1 + \sum_{i=1}^{n_a} a_i/h^{i\alpha})} \cdot \left[\sum_{i=0}^{n_b} \sum_{k=1}^{n_c} \frac{b_i}{h^{i\alpha} c_k} \cdot \sum_{j=0}^{[t/h]} w_j^{i\alpha} f_k(u(t-jh)) - \frac{1}{h^{i\alpha}} \sum_{j=1}^{[t/h]} y(t-jh) - \sum_{i=1}^{n_a} \frac{a_i}{h^{i\alpha}} \sum_{j=1}^{[t/h]} w_j^{i\alpha} y(t-jh) \right] + v(t). \quad (20)$$

The input-output relations can be written in the regression form as

$$y'(t) = y(t) + v(t) = \varphi^T(t) \cdot \theta + v(t), \quad (21)$$

where $y(t)$ is the noise-free output of the system and φ is the variable vector including input-output data, which can be expressed as

$$\begin{aligned} \varphi(t) &= [\varphi_a(t), \varphi_{b_0}(t), \varphi_{b_1}(t), \dots, \varphi_{b_{n_b}}(t)]^T, \\ Y_m(t) &= \sum_{j=1}^{[t/h]} w_j^{m\alpha} y(t-jh), \\ \varphi_a(t) &= [-Y_0(t), -Y_1(t), \dots, -Y_{n_a}(t)], \\ U_{i,p}(k) &= \sum_{j=0}^{[t/h]} w_j^{i\alpha} f_p(u(t-jh)), \\ \varphi_{b_i}(t) &= [U_{i,1}(k), U_{i,2}(k), \dots, U_{i,n_c}(k)]. \end{aligned} \quad (22)$$

The estimated vector θ is given by

$$\begin{aligned} \theta &= [\theta_a, \theta_{b_0}, \dots, \theta_{b_{n_b}}]^T, \\ \theta_a &= [a'_0, a'_1, \dots, a'_{n_a}], \\ \theta_{b_j} &= [b'_j c_1, b'_j c_2, \dots, b'_j c_{n_c}], \end{aligned} \quad (23)$$

where

$$\begin{aligned} a'_0 &= \frac{1}{1 + \sum_{i=1}^{n_a} a_i/h^{i\alpha}}, \\ a'_i &= \frac{(a_i/h^{i\alpha})}{1 + \sum_{i=1}^{n_a} a_i/h^{i\alpha}}, \\ b'_j &= \frac{(b_j/h^{i\alpha})}{1 + \sum_{i=1}^{n_a} a_i/h^{i\alpha}}, \quad i = [1, 2, \dots, n_a], j = [0, 1, 2, \dots, n_b]. \end{aligned} \quad (24)$$

It should be noted that the vector $\varphi_a(t)$ contains an unknown variable $y(t-jh)$, which makes it impossible to directly use the least squares algorithm or gradient descent method to solve the problem. In view of this, the auxiliary model is used to estimate the unknown variable $y(t-jh)$. It

can be seen from Figure 2 that $u(t)$ is used as an input to build an auxiliary model. The main idea of the auxiliary model is that the real output of the system $y(t)$ is replaced by the output of the auxiliary model $y_{\text{am}}(t)$; then, the parameter identification problem can be solved by $u(t)$ and $y_{\text{am}}(t)$. The parameters of the auxiliary model will eventually approximate the real ones after iteration.

According to Figure 2, $F_{\text{am}}(\cdot)$ is the nonlinear block of the auxiliary system, and $G_{\text{am}}(\cdot)$ represents the linear block of the auxiliary system. The expression of the auxiliary model regression form can be written as follows:

$$y_{\text{am}}(t) = \varphi_{\text{am}}^T(t) \cdot \theta_{\text{am}}, \quad (25)$$

where the estimate of $\hat{\varphi}(t)$ can be used as the value of the auxiliary model information vector $\varphi_{\text{am}}(t)$ and the parameter estimate of $\hat{\theta}$ can be used as the value of the auxiliary model parameter vector θ_{am} :

$$\begin{aligned} \varphi_{\text{am}}(t) &= \hat{\varphi}(t), \\ \theta_{\text{am}} &= \hat{\theta}. \end{aligned} \quad (26)$$

Then, the output of the auxiliary model $y_{\text{am}}(t)$ is used to represent the estimate of $\hat{\varphi}_a(t)$:

$$\begin{aligned} \hat{\varphi}(t) &= \left[\hat{\varphi}_a(t), \hat{\varphi}_{b_0}(t), \hat{\varphi}_{b_1}(t), \dots, \hat{\varphi}_{b_{n_b}}(t) \right]^T, \\ Y'_m(t) &= \sum_{j=1}^{\lfloor t/h \rfloor} w_j^{m\alpha} y_{\text{am}}(t - jh), \\ \hat{\varphi}_a(t) &= [-Y'_0, -Y'_1, \dots, -Y'_{n_a}]^T, \\ \hat{\varphi}_{b_i}(t) &= \varphi_{b_i}(t), \\ \hat{\theta} &= \left[\hat{\theta}_a, \hat{\theta}_{b_0}, \dots, \hat{\theta}_{b_{n_b}} \right]^T. \end{aligned} \quad (27)$$

Define the criterion function as

$$J(\hat{\theta}) = \frac{1}{2} \sum_{i=1}^t [y'(i) - \hat{\varphi}^T(i)\hat{\theta}]^2. \quad (28)$$

Then, the estimated $\hat{\theta}$ can be obtained by minimizing the criterion function as the following formula:

$$\frac{\partial J(\hat{\theta})}{\partial \hat{\theta}} = - \sum_{i=1}^t \hat{\varphi}^T(i) \cdot [y'(i) - \hat{\varphi}^T(i)\hat{\theta}] = 0. \quad (29)$$

Provided the inverse $[\sum_{i=1}^t \hat{\varphi}(i-1) \cdot \hat{\varphi}^T(i-1)]^{-1}$ exists, then $\hat{\theta}(t)$ can be obtained by using least squares estimation:

$$\hat{\theta}(t) = \left[\sum_{i=1}^t \hat{\varphi}(i-1) \cdot \hat{\varphi}^T(i-1) \right]^{-1} \cdot \sum_{i=1}^t \hat{\varphi}^T(i-1) y'(i). \quad (30)$$

Then, a recursive version is given as the following formula:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y'(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)], \quad (31)$$

$$L(t) = P(t-1)\hat{\varphi}(t) [1 + \hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)]^{-1}, \quad (32)$$

$$P(t) = [I_n - L(t)\hat{\varphi}^T(t)]P(t-1), \quad (33)$$

$$y_{\text{am}}(t) = \hat{\varphi}^T(t) \cdot \hat{\theta}(t), \quad (34)$$

where the adaptation gain matrix P is generally chosen to be started by

$$P(0) = \frac{1}{p_0} \cdot I_n, \quad 0 < p_0 \ll 1. \quad (35)$$

For the problem of overparameterization, to ensure the uniqueness of parameter identification, the value of c_1 can be assumed to be 1 without loss of generality. Then, the value of c_k and b'_j can be calculated as follows:

$$\begin{aligned} b'_j &= \theta_{b_j}(1), \\ c_k &= \sum_{j=0}^{nb} \frac{\theta_{b_j}(t)}{b'_j}. \end{aligned} \quad (36)$$

Based on the identification result of the ADELS algorithm, the OAMRLS algorithm cooperates with the following order identification method to complete the identification of the system parameters.

4.2. Fractional Order Identification Based on Steepest Descent Method. The initial value of the fractional order has been given by the ADELS algorithm; then, the iterative optimization of the fractional order is performed according to the steepest descent method. Define the criterion function as

$$J(\alpha) = \frac{1}{2} \sum_{t=1}^k [y'(t) - \hat{y}(t)]^2. \quad (37)$$

The criterion function can be minimized by the steepest descent method:

$$\alpha(n+1) = \alpha(n) - \gamma(n)\nabla J(\alpha(n)), \quad (38)$$

where $\gamma(n)$ satisfies $J(\alpha(n) - \gamma(n)\nabla J(\alpha(n))) < J(\alpha(n))$ and $\nabla J(\alpha(n))$ is the partial derivative of J to $\alpha(n)$ which can be expressed as follows:

$$\nabla J(\alpha(n)) = \sum_{i=1}^k \frac{\partial (y'(t) - \hat{y}(t))}{\partial \alpha} [y'(t) - \hat{y}(t)], \quad (39)$$

where

$$\frac{\partial (y'(t) - \hat{y}(t))}{\partial \alpha} = - \frac{\partial}{\partial \alpha} (\hat{G}(s^\alpha) u'(t)). \quad (40)$$

According to formula (10), $\hat{G}(s^\alpha)$ is defined as

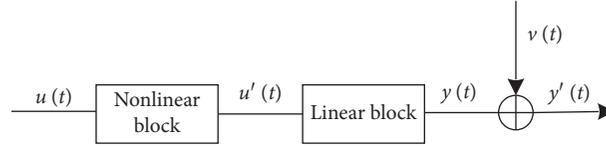


FIGURE 1: The fractional order Hammerstein model.

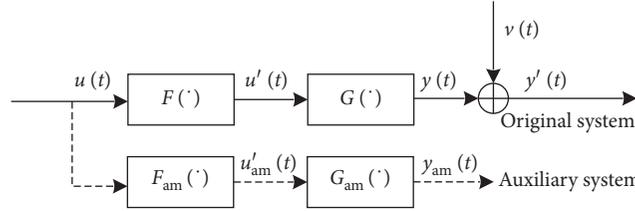


FIGURE 2: The fractional order Hammerstein based on the auxiliary model.

$$\widehat{G}(s^\alpha) = \frac{b_{n_b} s^{n_b \alpha} + b_{n_b-1} s^{n_b-1 \alpha} + \dots + b_0}{a_{n_a} s^{n_a \alpha} + a_{n_a-1} s^{n_a-1 \alpha} + \dots + a_0}, \quad (41)$$

$$\begin{aligned} -\frac{\partial}{\partial \alpha} (\widehat{G}(s^\alpha) u'(t)) = & - \left[\left(\frac{b_{n_b} s^{n_b \alpha} + \dots + b_0}{(a_{n_a} s^{n_a \alpha} + \dots + a_0)^2} \right) \cdot (n_a \cdot a_{n_a} s^{n_a \alpha} + \dots + a_1 s^\alpha) \right. \\ & \left. + \left(\frac{n_b \cdot b_{n_b} s^{n_b \alpha} + \dots + b_1 s^\alpha}{a_{n_a} s^{n_a \alpha} + \dots + a_0} \right) \right] \cdot \ln(s) \cdot u'(t), \end{aligned} \quad (42)$$

where $a_0 = 1$. As for the polynomial $\ln(s) \cdot u'(t)$, it can be replaced by $s^\alpha \cdot (\ln(s)/s^\alpha) \cdot u'(t)$ [46,47], the inverse Laplace transform of $(\ln(s)/s^\alpha)$ is

$$\frac{t^{\alpha-1}}{\Gamma(\alpha)} [\Psi(\alpha) - \ln(t)], \quad (43)$$

where $\Psi(\alpha)$ can be expressed as follows:

$$\Psi(\alpha) = \frac{1}{\Gamma(\alpha)} \frac{d\Gamma(\alpha)}{d\alpha}. \quad (44)$$

Then, $\ln(s) \cdot u'(t)$ can be expressed as follows:

$$D^\alpha \left[\left(\frac{t^{\alpha-1}}{\Gamma(\alpha)} (\Psi(\alpha) - \ln(t)) \right) * u'(t) \right] = D^\alpha \left[\Psi(\alpha) D^\alpha u'(t) - \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} \ln(t-\tau) u'(\tau) d\tau \right]. \quad (45)$$

Consequently, $\nabla J(\alpha(n))$ can be obtained by equations (39)–(45).

Combined with ADELS and OAMRLS algorithms, the identification process of the fractional Hammerstein model is shown as follows.

5. Experimental Results

5.1. Benchmark Functions. Firstly, 15 benchmark functions will be used to verify the effectiveness of the ADELS algorithm, and the results will be compared with other intelligent optimization algorithms proposed in recent years. These functions include unimodal and multimodal functions. The function expression and search range are shown in Table 1.

In this paper, the ADELS algorithm will be compared with three algorithms including JADE [48], GWO [49], and WOA [50]. Their parameter settings are given in Table 2.

In this section, the population size of the four algorithms is set to 50, the number of iterations is 500, and the simulation dimensions of the benchmark function are 10 and 30. The simulation results are shown in Tables 3 and 4.

The convergence curves of ADELS, GWO, WOA, and JADE are as shown in Figures 3 and 4; ADELS is very competitive with other metaheuristic algorithms due to its adaptive operators and local search with the Chebyshev chaotic map.

The unimodal function can be used to evaluate the exploitation capability of the optimization algorithm, whereas the multimodal function is very effective to evaluate

- (1) Define the objective function $f(X)$, $X = (x_1, x_2, \dots, x_D)^T$;
- (2) Initialize parameters of the Chebyshev map: m and h_1 ;
- (3) Initialize individuals X ;
- (4) Evaluate all the individuals in the population by the objective function $f(X)$;
- (5) Initialize the number of iteration $k = 1$;
- (6) While ($k < \text{max number of iterations } N$)
- (7) For each individual
- (8) Update operators adaptively (equations (13)–(15));
- (9) The mutation vector v_i^G is obtained by mutation (equations (16));
- (10) If the generated mutation vector exceeds the boundary, a new mutation vector is generated randomly, until it is within the boundary;
- (11) The trial vectors u_i^G is obtained by equations (7);
- (12) The best individuals X is obtained by greedy selection (equations (8));
- (13) Find the current best X according to the local search strategy (equations (17)–(19));
- (14) If a new optimal individual X is obtained, it will be randomly copied to ten individuals of the next generation population;
- (15) End
- (16) $k = k + 1$;
- (17) End while;
- (18) Postprocess results and visualization.

ALGORITHM 1: ADELS.

TABLE 1: Benchmark functions and search range.

Function	Range
$F1(x) = \sum_{i=1}^D x_i^2$	$[-100,100]$
$F2(x) = \sum_{i=1}^D x_i + \prod_{i=1}^D x_i $	$[-10,10]$
$F3(x) = \sum_{i=1}^D x_i + 0.5 ^2$	$[-100,100]$
$F4(x) = \sum_{i=1}^D (-x_i \cdot \sin(\sqrt{ x_i }))$	$[-500,500]$
$F5(x) = \sum_{i=1}^D i \cdot x_i^2$	$[-10,10]$
$F6(x) = x_1^2 + 10^6 \sum_{i=2}^D x_i^2$	$[-100,100]$
$F7(x) = \sum_{i=1}^{D-1} (100 \cdot (x_i^2 - x_{i+1})^2 + (x_i - 1)^2)$	$[-100,100]$
$F8(x) = \sin^2(\pi w_1) + (w_D - 1)^2 [1 + \sin^2(2\pi w_D)] + \sum_{i=1}^{D-1} (w_i - 1)^2 \cdot [1 + 10 \sin^2(\pi w_i + 1)]$, $w_i = 1 + ((x_i - 1)/4)$, $i = 1, \dots, D$	$[-100,100]$
$F9(x) = \sum_{i=1}^D (10^6)^{i-1/D-1} \cdot x_i^2$	$[-100,100]$
$F10(x) = 10^6 \cdot x_1^2 + \sum_{i=2}^D x_i^2$	$[-100,100]$
$F11(x) = -20 \exp(-0.2 \sqrt{(1/D) \sum_{i=1}^D x_i^2}) - \exp((1/D) \sum_{i=1}^D \cos(2\pi x_i)) + 20 + e$	$[-32,32]$
$F12(x) = \sum_{i=1}^D (\sum_{k=0}^{20} [a^k \cos(2\pi b^k (x_i + 0.5))]) - D \sum_{k=0}^{20} a^k \cos(2\pi b^k \cdot 0.5)$, $a = 0.5, b = 3$	$[-100,100]$
$F13(x) = \sum_{i=1}^D (x_i^2/4000) - \prod_{i=1}^D \cos(x_i/\sqrt{i}) + 1$	$[-100,100]$
$F14(x) = (10/D^2) \prod_{i=1}^D (1 + i \sum_{j=1}^{32} (2^j x_i - \text{round}(2^j x_i) /2^j))^{10/D^{1.2}} - (10/D^2)$	$[-100,100]$
$F15(x) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10)$	$[-5.12, 5.12]$

TABLE 2: Parameter settings.

Algorithm	Parameter settings
JADE	$CR = 0.5, F = 0.5, \text{size} = 100, P_0 = 0.05$
GWO	Size = 50
WOA	Size = 50
ADELS	Size = 50, $F_{\min} = 0.1, F_{\max} = 0.9, CR_{\min} = 0.1, CR_{\max} = 0.9$

the exploration capability of the algorithm. The results reported in Tables 3 and 4 indicate that the exploitation and exploration capabilities of ADELS are significantly improved compared to other algorithms.

5.2. Identification of Fractional Hammerstein Model. In this section, several different algorithms will be compared to prove the effectiveness of the proposed algorithm. The fractional Hammerstein model is presented as follows:

TABLE 3: Comparison of optimization results obtained for the 30-dimensional benchmark functions.

F	JADE		GWO		WOA		ADELS	
	Ave	Std	Ave	Std	Ave	Std	Ave	Std
F1	2.55E-22	2.43E-22	5.46E-33	4.78E-33	3.35E-86	2.52E-83	0	0
F2	7.65E-11	5.09E-11	7.26E-20	6.93E-20	2.13E-54	5.88E-52	1.80E-190	0
F3	2.01E-22	1.98E-22	0.4587	0.20	0.086	6.7E-02	0	0
F4	-11684.98	142.47	-6281.5	861.64	-11535.27	1685.05	-12391.83	76.85
F5	1.96E-21	4.00E-21	3.46E-32	7.43E-32	3.12E-85	2.68E-84	0	0
F6	1.01E-15	9.11E-16	1.14E-27	2.34E-27	3.15E-78	7.32E-79	0	0
F7	17.69	2.06	26.67	0.87	27.54	0.33	0.2391	19.20
F8	2.3E-02	0.10	0.94	0.22	0.49	0.28	1.50E-32	5.62E-48
F9	1.78E-17	6.95E-18	3.52E-30	1.63E-29	1.79E-80	8.55E-82	0	0
F10	5.25E-21	5.43E-21	7.65E-33	4.43E-33	2.44E-83	1.12E-80	0	0
F11	3.05E-12	4.89E-12	4.42E-14	4.53E-15	3.73E-15	2.55E-15	1.15E-15	3.63E-15
F12	-675.49	10.04	-476.51	66.40	-870.00	1.02E-12	-870.00	3.61E-13
F13	9.60E-10	4.28E-09	1.4E-03	7.8E-03	5.1E-03	0.0078	0	0
F14	9.11E-12	6.81E-13	3.68E-11	5.30E-12	0	0	0	0
F15	26.51	3.41	2.69	3.38	0	0	0	0

TABLE 4: Comparison of optimization results obtained for the 10-dimensional benchmark functions.

F	JADE		GWO		WOA		ADELS	
	Ave	Std	Ave	Std	Ave	Std	Ave	Std
F1	9.16E-86	3.47E-85	8.14E-70	3.06E-69	8.55E-89	2.07E-88	0	0
F2	6.74E-56	2.22E-55	2.88E-40	2.65E-40	2.43E-57	9.34E-57	5.77E-191	0
F3	1.46E-04	1.78E-04	2.56E-06	7.58E-07	1.99E-04	2.2E-04	0	0
F4	-3542.12	672.55	-2806.47	322.09	-3439.87	567.21	-4189.83	60.45
F5	2.74E-88	1.22E-87	1.01E-68	4.31E-68	1.34E-89	4.04E-89	0	0
F6	1.42E-80	6.29E-80	1.94E-64	4.90E-64	1.79E-84	5.36E-84	0	0
F7	6.67	6.38E-01	6.53	5.31E-01	6.60	4.79E-01	1.99E-03	2.61E-03
F8	6.81E-02	1.01E-01	1.02E-01	7.3E-02	4.77E-02	8.06E-02	1.50E-32	5.62E-48
F9	5.18E-85	2.15E-84	2.08E-67	4.60E-67	1.58E-84	3.50E-84	0	0
F10	6.55E-87	2.32E-86	7.91E-69	1.83E-68	6.55E-87	2.91E-86	0	0
F11	3.02E-15	2.13E-15	6.75E-15	1.74E-15	3.73E-15	2.47E-15	4.80E-14	1.09E-14
F12	-90.00	0	-82.63	7.16	-90.00	0	-90.00	0
F13	6.80E-02	1.02E-01	1.8E-02	2.7E-02	5.59E-02	8.67E-02	3.11E-02	2.67E-02
F14	0	0	2.16E-11	1.85E-11	0	0	0	0
F15	0	0	5.99E-01	1.85	0	0	0	0

$$\begin{aligned}
 a_2 D^{2\alpha} y(t) + a_1 D^\alpha y(t) + y(t) &= b_1 D^\alpha u'(t) + b_0 u'(t), \\
 u'(t) &= u(t) + c_2 u^2(t) + c_3 u^3(t) \\
 &\quad + c_4 u^4(t), \\
 y'(t) &= y(t) + v(t).
 \end{aligned}
 \tag{46}$$

The parameter vector to be identified is

$$\theta = [a_1, a_2, b_0, b_1, c_2, c_3, c_4, \alpha]^T = [1.3, 1.5, 1.9, 1.7, 0.5, 0.3, 0.1, 0.3]^T.
 \tag{47}$$

The input u is a persistent excitation signal sequence with zero mean and unit variance. $v(t)$ is the stochastic Gaussian noise with zero mean and $\sigma^2 = 0.001$. Then, the output $y(t)$ is generated by the corresponding fractional Hammerstein model.

The ADELS algorithm is used to estimate initial parameter values of the nonlinear and linear parts of the system. The population size of the algorithm is set to 40, the minimum and maximum values of the mutation operator

are 0.1 and 0.9, the minimum and maximum values of the crossover operator are 0.1 and 0.9, and the number of iterations of the algorithm is set to 200. The parameters estimated by the optimization algorithm converge to the true value, and the relative quadratic error is as follows:

$$RQE = \sqrt{\frac{\widehat{\theta} - \theta^2}{\theta^2}},
 \tag{48}$$

where θ is the true parameter vector and $\widehat{\theta}$ is the estimated parameter vector. The simulation results compared with other approaches are shown in Table 5. It reveals that more accurate initial values can be obtained by ADELS than other heuristic algorithms.

Simulation results obtained by different algorithms are given in Table 6. In the case of the same initial value obtained by ADELS, the accuracy of OAMRLS compared to KTRLS has been improved effectively.

The performance of the intelligent optimization algorithm to search for the initial value can be measured by Root Mean Square Error (RMSE) and Mean Square Error (MSE), which are defined as follows:

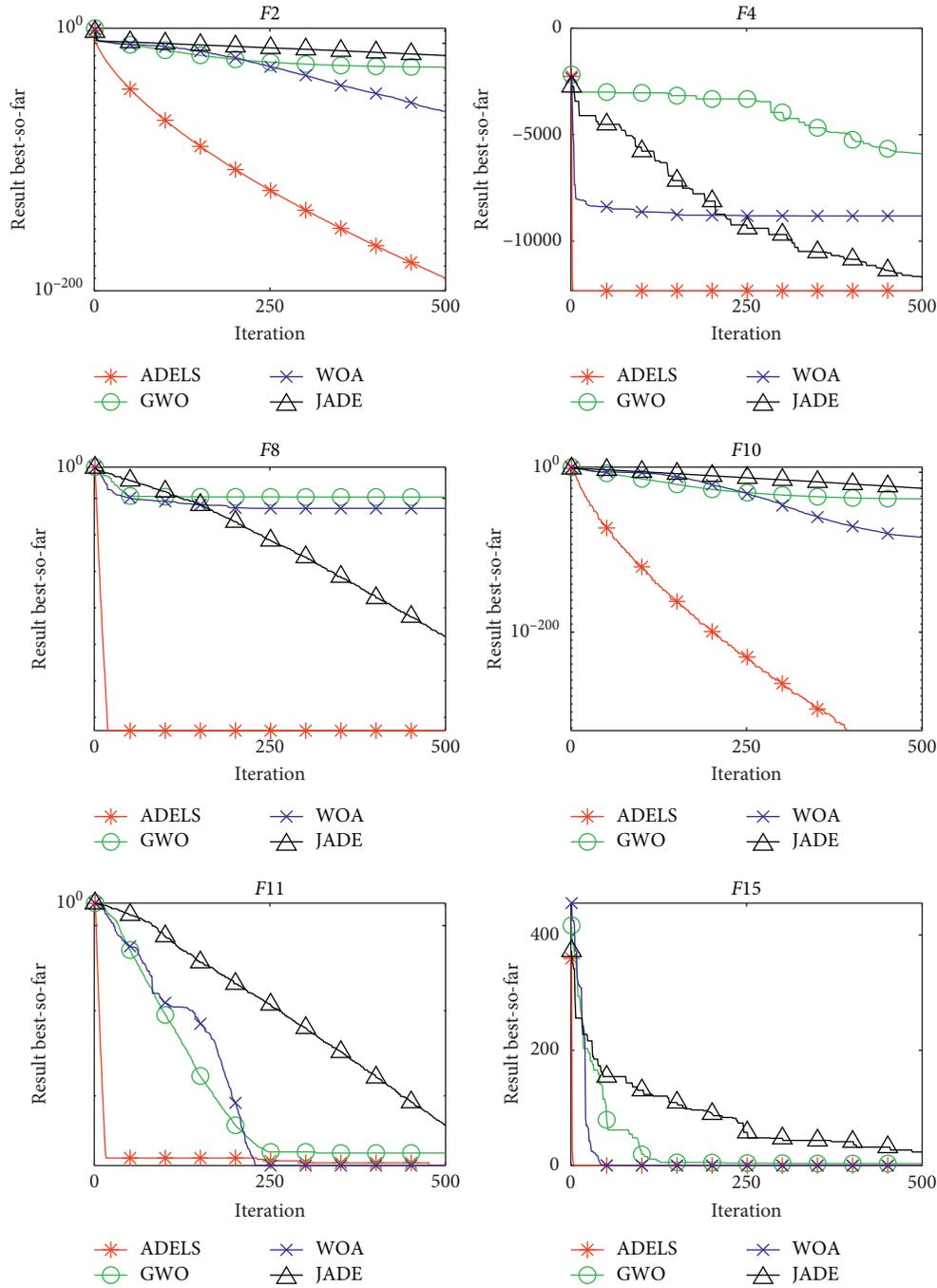


FIGURE 3: Comparison of convergence curves of WOA and literature algorithms obtained in some of the 30-dimensional benchmark problems.

$$\begin{aligned} \text{RMSE} &= \sqrt{\text{MSE}}, \\ \text{MSE} &= \frac{\sum_{n=1}^{i-1} (y_i - \hat{y}_i)^2}{n}, \end{aligned} \quad (49)$$

where y_i is the system output and \hat{y}_i is the estimated system output. The output estimation errors of different intelligent optimization algorithms are depicted in Figure 5.

Using the initial estimation value, OAMRLS and the steepest descent method are combined to estimate the accurate parameter vector. Figure 6 shows the actual outputs and the identified outputs regarding Algorithm 2, and the output error of the system identified by Algorithm 2 is shown in Figure 7. It can be seen from Figure 8 that the identified model can stably track the real output of the system under the step response. Figure 9 illustrates the trend of estimated fractional order convergence curve.

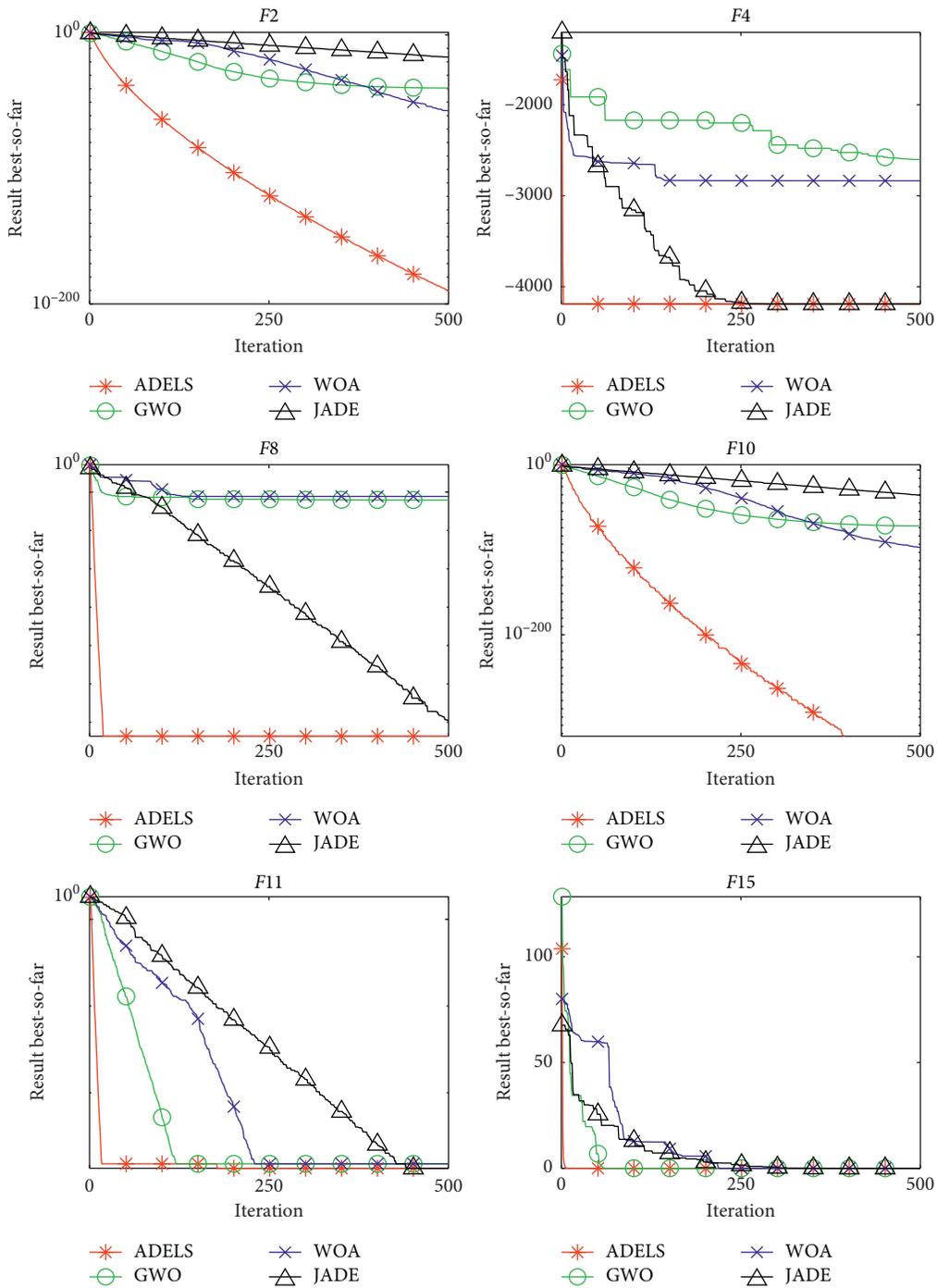


FIGURE 4: Comparison of convergence curves of WOA and literature algorithms obtained in some of the 10-dimensional benchmark problems.

TABLE 5: Initial values obtained by three different heuristic algorithms.

Approach		Parameter								Rqe
		a_1	a_2	b_0	b_1	c_2	c_3	c_4	α	
1	GWO	1.0863	1.7769	1.8239	1.8615	0.5159	0.2947	0.0932	0.2841	0.1459
2	JADE	0.3828	1.9053	1.6346	1.5920	0.5399	0.2728	0.0817	0.2603	0.3875
3	ADELS	1.1730	1.5398	1.8516	1.7289	0.5065	0.2882	0.0950	0.2970	0.0539

TABLE 6: Simulation results obtained by different algorithms.

Approach		Parameter								Rqe
		a_1	a_2	b_0	b_1	c_2	c_3	c_4	α	
1	GWO & OAMRLS	1.3233	1.4821	1.8995	1.7173	0.5039	0.2912	0.0953	0.3022	0.0133
2	JADE & OAMRLS	1.3699	1.4582	1.9044	1.7339	0.5065	0.2914	0.0946	0.3063	0.0332
3	ADELS & KTSG	1.1932	1.5674	1.8867	1.7289	0.5065	0.2882	0.0950	0.2969	0.0486
4	ADELS & OAMRLS	1.3060	1.4899	1.8977	1.7100	0.5029	0.2912	0.0956	0.3007	0.0069

- (1) Collect the input/output data $\{u(t), y(t)\}$;
- (2) Obtain the initial of unknown parameters by using Algorithm 1;
- (3) While ($t < \text{max number of iterations } N$)
- (4) Estimate the value of system coefficient according to the equation (31);
- (5) Update the fractional order according to the equation (38);
- (6) If the criterion function value J satisfies the error accuracy
- (7) Break;
- (8) End;
- (9) $t = t + 1$;
- (10) End while;
- (11) Postprocess results and visualization.

ALGORITHM 2: Identification process of the fractional Hammerstein model.

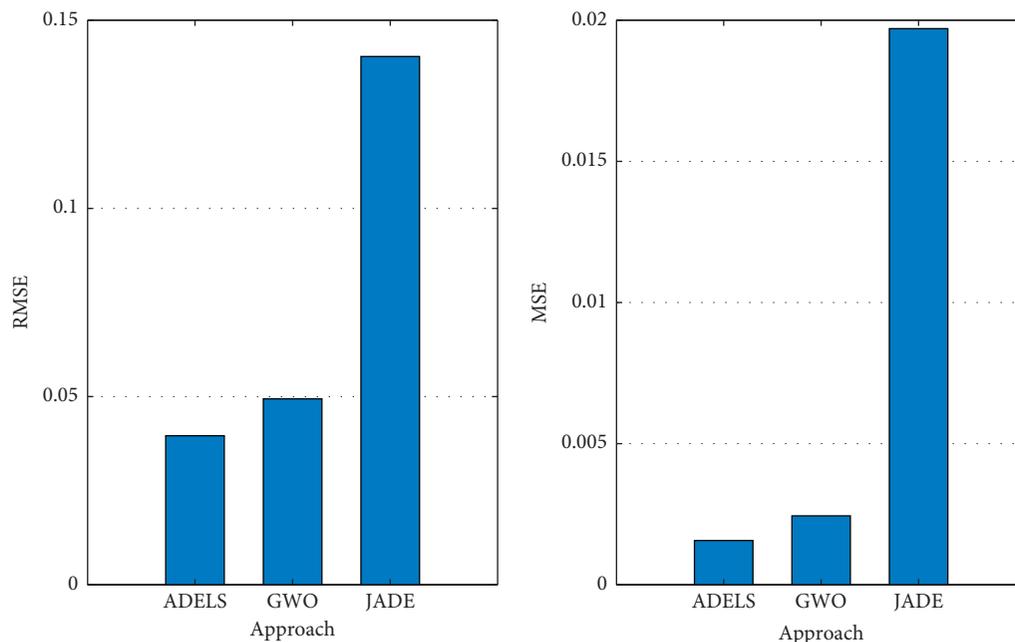


FIGURE 5: The output estimation errors of different intelligent optimization algorithms.

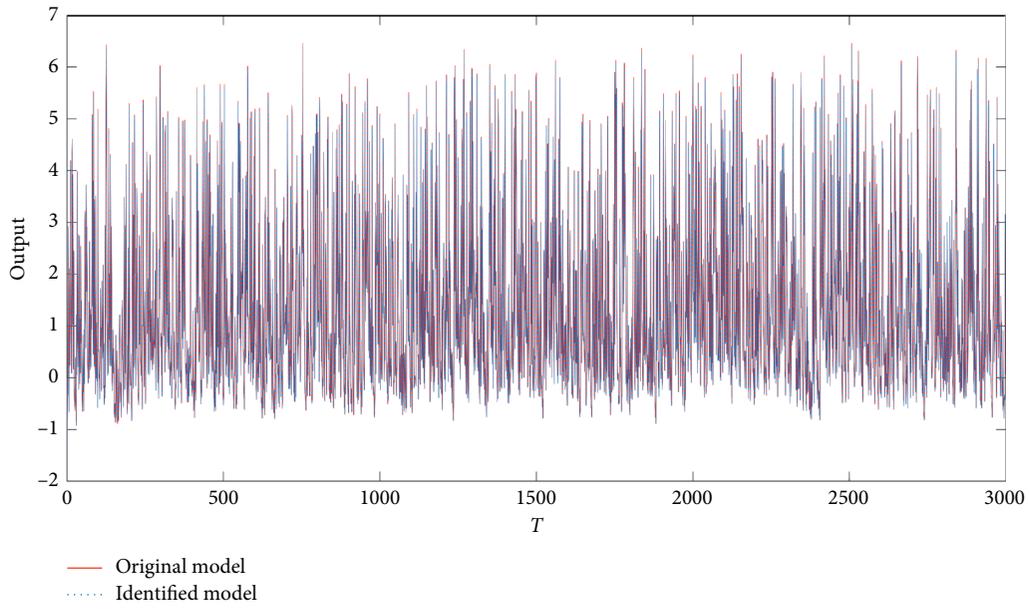


FIGURE 6: The actual outputs and the identified outputs regarding Algorithm 2.

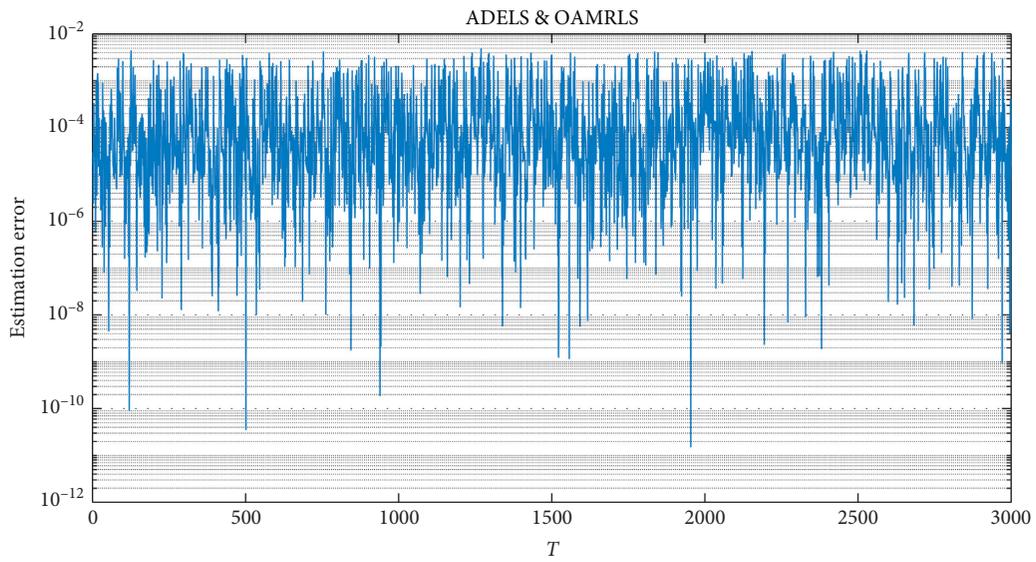


FIGURE 7: The error between the estimated and the actual outputs regarding Algorithm 2.

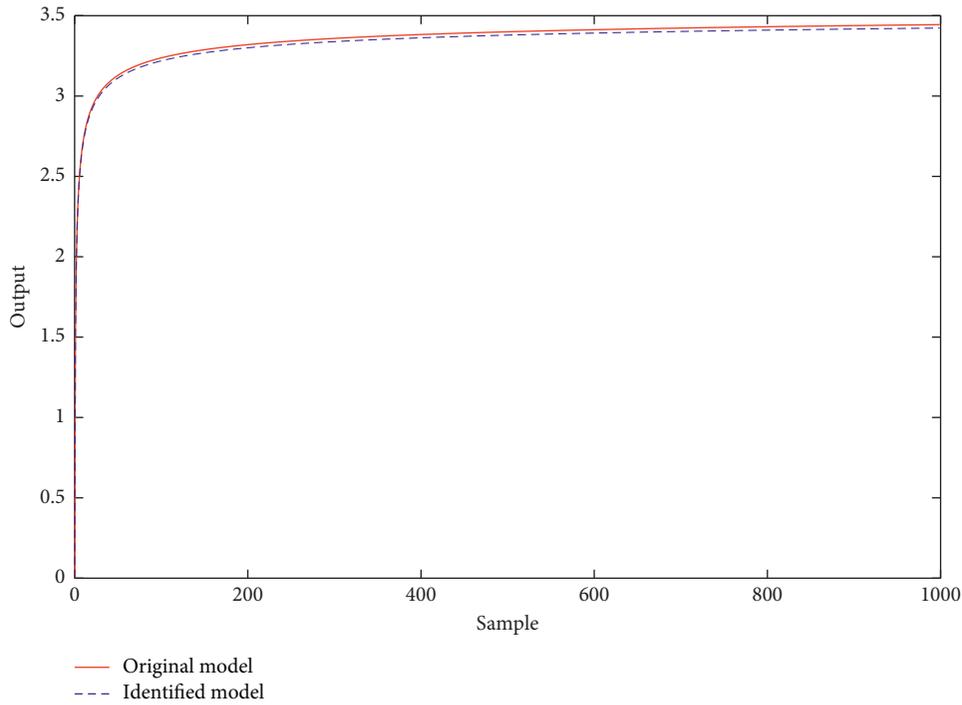


FIGURE 8: Step response of real system and estimation model regarding Algorithm 2.

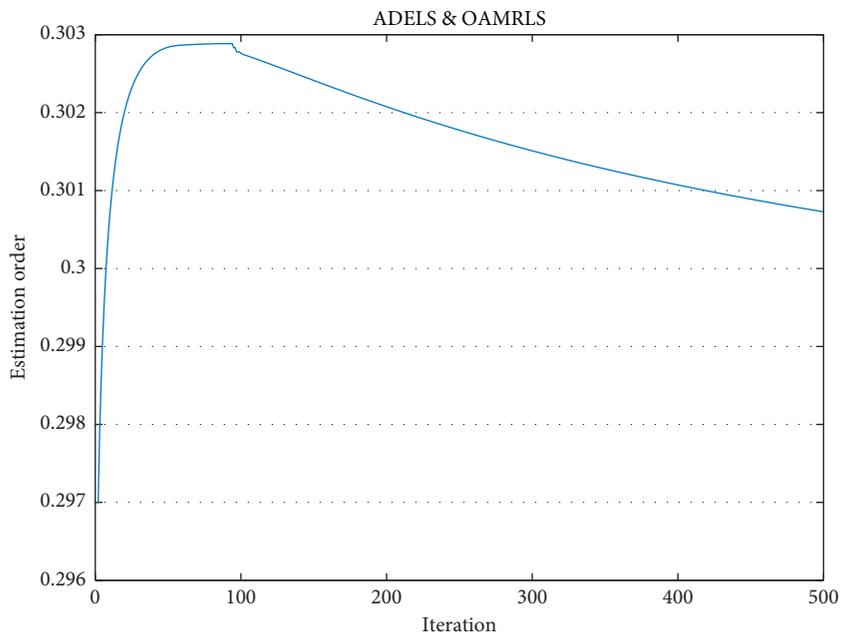


FIGURE 9: The estimated fractional order regarding Algorithm 2.

6. Summary

This paper outlines the identification methods for fractional order Hammerstein models. To improve the accuracy of identification, the heuristic algorithm is considered to search the initial value. Simulation results show that the heuristic algorithm is easy to fall into local optimum. To solve this

problem, ADELS algorithm is proposed. In this algorithm, combined with Chebyshev mapping, genes are adaptively searched in each generation, which leads to faster convergence and more accurate results. The initial values of the coefficients of the linear part and the nonlinear part are obtained by the ADELS. Furthermore, the fractional order of the transfer function obtains the initial value through the

ADELS. Then, the method of fractional order Hammerstein model coefficient estimation is proposed and compared with other algorithms. When the initial value is obtained, the parameters can be accurately identified by the corresponding OAMRLS algorithm. The fractional order is iterated by the steepest descent algorithm through derivation. Through comparison with other algorithms, the simulation results show the effectiveness of the proposed algorithm. The proposed methods in this paper can be applied to other literatures [51–56] such as parameter identification problems of different systems, engineering applications, fault diagnosis, and so on.

Data Availability

The code used in this paper was written in MATLAB 2018a; the data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61673004) and the Fundamental Research Funds of the Central Universities of China (XK1802-4).

References

- [1] F. Ding, L. Lv, J. Pan, X. Wan, and X.-B. Jin, “Two-stage gradient-based iterative estimation methods for controlled autoregressive systems using the measurement data,” *International Journal of Control, Automation and Systems*, vol. 18, no. 4, pp. 886–896, 2020.
- [2] X. Zhang, F. Ding, L. Xu, and E. Yang, “Highly computationally efficient state filter based on the delta operator,” *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 6, pp. 875–889, 2019.
- [3] X. Zhang, F. Ding, and E. Yang, “State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors,” *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1157–1173, 2019.
- [4] M. Li, X. Liu, and F. Ding, “Filtering-based maximum likelihood gradient iterative estimation algorithm for bilinear systems with autoregressive moving average noise,” *Circuits, Systems, and Signal Processing*, vol. 37, no. 11, pp. 5023–5048, 2018.
- [5] M. Chen, F. Ding, R. Lin, T. Y. Ng, Y. Zhang, and W. Wei, “Maximum likelihood least squares-based iterative methods for output-error bilinear-parameter models with colored noises,” *International Journal of Robust and Nonlinear Control*, vol. 30, no. 15, pp. 6262–6280, 2020.
- [6] X. Jin, H. X. Wang, X. Y. Wang et al., “Adaptive gradient-based iterative algorithm for multivariable controlled autoregressive moving average systems using the data filtering technique,” *Complexity*, vol. 2018, Article ID 9598307, 11 pages, 2018.
- [7] X.-B. Jin, H.-X. Wang, X.-Y. Wang, Y.-T. Bai, T.-L. Su, and J.-L. Kong, “Deep-learning prediction model with serial two-level decomposition based on Bayesian optimization,” *Complexity*, vol. 2020, Article ID 4346803, 14 pages, 2020.
- [8] Z. Zou, D. Zhao, X. Liu et al., “Pole-placement self-tuning control of nonlinear Hammerstein system and its application to pH process control,” *Chinese Journal of Chemical Engineering*, vol. 23, no. 8, pp. 1364–1368, 2015.
- [9] J. G. Smith, S. Kamat, and K. P. Madhavan, “Modeling of pH process using wavenet based Hammerstein model,” *Journal of Process Control*, vol. 17, no. 6, pp. 551–561, 2007.
- [10] R. Quachio and C. Garcia, “MPC relevant identification method for hammerstein and wiener models,” *Journal of Process Control*, vol. 80, pp. 78–88, 2019.
- [11] K.-K. Xu, H.-D. Yang, and C.-J. Zhu, “A novel extreme learning machine-based hammerstein-wiener model for complex nonlinear industrial processes,” *Neurocomputing*, vol. 358, pp. 246–254, 2019.
- [12] H. O. Garcés, J. M. Palma, A. J. Rojas, and V. Valdebenito, “Model analysis of heat transfer by hammerstein systems and optical instrumentation,” *IFAC-Papers OnLine*, vol. 52, no. 1, pp. 442–447, 2019.
- [13] S. Dong, L. Yu, W.-A. Zhang, and B. Chen, “Robust extended recursive least squares identification algorithm for hammerstein systems with dynamic disturbances,” *Digital Signal Processing*, vol. 101, 2020.
- [14] L. Ma and X. Liu, “Recursive maximum likelihood method for the identification of hammerstein ARMAX system,” *Applied Mathematical Modelling*, vol. 40, no. 13-14, pp. 6523–6535, 2016.
- [15] G. Giordano and J. Sjöberg, “Maximum likelihood identification of wiener-hammerstein system with process noise,” *IFAC-Papers OnLine*, vol. 51, no. 15, pp. 401–406, 2018.
- [16] M. A. Dhaifallah and D. T. Westwick, “Support vector machine identification of output error hammerstein models,” *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 13948–13953, 2011.
- [17] M. Al Dhaifallah and D. T. Westwick, “Identification of NARX hammerstein models based on support vector machines,” *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 4999–5004, 2008.
- [18] S. Cheng, Y. Wei, D. Sheng, Y. Chen, and Y. Wang, “Identification for Hammerstein nonlinear ARMAX systems based on multi-innovation fractional order stochastic gradient,” *Signal Processing*, vol. 142, pp. 1–10, 2018.
- [19] L. Li, X. Ren, and F. Guo, “Modified multi-innovation stochastic gradient algorithm for Wiener-Hammerstein systems with backlash,” *Journal of the Franklin Institute*, vol. 355, no. 9, pp. 4050–4075, 2018.
- [20] W. Mi, H. Rao, T. Qian, and S. Zhong, “Identification of discrete hammerstein systems by using adaptive finite rational orthogonal basis functions,” *Applied Mathematics and Computation*, vol. 361, pp. 354–364, 2019.
- [21] L. Jia, X. Li, and M.-S. Chiu, “The identification of neuro-fuzzy based MIMO hammerstein model with separable input signals,” *Neurocomputing*, vol. 174, pp. 530–541, 2016.
- [22] F. Li, L. Jia, D. Peng, and C. Han, “Neuro-fuzzy based identification method for hammerstein output error model with colored noise,” *Neurocomputing*, vol. 244, pp. 90–101, 2017.
- [23] W. Zhao, “Recursive identification for hammerstein systems with diminishing excitation signals,” *IFAC-PapersOnLine*, vol. 52, no. 24, pp. 151–157, 2019.

- [24] M. A. Taleb, O. Béthoux, and E. Godoy, "Identification of a PEMFC fractional order model," *International Journal of Hydrogen Energy*, vol. 42, no. 2, pp. 1499–1509, 2017.
- [25] S. Kumar and A. Ghosh, "Identification of fractional order model for a voltammetric E-tongue system," *Measurement*, vol. 150, Article ID 107064, 2020.
- [26] Q. Zhang, Y. Shang, Y. Li, N. Cui, B. Duan, and C. Zhang, "A novel fractional variable-order equivalent circuit model and parameter identification of electric vehicle Li-ion batteries," *Isa Transactions*, vol. 97, pp. 448–457, 2020.
- [27] G. Liu, L. Wang, W. L. Luo, J. K. Liu, and Z. R. Lu, "Parameter identification of fractional order system using enhanced response sensitivity approach," *Communications in Nonlinear Science and Numerical Simulation*, vol. 67, pp. 492–505, 2019.
- [28] Y. Lu, Y. Tang, X. Zhang, and S. Wang, "Parameter identification of fractional order systems with nonzero initial conditions based on block pulse functions," *Measurement*, vol. 158, 2020.
- [29] M. Chetoui and M. Aoun, "Instrumental variables based methods for linear systems identification with fractional models in the EVI context," in *Proceedings of the 16th International Multi-Conference on Systems, Signals & Devices*, pp. 90–95, IEEE, New York, NY, USA, March 2019.
- [30] K. Hammar, T. Djama, and M. Bettayeb, "Fractional hammerstein system identification using particle swarm optimization," in *Proceedings of the International Conference on Modelling*, Jeju Island, South Korea, September 2015.
- [31] K. Hammar, T. Djama, and M. Bettayeb, "Fractional Hammerstein system identification based on two decomposition principles," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 206–210, 2019.
- [32] L. Xu, B. Song, M. Cao, and Y. Xiao, "A new approach to optimal design of digital fractional-order PI D controller," *Neurocomputing*, vol. 363, pp. 66–77, 2019.
- [33] D. Miao, W. Chen, W. Zhao, and T. Demsas, "Parameter estimation of PEM fuel cells employing the hybrid grey wolf optimization method," *Energy*, vol. 193, 2020.
- [34] Y. Dai, D. Wu, S. Yu, and Y. Yan, "Robust control of underwater vehicle-manipulator system using grey wolf optimizer-based nonlinear disturbance observer and H-infinity controller," *Complexity*, vol. 2020, Article ID 6549572, 17 pages, 2020.
- [35] M. Du, Z. Cheng, Y. Zhang, and S. Wang, "Multiobjective optimization of tool geometric parameters using genetic algorithm," *Complexity*, vol. 2018, Article ID 9692764, 14 pages, 2018.
- [36] A. Wu and Z.-L. Yang, "An elitist transposon quantum-based particle swarm optimization algorithm for economic dispatch problems," *Complexity*, vol. 2018, Article ID 7276585, 15 pages, 2018.
- [37] P. Gong, W.-Q. Wang, F. Li, and H. C. So, "Sparsity-aware transmit beamspace design for FDA-MIMO radar," *Signal Processing*, vol. 144, pp. 99–103, 2018.
- [38] J. Pan, X. Jiang, X. Wan, and W. Ding, "A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems," *International Journal of Control, Automation and Systems*, vol. 15, no. 3, pp. 1189–1197, 2017.
- [39] J. Pan, W. Li, and H. Zhang, "Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control," *International Journal of Control, Automation and Systems*, vol. 16, no. 6, pp. 2878–2887, 2018.
- [40] I. Podlubny, L. Dorcak, and I. Kostial, "On fractional derivatives, fractional-order dynamic systems and PI $\lambda D\mu$ -controllers," in *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 5, pp. 4985–5499, San Diego, CA, USA, 1997.
- [41] A. Dzieliński and D. Sierociuk, "Stability of discrete fractional order state-space systems," *IFAC Proceedings Volumes*, vol. 39, no. 11, pp. 505–510, 2006.
- [42] I. Podlubny, *Fractional Differential Equations*, Academic Press, Cambridge, MA, USA, 1999.
- [43] S. Birogul, "Hybrid harris hawk optimization based on differential evolution (HHODE) algorithm for optimal power flow problem," *IEEE Access*, vol. 7, pp. 184468–184488, 2019.
- [44] K. Miao and Z. Wang, "Neighbor-induction and population-dispersion in differential evolution algorithm," *IEEE Access*, vol. 7, pp. 146358–146378, 2019.
- [45] Y.-X. Zhang and J. Gou, "Adaptive differential evolution algorithm based on restart mechanism and direction information," *IEEE Access*, vol. 7, pp. 166803–166814, 2019.
- [46] A. D. Polyanin and A. V. Manzhirov, *Handbook of Integral Equations*, CRC Press, Washington, DC, USA, 2nd edition, 2008.
- [47] J. Wang, Y. Wei, T. Liu, A. Li, and Y. Wang, "Fully parametric identification for continuous time fractional order hammerstein systems," *Journal of the Franklin Institute*, vol. 357, no. 1, pp. 651–666, 2020.
- [48] J. Zhang and A. C. Sanderson, "Adaptive differential evolution with optional external archive," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 945–958, 2009.
- [49] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [50] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.
- [51] J. Ding, "The hierarchical iterative identification algorithm for multi-input-output-error systems with autoregressive noise," *Complexity*, vol. 2017, Article ID 5292894, 11 pages, 2017.
- [52] J. Ding, J. Chen, J. Lin, and L. Wan, "Particle filtering based parameter estimation for systems with output-error type model structures," *Journal of the Franklin Institute*, vol. 356, no. 10, pp. 5521–5540, 2019.
- [53] X. Zhao, Z. Lin, Bo Fu, Li He, and Na Fang, "Research on automatic generation control with wind power participation based on predictive optimal 2-degree-of-freedom pid strategy for multi-area interconnected power system," *Energies*, vol. 11, 3325 pages, 2018.
- [54] L. Wang, H. Liu, Le Van Dai, and Y. Liu, "Novel method for identifying fault location of mixed lines," *Energies*, vol. 11, 2018.
- [55] R. Hong, C. Guorong, L. Yao, and H. Hongli, "Research on rfid fault diagnosis method based on SVM and PSO algorithm," in *Proceedings of the 2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*, pp. 122–126, Chongqing, China, December 2018.
- [56] B. Li, L. Zhang, B. Zhang, W. Wang, and Z. Hao, "Fault diagnosis method based on-IBP neural network," in *Proceedings of the 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFE-PROCESS)*, pp. 93–97, Xiamen, China, July 2019.