Received February 1, 2019, accepted February 25, 2019, date of publication March 5, 2019, date of current version April 11, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2902958*

# Complex-Valued Ordinary Differential Equation Modeling for Time Series Identification

**BIN YANG[1] AND WENZHENG BAO[2]**

[1]School of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160, China
[2]School of Information and Electrical Engineering, Xuzhou University of Technology, Xuzhou 221018, China

Corresponding author: Wenzheng Bao (baowz55555@126.com)

This work was supported in part by the Natural Science Foundation of China under Grant 61702445, Grant 61873270, Grant 61732012, and Grant 61861146002, and in part by the Ph.D. Programs Foundation of the Ministry of Education of China under Grant 201200721100400, in part by the Shandong Provincial Natural Science Foundation, China, under Grant ZR2018LF005 and Grant ZR2015PF007, in part by the PhD Research Startup Foundation of Zaozhuang University under Grant 2014BS13, and in part by the Zaozhuang University Foundation under Grant 2015YY02.

**ABSTRACT** Time series identification is one of the key approaches to dealing with time series data and discovering the change rules. Therefore, time series forecasting can be treated as one of the most challenging issues in this field. In order to improve the forecasting performance, we propose a novel time series prediction model based on a complex-valued ordinary differential equation (CVODE) to predict time series. A multi expression programming (MEP) algorithm is utilized to optimize the structure of the CVODE model. So as to achieve the optimal complex-valued coefficients, a novel optimization algorithm based on a complex-valued crow search algorithm (CVCSA) is proposed. The chaotic Mackey-Glass time series, small-time scale traffic measurements, Nasdaq-100 index, and Shanghai stock exchange composite index are utilized to evaluate the performance of our method. The results prove that our proposed method could predict more accurately than state-of-the-art real-valued neural networks and an ordinary differential equation (ODE). The CVCSA has faster convergence speed and stronger optimization ability than the crow search algorithm (CSA) and particle swarm optimization (PSO).

**INDEX TERMS** Complex-valued, ordinary differential equation, crow search algorithm, time series.

## I. INTRODUCTION

Time series forecasting method is to observe and learn time series data, and focus on discovering the change rule, so as to forecast the level that may be reached in the next period of time or in the next few years [1]–[4]. Time series data can hardly overcome some shortcomings, including high noise, randomness and nonlinearity. Considering those limitations, the modeling and prediction of such type of data are always research hot and difficulty points in this field [5], [6]. As the most classical modeling model, neural network (NN) has been utilized to solve these practical problems [7]–[10]. Generally, neural network can be regarded as a black box, and researchers could not understand the specific functional relationship among these variables, which make it difficult to analyze the internal mechanism of the system. Recently, mathematical formulations especially ordinary differential equations (ODE) have been proposed to model the relationships between input and output variables by intelligent computing methods, whose theory analysis can offer the guidance for practical systems [11]–[14].

Ordinary differential equation model has proposed to predict time series data and model the real systems [15]–[19]. Hass *et al.* proposed fast integration-based prediction bands to assess ODE model's uncertainty for cellular signaling [20]. Zjavka *et al.* proposed differential polynomial neural network (DPNN) with ODE substitutions to predict Short-term power load [21]. Linares et al. presented ODE and partial differential equation (PDE) to accurately quantify the basic phenomenon of methadone mass transfer during hemodialysis in order to provide the methadone dosage guidance for the doctors [22]. Cao *et al.* utilized ODE model to identify gene regulatory

The associate editor coordinating the review of this manuscript and approving it for publication was Yucong Duan.

network (GRN) with the time-series data [23]. Ilea et al. applied ODE model in molecular biology, which conforms well to many biological laws and relations [24].

Complex number could have richer representation ability than real value, and also improve the memory retrieval mechanism of noise robust. Many complex-valued models have been proposed to predict time series data [25]–[29]. Xiong *et al.* presented complex-valued radial basis function neural networks (FCRBFNNs) to forecast real interval stock price time series data [30]. Saoud et al. presented fully complex-valued wavelet network (FCWN) to predict the global solar irradiation and complex-valued gradient descent-learning method was proposed to search the complex-valued parameters [31]. These experiment results revealed that complex-valued models could perform better than the corresponding real-valued versions.

In order to accurately predict time series, a novel complex-valued time series prediction model based on ordinary differential equation (CVODE) is proposed. In a CVODE model, coefficients and functions are complex-valued. A novel hybrid evolutionary method based on Multi expression programming (MEP) algorithm and complex-valued crow search algorithm (CVCSA) is proposed to optimize the structure and complex-valued coefficients of CVODE models. The simulated time series from chaotic Mackey-Glass differential delay equation and three real time series datasets from small-time scale traffic measurements, Nasdaq-100 index and Shanghai stock exchange composite index are utilized as the standard datasets.

## II. METHOD

### A. COMPLEX-VALUED ORDINARY DIFFERENTIAL EQUATION

Complex-valued ordinary differential equation (CVODE) is the complex-valued version of ODE model. In a CVODE, input variables, output variables, functions and coefficients are complex-valued, which is described as follows.

$$\frac{dZ}{dt} = W \cdot F(Z, t). \tag{1}$$

where, $W$ is the complex-value coefficient vector, $F(\cdot)$ is complex-valued function and $Z$ is complex-valued input vector.

### B. MULTI EXPRESSION PROGRAMMING

#### 1) CHROMOSOME STRUCTURE

Multi expression programming (MEP) is a novel evolutionary algorithm based on structure proposed by Oltean, which is an improved version of genetic programming (GP) [32], [33]. Compared to GP, MEP algorithm has the following advantages: (1) each chromosome has the linear structure; (2) the chromosome could store multiple solutions of the problem, and generally the best solution is utilized as the individual's solution; (3) the chromosome expression can reuse the code, and no code area is rich and diverse. So MEP has been widely applied in many areas, such as image processing [34], bioinformatics [35], and time series prediction [36].

| | | |
|---|---|---|
| 1: | 0.6+1.2 *i* | |
| 2: | z3 | |
| 3: | sin | 2 1 |
| 4: | + | 1 3 |
| 5: | z2 | |
| 6: | × | 5 2 |
| 7: | z4 | |
| 8: | / | 6 4 |

**FIGURE 1. An example of the chromosome in MEP with eight genes. The first column denotes gene sequence number, the second column contains the function or terminal symbol selected randomly and the last column contains gene sequence numbers of operands of function symbol.**

| | |
|---|---|
| gene1: | 0.6+1.2 *i* |
| gene2: | z3 |
| gene3: | sin(z3) |
| gene4: | sin(z3)+ 0.6+1.2 *i* |
| gene5: | z2 |
| gene6: | z2×z3 |
| gene7: | z4 |
| gene8: | z2×z3/(sin(z3)+ 0.6+1.2 *i*) |

**FIGURE 2. The corresponding eight expressions of eight genes of the chromosome in MEP. The expression of gene with function symbol could be composed by function symbol and the expressions of gene sequence numbers of operands.**

In MEP algorithm, each chromosome contains multiple expressions, each of which is called a gene. There are three kinds of gene symbols: function symbol, terminal symbol and gene sequence number. In this paper, MEP is utilized to optimize the structure of CVODE model. Suppose that function symbol set is $F = \{+, -, \times, /, \sin, \cos, e^x\}$, terminal symbol set is $T = \{z_1, z_2, \ldots, z_5, R\}$ ($R$ is complex-valued constant), and the number of genes is set as 8. An example of the chromosome is created in Fig. 1. Each gene is a complex-valued expression or CVODE model, which is depicted in Fig. 2. Each gene is also a candidate solution or candidate CVODE model. In order to represent the complex-valued coefficients, each gene in chromosome of MEP is assigned a complex value, which is described in Fig. 3. The fitness values of all

| | | | | |
|---|---|---|---|---|
| 1: | 0.6+1.2 *i* | | | w1 |
| 2: | z3 | | | w2 |
| 3: | sin | 2 | 1 | w3 |
| 4: | + | 1 | 3 | w4 |
| 5: | z2 | | | w5 |
| 6: | × | 5 | 2 | w6 |
| 7: | z4 | | | w7 |
| 8: | / | 6 | 4 | w8 |

(*a*)

| | |
|---|---|
| gene1: | (0.6+1.2 *i* )×w1 |
| gene2: | w2×z3 |
| gene3: | w3×sin(w2×z3) |
| gene4: | w4×(w3×sin(z3)+ (0.6+1.2 *i*)×w1) |
| gene5: | w5×z2 |
| gene6: | w5×w2×z2×z3 |
| gene7: | w7×z4 |
| gene8: | w8×w5×w2×z2×z3/(w4×(w3×sin(z3)+ (0.6+1.2 *i*)×w1) ) |

(*b*)

**FIGURE 3. (a) Chromosome with complex-valued parameters in MEP. Each gene is assigned a complex-valued coefficient. (b) The corresponding eight expressions of eight genes of the chromosome with complex-valued parameters in MEP.**

genes in one chromosome are calculated, and the best fitness value is selected as the fitness value of the individual.

### 2) CHROMOSOME REPRODUCTION

In order to search for the best chromosome, three genetic operators, including selection, crossover and mutation, are utilized. Selection operation is utilized to select the better solution to the next generation. The crossover operation exchanges the partial genes of two chromosomes. Mutation operation is utilized to change some genes in chromosomes. The specific operator processes are introduced in Ref [32].

## *C. PARAMETERS OPTIMIZATION*

### 1) CROW SEARCH ALGORITHM

Crow search algorithm (CSA) is a novel meta heuristic optimization algorithm, which was proposed in 2016 and originates in simulating the behavior of crows searching for food [37]. Crows live in groups. Each crow remembers the best place of the food. Crows could observe where other birds hide their food and steal them when they leave. If a crow commits theft, it can protect its food against theft with a certain probability.

Because CSA is simple, and has less parameters and computation complexity, this algorithm performs better than bat algorithm (BA), genetic algorithm (GA), particle swarm optimization (PSO), group search optimization (GSO), mine blast algorithm (MBA), and harmony search (HS) [37], [38].

CSA is described as follows:

a) Initialize the population and parameters. Suppose that the population has $N$ crows $X = [X_1, X_2, \ldots, X_N]$, $M_i^t$ represents the best place of hiding food of *i-th* crow at *t-th* time point, $AP_i^t$ is awareness probability of *i-th* crow at *t-th* time point, and $fl_i^t$ is flight length of *i-th* crow at *t-th* time point.

b) Calculate the fitness values of the population $f(X)$.

c) Update the places of crow group. The update strategies contain two cases. The first case is that *j-th* crow does not know that *i-th* crow tracks it and *i-th* crow will get closer to the best place $M_j^t$ of hiding food of *j-th* crow at *t-th* time point. The second one is that if *j-th* crow finds that *i-th* crow tracks it, *j-th* crow could deliberately take *i-th* crow to a random location. The update formula is shown as follows.

$$X_{i+1}^t = \begin{cases} X_i^t + r_i \times fl_i^t \times (M_j^t - X_i^t), & r_j \geq AP_j^t \\ \text{a random position}, & \text{otherwise} \end{cases} \tag{2}$$

where, $r_i$ and $r_j$ are random variables in [0, 1].

d) Judge the feasibility of the new location of *i-th* crow and evaluate the fitness of the new location of *i-th* crow. The place of hiding food of *i-th* crow is updated with the following formula.

$$M_i^{t+1} = \begin{cases} X_i^{t+1}, & f(X_i^{t+1}) > f(M_i^t) \\ M_i^t, & \text{otherwise} \end{cases} \tag{3}$$

e) If the end condition is satisfied, algorithm stops; otherwise go to step c).

### 2) COMPLEX-VALUED CROW SEARCH ALGORITHM

Compared with the real-valued optimization method, complex-valued evolutionary methods have the higher population diversity and are easy to search the global optimal solution due to the higher dimension space. Thus complex-valued crow search algorithm (CVCSA) is firstly proposed to optimize the complex-valued coefficients of CVODE models in this paper.

In CVCSA, initialize $n$ complex-valued crows according to the number of complex-valued coefficients in CVODE model. Each crow is evaluated after the complex number is converted into real number. The real and imaginary part of each complex-valued individual are updated separately. The pseudo code of CVCSA is described in **Algorithm 1**.

## *D. THE FLOWCHART OF TIME SERIES PREDICTION WITH CVODE MODEL*

The flowchart of time series prediction is depicted in Fig. 4, which is introduced in detailed as follows.

*Step 1 (Data Preprocessing):* Due to that the input data of CVODE model are complex-valued and the predicted

**Algorithm 1** Pseudo code of CVCSA.

Initialize $N$ complex-valued crows $[X_1, X_2, \ldots, X_N](X_k = (x_k^{1,R} + x_k^{1,I}\ i, x_k^{2,R} + x_k^{2,I}\ i, \ldots, x_k^{n,R} + x_k^{n,I}\ i))$ with the $n$ dimension and value range $[V_{min}, V_{max}]$;
**for** $k = 1; k \le N; k{+}{+}$ **do**
  Calculate the fitness $F(X_k)$;
**end**
Initialize $(M_k^{1,R} + M_k^{1,I}\ i, M_k^{2,R} + M_k^{2,I}\ i, \ldots, M_k^{n,R} + M_k^{n,I}\ i))$ of each complex-valued crow (for example $k$);
**while** $t < t_{max}$ **do**
  **for** $k = 1; k \le N; k{+}{+}$ **do**
    Select a crow randomly ($j$);
    $r_j \leftarrow$ a random variable;
    **if** $r_j \ge AP_j^t$ **then**
      $X_i^R(t+1) \leftarrow X_i^R(t) + r_i * fl_i^R(t) * (M_j^R(t) - X_i^R(t))$;
      $X_i^I(t+1) \leftarrow X_i^I(t) + r_i * fl_i^I(t) * (M_j^I(t) - X_i^I(t))$;
    **end**
    **else**
      $X_i^R(t+1) \leftarrow$ a random position;
      $X_i^I(t+1) \leftarrow$ a random position;
    **end**
    Evaluate the new positions of the complex-valued crows;
    Update $M_k$ with Eq.(3);
  **end**
**end**
Store the best solution obtained;

**Algorithm 2** Real_valued input vector is converted into complex_valued one.

**Input** : Real_valued input data $[g_1, g_2, \ldots, g_m]$ ($m$ is the number of sample points);
**Output**: Complex_valued vector $[z_1, z_2, \ldots, z_m]$;
Calculate the maximum and minimum values of input data $g_{max}$ and $g_{min}$;
//Real data are converted into complex_valued values;
**for** $i = 1; i \le m; i{+}{+}$ **do**
  $\varphi_i = \frac{g_i - g_{min}}{g_{max} - g_{min}}(2\pi - \delta)$;
  $z_i = e^{i\varphi_i}$;
**end**

time series data are real-valued, it is necessary to convert time series into complex data. The specific process is shown in **Algorithm 2.**

*Step 2 (Training Phase):* Combined with complex-valued time series data, MEP is utilized to optimize the structure of CVODE model, and CVCSA is utilized to optimize the

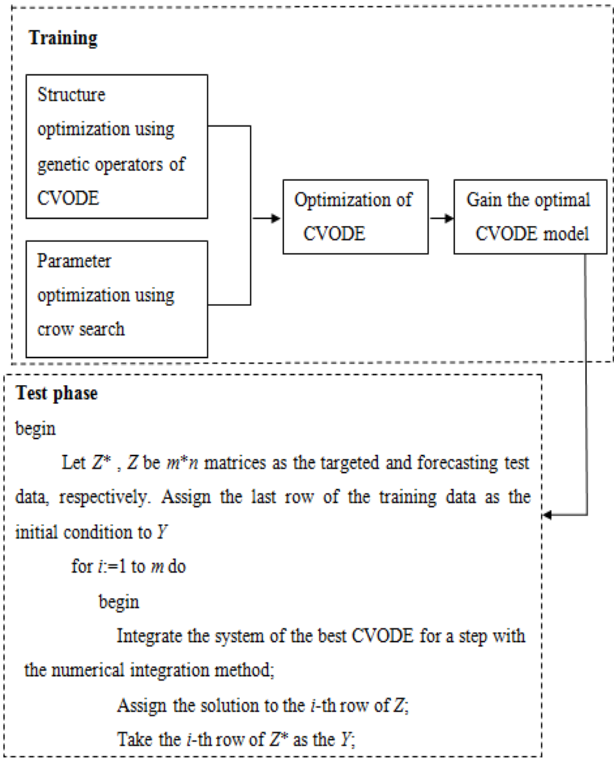



**FIGURE 4.** The flowchart of time series data prediction by CVODE model.

**Algorithm 3** Complex_valued input vector is converted into real_valued one.

**Input** : Complex_valued vector $[z_1, z_2, \ldots, z_m]$;
**Output**: Real_valued vector $[y_1, y_2, \ldots, y_m]$;
**for** $i = 1; i \le m; i{+}{+}$ **do**
  $arg\ z_i = \varphi_i$;
  $y_i = \frac{\varphi_i(g_{max} - g_{min})}{2\pi - \delta_i} + g_{min}$;
**end**

complex-valued coefficients of CVODE model. The specific optimization process is described as follows:

(1) Initialize the CVODE population, containing the complex-valued structure and parameters of the model.

(2) The fitness value of each CVODE model is calculated with **Algorithm 4**.

(3) The structure is optimized by MEP with selection, crossover and mutation operators. And the corresponding complex-valued parameter of each gene in MEP chromosome is evolved by complex-valued crow search algorithm.

(4) If the maximum generation is reached, algorithm ends; otherwise, go to (2).

*Step 3 (Test Phase):* The optimal CVODE model is utilized to predict the time series of the next time point. The training data at the last time point is selected as the initial input data of the optimal CVODE model.

**Algorithm 4** Calculate the fitness value of the *k-th* CVODE model.

**Input** : Real_valued output vector $[o_1, o_2, \ldots, o_m]$, Complex_valued input vector $Z = [z_1, z_2, \ldots, z_m]$, step size $h$, CVODE model $\frac{dZ}{dt} = H(t, Z)$, and fitness function $F(\cdot)$;

**Output**: Fitness value $f_k$ of the $k - th$ CVODE model;

```
f_k = 0;
for t = 1; t <= m; t++ do
    k1 = H(t, z_t);
    k2 = f(t + h/2, z_t + h * k1/2);
    k3 = f(t + h/2, z_t + h * k2/2);
    k4 = f(t + h, z_t + h * k3);
    zo = z_t + h * (k1+2k2+2k3+k4)/6;
    ro_t <- convert zo into real-valued data with Algorithm 3;
end
f_k = F(ro, Z);
```

## III. EXPERIMENTS

### A. DATA SETS AND EVALUATION CRITERIA

One simulated dataset and three real datasets are utilized to test the performance of CVODE model. The simulated time series are sampled from chaotic Mackey–Glass differential delay equation [39]. Three real time series data are from small-time scale traffic measurements, Nasdaq-100 index from 11 January 1995 to 11 January 2002, Shanghai stock exchange composite index (Shanghai index) from 04 January, 2011 to 01 January, 2015 [40], [41]. The parameters are selected empirically and listed in Table 1.

**TABLE 1. Parameters of our methods.**

| Parameters | Values |
|---|---|
| Number of genes in MEP | 8 |
| Population size in MEP | 30 |
| Maximum generation in MEP | 100 |
| Population size in CVCSA | 50 |
| Maximum generation in CVCSA | 200 |
| $AP$ in CVCSA | 0.1 |
| $fl$ in CVCSA | 2 |

The root mean squared error (RMSE), maximum absolute percentage error (MAP), means absolute percentage error (MAPE) are utilized to test the performance of the method. RMSE is selected as the fitness function of our method. Three criterions are defined as followed.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f^i_{target} - f^i_{forecast}\right)^2} \tag{4}$$

$$MAP = \max\left(\frac{|f^i_{target} - f^i_{forecast}|}{f^i_{forecast}} \times 100\right) \tag{5}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|f^i_{target} - f^i_{forecast}|}{f^i_{forecast}}\right) \times 100 \tag{6}$$

where $N$ represents the total time points, $f^i_{target}$ is the actual index value on day $i$ and $f^i_{forecast}$ is the forecasting index value on day $i$.

### B. CHAOS TIME SERIES PREDICTION

The chaotic Mackey–Glass differential delay equation has been utilized to model various physiological systems and its time series have been recognized as a common benchmark dataset for testing the performance of nonlinear system models. Mackey-Glass chaos time series are created with the following equation.

$$x(t+1) = (1-a)x(t) + \frac{bx(t-\tau)}{1 + x^{10}(t-\tau)}. \tag{7}$$

where $a = 0.1$, $b = 0.2$, and $\tau = 17$. In order to make the comparison with other methods fairly, input vector $[x(t-18), x(t-12), x(t-6), x(t)]$ is utilized to predict $x(t+6)$ and 1000 sample points are utilized. 500 samples are utilized for training, and the remaining 500 samples are utilized for prediction.

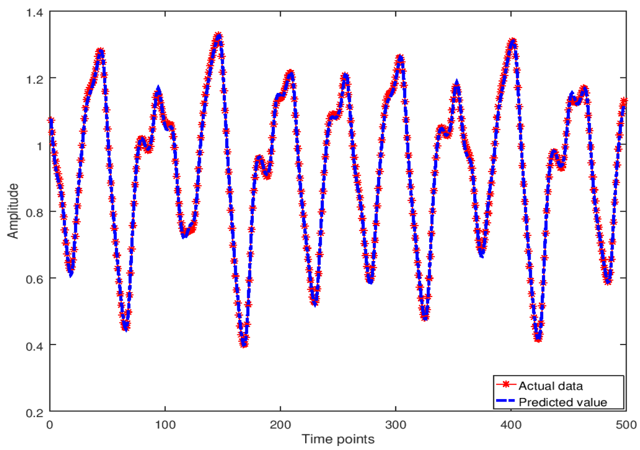



**FIGURE 5.** The predicted performance of Machey-Glass chaos time series.

The prediction results are shown in Fig. 5 and the correlation coefficient ($R^2$) between predicted data and real data is calculated, which is 0.99565. The results reveal that the predicted time series data are very close to the real ones. The predicted errors are depicted in Fig. 6, which show that the predicted errors mainly concentrate in the vicinity of zero $[-0.01, 0.01]$. The prediction results show that CVODE model could accurately predict chaotic time series data.

Table 2 lists the testing RMSE performance of eleven methods for Machey-Glass chaos time series. It could be clearly seen that CVODE model has smaller RMSE than the classical methods such as Auto-regressive model and

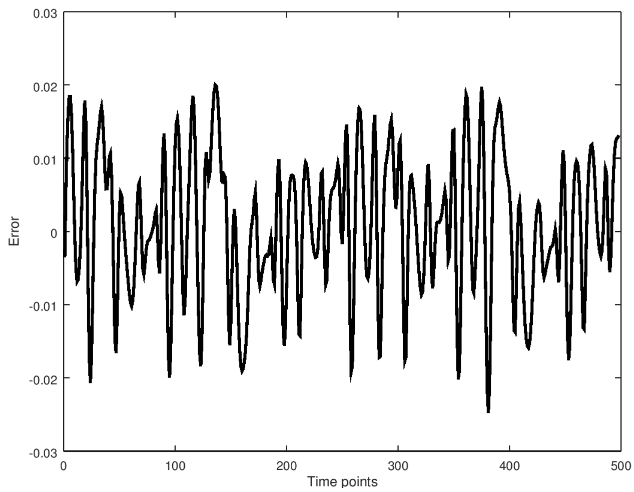



**FIGURE 6. The predicted errors of Machey-Glass chaos time series.**

**TABLE 2. Predicted performance of eleven methods with Machey-Glass chaos time series.**

| Methods | Testing RMSE |
|---|---|
| Auto-regressive model | 0.19 |
| Genetic algorithm and fuzzy system [42] | 0.049 |
| D-RBF[43] | 0.0125 |
| SORBF[44] | 0.0107 |
| RRBF [45] | 0.0260 |
| Stochastic ANN+PSO[47] | 0.0138 |
| ANN+PSO | 0.088036 |
| FNT | 0.024557 |
| ODE [15] | 0.01946 |
| **CVODE** | **0.009578** |

Genetic algorithm and fuzzy system, ANN, and state-of-the-art methods such as RBF neural networks (D-RBF, SORBF and RRBF), FNT and DE-ELM. Because of complex number, CVODE performs better than its real-valued version (ODE).

In order to test the prediction ability of the algorithm for noisy time series data, Gaussian noise is added to the Machey-Glass chaos time series, and the standard deviations of noises are 1%, 5% and 10%, respectively. The chaos data with three kinds of noises are shown in Fig. 7. As can be seen from Fig. 7, when the noise is up to 10%, the shape and characteristics of the time series have been changed. The predicted results and errors of chaos data with noises are depicted in Fig. 8. CVODE model could accurately predict the data with 1% and 5% noises and has the relatively small errors. Due to that 10% noise changes the data, our method has the bad performance.

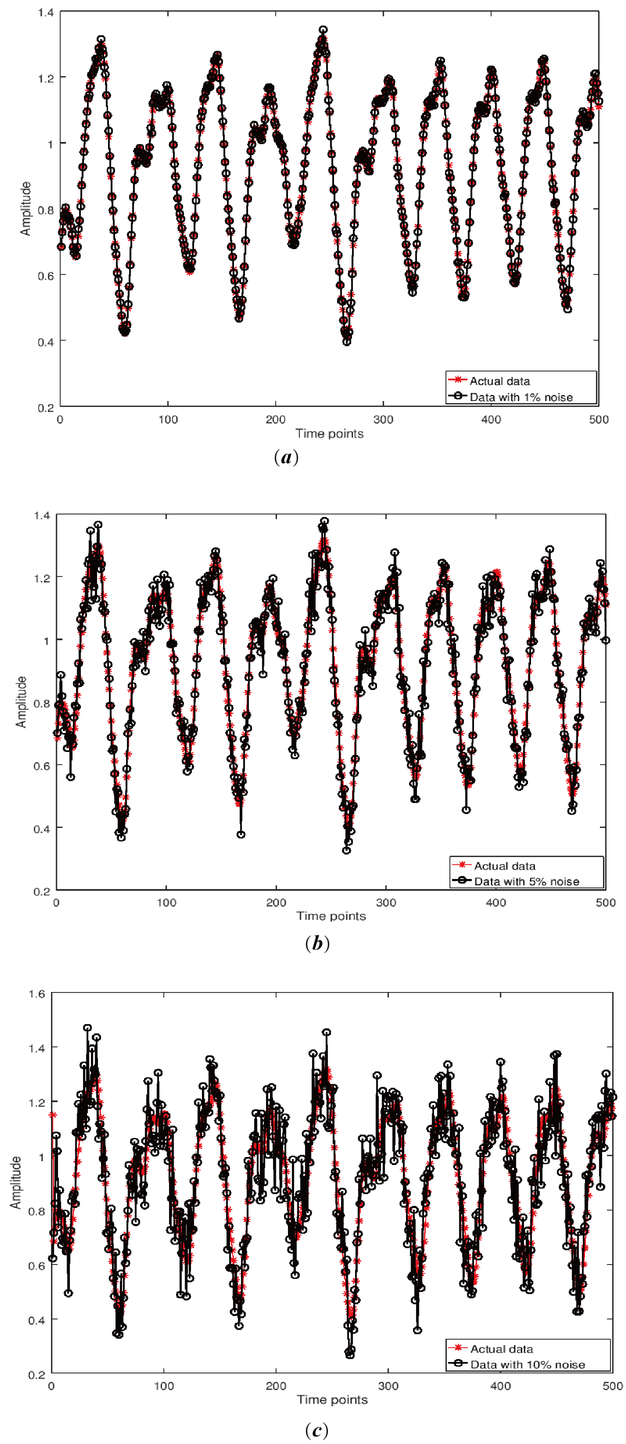
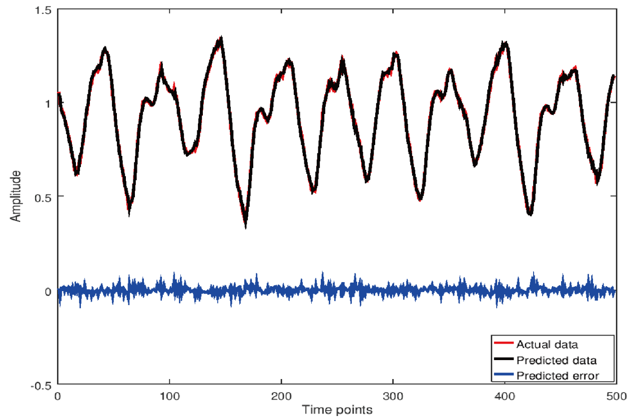
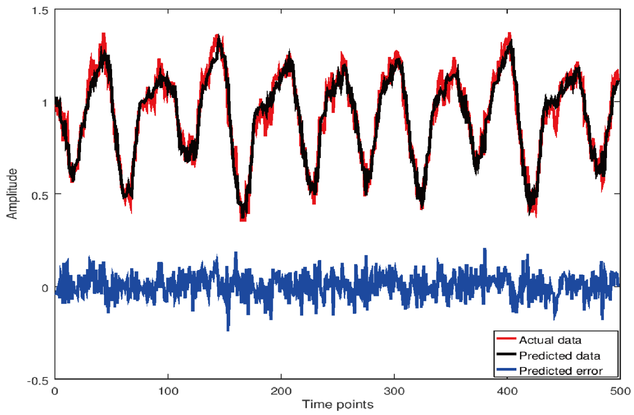



**FIGURE 7. Machey-Glass chaos time series with 1% noise (a), 5% noise (b) and 10% noise (c).**

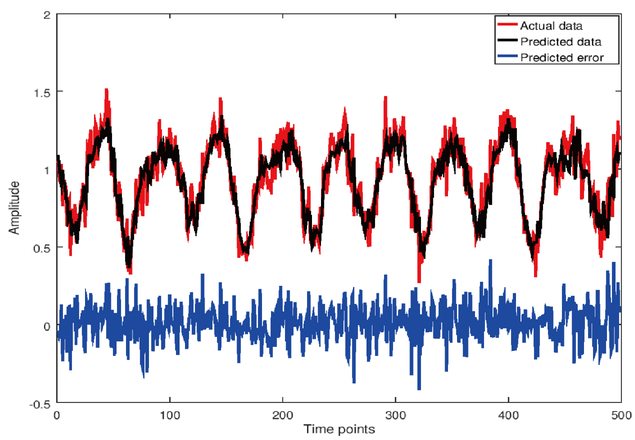
With the noise chaos time series data, CVODE model are made the comparison with ANN and ODE. The comparison results are listed in Table 3. CVODE has the smallest




(*a*)




(*b*)




(*c*)

**FIGURE 8. Prediction results of Machey-Glass chaos time series with 1% noise (a), 5% noise (b) and 10% noise (c).**

predicted RMSEs among three methods, which reveal that CVODE model has more robust performance than ANN and ODE.

**TABLE 3. RMSE performances of different methods with different noise rates.**

| Methods | 1% | 5% | 10% |
|---|---|---|---|
| ANN+PSO[47] | 0.089686 | 0.11589 | 0.154239 |
| ODE[15] | 0.021645 | 0.070025 | 0.134355 |
| CVODE | 0.015855 | 0.06524 | 0.124268 |

**TABLE 4. The optimal CVODE models of three datasets.**

| Type of data set | Optimal CVODE models |
|---|---|
| Traffic data | $\frac{dy}{dt} = (1.818099\text{-}0.426584\ i)\times(x_2+x_3)$ $+(\text{-}1.366544+0.853341\ i)x_1$ |
| Nasdaq-100 index | $\frac{dy}{dt} = (0.038086\text{-}0.080661\ i)\times x_4^2+$ $(13.587897+29.326102\ i)\times(x_4\text{-}x_1)$ |
| Shanghai index | $\frac{dy}{dt} = (\text{-}1.677217+0.453373\ i)\times\cos(x_2)+$ $(\text{-}3.130415+0.313458\ i)\times x_5\times x_9$ |

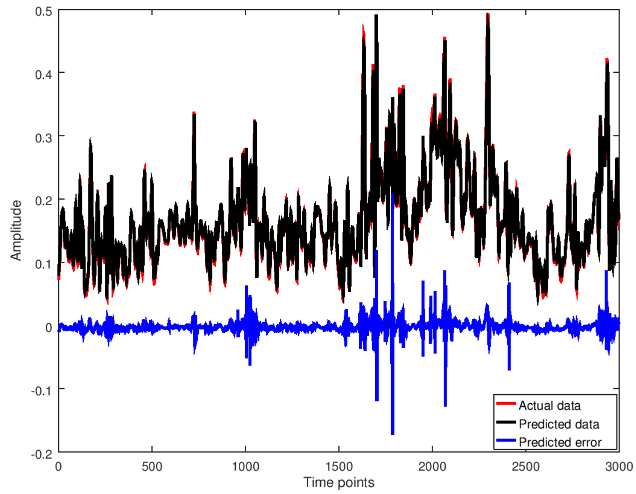



**FIGURE 9. Prediction results and errors with the traffic data.**

### C. REAL TIME SERIES PREDICTION

Nasdaq-100 index, Shanghai index and small-time scale traffic measurements are utilized to test the prediction performance of CVODE model. In order to make the comparison fairly, the partition of training and testing sets is the same as literatures. Through several runs, three optimal CVODE models achieved are listed in Table 4 for three kinds of datasets. The CVODE forms reveal that our proposed method could select automatically the important and proper features.

The prediction results and errors of three kinds of real datasets are depicted in Fig. 9, Fig. 10 and Fig. 11, respectively. From these figures, it can be seen that CVODE model

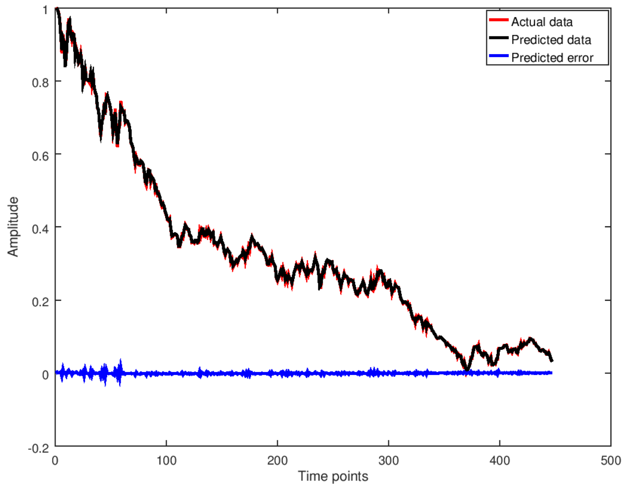



**FIGURE 10.** Prediction results and errors with the Nasdaq-100 index.

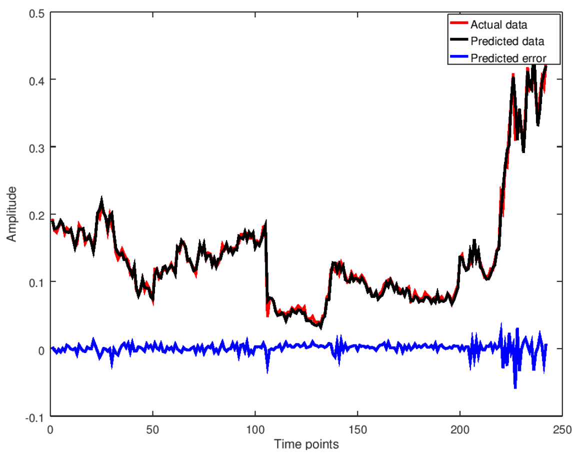



**FIGURE 11.** Prediction results and errors with the Shanghai index.

could predict the real time series data well and provide the small forecasting errors. To evaluate the performance of CVODE model, NN, WNN, FNT, ODE and S-system are also utilized to forecast three real time series datasets. The comparison results are listed in Table 5, Table 6 and Table 7, respectively. With traffic data and Nasdaq-100 index, CVODE has the highest RMSE and MAPE except for MAP. S-system has the best MAP performance, which reveals that the model has the smallest prediction error at some time point. With Shanghai index, CVODE obtains the most convincing performance among six state-of-the art methods. On the whole, our method performs best.

**TABLE 5.** Predicted results of five methods with traffic data.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| NN-PSO[48] | 0.0189 | 77.29 | 9.609 |
| FNT[40] | 0.0129 | 159.46 | 5.01 |
| ODE[15] | 0.0124 | 157.3 | 4.84 |
| S-system[50] | 0.0127 | **134.87** | 4.393 |
| CVODE | **0.01167** | 152.87 | **4.0456** |

**TABLE 6.** Predicted results of six methods with Nasdaq-100 index.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| NN-PSO[48] | 0.01864 | 141.363 | 6.528 |
| WNN-PSO[49] | 0.01789 | 152.754 | 6.570 |
| FNT[40] | 0.01882 | 98.107 | 6.205 |
| ODE[15] | 0.01675 | 85.682 | 6.127 |
| S-system[50] | 0.00832 | **45.398** | 5.768 |
| CVODE | **0.002931** | 76.586 | **1.1909** |

**TABLE 7.** Predicted results of six methods with Shanghai index.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| NN-PSO[48] | 0.01325 | 75.865 | 3.092 |
| WNN-PSO [49] | 0.01385 | 63.240 | 3.408 |
| FNT[40] | 0.01289 | 58.305 | 3.328 |
| ODE [15] | 0.01302 | 65.278 | 2.965 |
| S-system[50] | 0.01166 | 56.622 | 2.813 |
| CVODE | **0.008395** | **33.748** | **2.631** |

## IV. DISCUSSION

In this paper, CVCSA is proposed to optimize the complex-valued coefficients of CVODE model. In order to investigate the performance of CVCSA, we make the comparison experiments with CSA and PSO. Three real time series datasets are utilized as testing data. For these optimization algorithms, population size is set as 50, maximum generation is set as 500 and other parameters are set empirically. Evolution curves of fitness values with the three real datasets are described in Fig. 12, Fig. 13 and Fig. 14, respectively. From the results, it could be clearly seen that CVCSA could search the optimal solution faster than real-valued optimization algorithms (CSA and PSO), which is because that complex-valued encoding method can improve population diversity.

Through 30 runs, the averaged predicting performances (mean±standard deviation (SD)) of three real time series datasets by three optimization methods are listed in Table 8, Table 9 and Table 10, respectively. From the results, it could be clearly seen that CVCSA has the smallest mean and

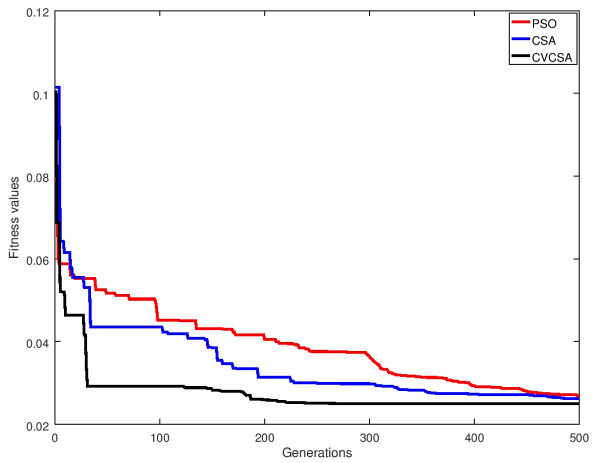



**FIGURE 12.** Evolution curves of fitness values with traffic data.

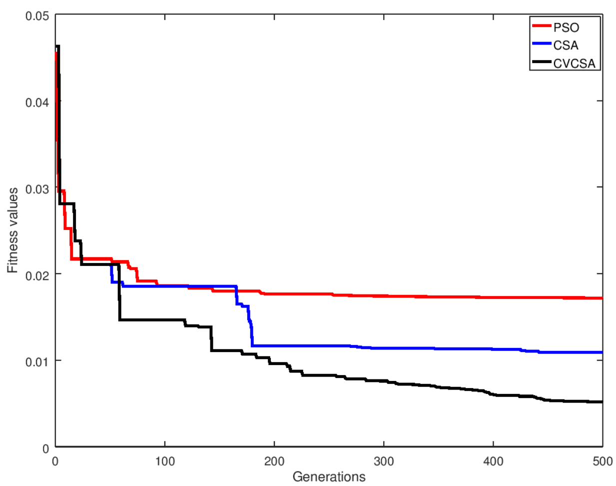



**FIGURE 13.** Evolution curves of fitness values with Nasdaq-100 index.

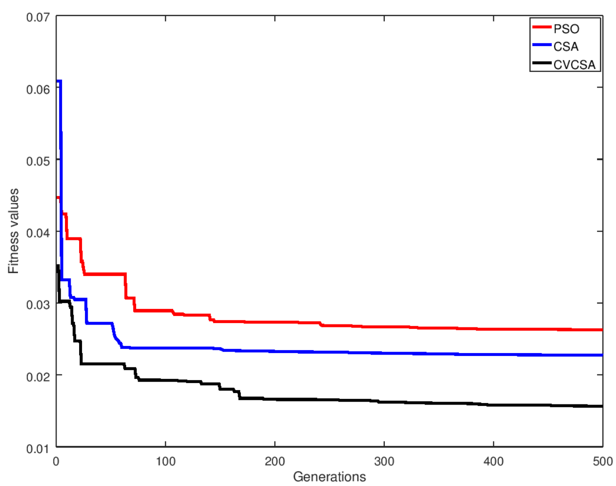



**FIGURE 14.** Evolution curves of fitness values with Shanghai index.

SD performances among three optimization methods, which reveal that CVCSA has stronger and more robust optimization ability than PSO and CSA.

**TABLE 8.** Averaged performances of three optimization methods for traffic data.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| PSO | $0.0149 \pm 1.26\times10^{-3}$ | $164.47 \pm 9.54$ | $4.809 \pm 1.208$ |
| CSA | $0.0137 \pm 7.45\times10^{-4}$ | $162.76 \pm 7.62$ | $4.531 \pm 0.891$ |
| CVCSA | $0.0129 \pm 2.41\times10^{-4}$ | $156.57 \pm 3.08$ | $4.208 \pm 0.239$ |

**TABLE 9.** Averaged performances of three optimization methods for Nasdaq-100 index.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| PSO | $0.00504 \pm 8.05\times10^{-4}$ | $89.31 \pm 5.31$ | $3.79 \pm 1.853$ |
| CSA | $0.00443 \pm 5.62\times10^{-4}$ | $83.62 \pm 2.89$ | $2.98 \pm 0.541$ |
| CVCSA | $0.00331 \pm 2.37\times10^{-4}$ | $79.03 \pm 1.22$ | $1.563 \pm 0.438$ |

**TABLE 10.** Averaged performances of three optimization methods for Shanghai index.

| | RMSE | MAP | MAPE |
|---|---|---|---|
| PSO | $0.012 \pm 1.05\times10^{-3}$ | $45.29 \pm 6.21$ | $3.38 \pm 1.09$ |
| CSA | $0.0101 \pm 8.71\times10^{-4}$ | $40.98 \pm 5.99$ | $3.32 \pm 0.94$ |
| CVCSA | $0.0097 \pm 3.28\times10^{-4}$ | $38.17 \pm 3.26$ | $3.02 \pm 0.58$ |

## V. CONCLUSIONS

In this paper, complex-valued ordinary differential equation (CVODE) is proposed to forecast the time series datasets with a novel hybrid evolutionary method based on Multi expression programming (MEP) algorithm and complex-valued crow search algorithm (CVCSA). Our proposed method has the better performance than real-valued neural networks (ANN, RBF, WNN and FNT), ELM and nonlinear ordinary differential equations. We investigate the prediction performance of CVCSA and the results reveal that CVCSA has faster convergence speed and stronger optimization ability than crow search algorithm and particle swarm optimization.

## REFERENCES

[1] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1506–1518, Nov. 2003.

[2] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decis. Support Syst.*, vol. 47, no. 2, pp. 115–125, May 2009.

[3] C. Lemke and B. Gabrys, "Meta-learning for time series forecasting and forecast combination," *Neurocomputing*, vol. 73, nos. 10–12, pp. 2006–2016, Jun. 2010.

[4] R. Webby and M. O'Connor, "Judgemental and statistical time series forecasting: A review of the literature," *Int. J. Forecasting*, vol. 12, no. 1, pp. 91–118, 1996.

[5] S. Ohgi, S. Morita, K. K. Loo, and C. Mizuike, "Time series analysis of spontaneous upper-extremity movements of premature infants with brain injuries," *Phys. Therapy*, vol. 88, no. 9, pp. 1022–1033, 2008.

[6] T. Mazloum and A. Sibille, "Analysis of secret key randomness exploiting the radio channel variability," *Int. J. Antennas Propag.*, vol. 2015, Sep. 2015, Art. no. 106360.

[7] M. Y. Rafiq, G. Bugmann, and D. J. Easterbrook, "Neural network design for engineering applications," *Comput. Struct.*, vol. 79, no. 17, pp. 1541–1552, 2001.

[8] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Proc. Int. Joint Conf. Neural Netw.*, vol. 1, 1989, pp. 593–605.

[9] A. Graves *et al.*, "Hybrid computing using a neural network with dynamic external memory," *Nature*, vol. 538, no. 7626, pp. 471–476, 2016.

[10] T. Wang, H. Gao, and J. Qiu, "A combined adaptive neural network and nonlinear model predictive control for multirate networked industrial process control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 2, pp. 416–425, Feb. 2016.

[11] B. Tran, B. Xue, and M. Zhang, "Genetic programming for feature construction and selection in classification on high-dimensional data," *Memetic Comput.*, vol. 8, no. 1, pp. 3–15, 2016.

[12] K. Krawiec, *Behavioral Program Synthesis With Genetic Programming* (Studies in Computational Intelligence). Cham, Switzerland: Springer, vol. 33, no. 1. 2016, pp. 101–112.

[13] Y. Mei, N. Su, and M. Zhang, "Evolving time-invariant dispatching rules in job shop scheduling with genetic programming," in *Proc. Eur. Conf. Genet. Program.*, 2017, pp. 147–163.

[14] M. M. Mostafa and A. A. El-Masry, "Oil price forecasting using gene expression programming and artificial neural networks," *Econ. Model.*, vol. 54, no. 1, pp. 40–53, 2016.

[15] Y. Chen, B. Yang, Q. Meng, Y. Zhao, and A. Abraham, "Time-series forecasting using a system of ordinary differential equations," *Inf. Sci.*, vol. 181, no. 1, pp. 106–114, 2011.

[16] Z. Li, P. Li, A. Krishnan, and J. D. Liu, "Large-scale dynamic gene regulatory network inference combining differential equation models with local dynamic Bayesian network analysis," *Bioinformatics*, vol. 27, no. 19, pp. 2686–2691, 2011.

[17] M. Gustafsson and M. Hörnquist, "Stability and flexibility from a system analysis of gene regulatorynetworks based on ordinary differential equations," *Open Bioinf. J.*, vol. 5, no. 1, pp. 26–33, 2011.

[18] S. Chen, A. Shojaie, and D. M. Witten, "Network reconstruction from high-dimensional ordinary differential equations," *J. Amer. Stat. Assoc.*, vol. 112, no. 520, pp. 1697–1707, 2017.

[19] S. Ali, F. M. Mahomed, and A. Qadir, "Complex Lie symmetries for scalar second-order ordinary differential equations," *Nonlinear Anal. Real World Appl.*, vol. 10, no. 6, pp. 3335–3344, 2009.

[20] H. Hass, C. Kreutz, J. Timmer, and D. Kaschek, "Fast integration-based prediction bands for ordinary differential equation models," *Bioinformatics*, vol. 32, no. 8, pp. 1204–1210, 2015.

[21] L. Zjavka and V. Snášel, "Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks," *Electr. Power Syst. Res.*, vol. 137, pp. 113–123, Aug. 2016.

[22] O. A. Linares *et al.*, "In silico ordinary differential equation/partial differential equation hemodialysis model estimates methadone removal during dialysis," *J. Pain Res.*, vol. 8, pp. 417–429, Jul. 2015.

[23] J. Cao, X. Qi, and H. Zhao, "Modeling gene regulation networks using ordinary differential equations," in *Next Generation Microarray Bioinformatics* (Methods in Molecular Biology). New York, NY, USA: Humana Press, vol. 802. 2012, pp. 185–197.

[24] M. Ilea, M. Turnea, and M. Rotariu, "Ordinary differential equations with applications in molecular biology," *Rev. Med. Chirurgicala Soc. Med. Natural Iasi*, vol. 116, no. 1, pp. 347–352, 2012.

[25] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, and S. Subramanian, "Deep complex networks," in *Proc. ICLR*, 2018, pp. 1–19.

[26] X. Chen, Z. Tang, and S. Li, "An modified error function for the complex-value backpropagation," *Neural Inf. Process. Lett. Rev.*, vol. 8, no. 1, pp. 1–8, 2005.

[27] C. You and D. Hong, "Nonlinear blind equalization schemes using complex-valued multilayer feedforward neural networks," *IEEE Trans. Neural Netw.*, vol. 9, no. 6, pp. 1442–1455, Nov. 1998.

[28] S. Chen, S. Mclaughlin, and B. Mulgrew, "Complex-valued radial basis function network, Part II: Application to digital communications channel equalisation," *Signal Process.*, vol. 36, no. 2, pp. 175–188, 1994.

[29] C.-A. Popa, "Complex-valued deep belief networks," in *Advances in Neural Networks* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, vol. 10878. 2018, pp. 72–78.

[30] T. Xiong, Y. Bao, Z. Hu, and R. Chiong, "Forecasting interval time series using a fully complex-valued RBF neural network with DPSO and PSO algorithms," *Inf. Sci.*, vol. 305, pp. 77–92, Jun. 2015.

[31] L. S. Saoud, F. Rahmoune, V. Tourtchine, and K. Baddari, "Fully complex valued wavelet network for forecasting the global solar irradiation," *Neural Process. Lett.*, vol. 45, no. 2, pp. 475–505, 2016.

[32] M. Oltean and C. Groşan, "Evolving evolutionary algorithms using multi expression programming," in *Proc. Eur. Conf. Artif. Life*, 2003, pp. 651–658.

[33] M. Bagheri, A. H. Gandomi, M. Bagheri, and M. Shahbaznezhad, "Multi-expression programming based model for prediction of formation enthalpies of nitro-energetic materials," *Expert Syst.*, vol. 30, no. 1, pp. 66–78, 2013.

[34] W. Wang, W. Lin, and Q. Li, "Image retrieval based on multi expression programming algorithms," *Crit. Criminology*, vol. 21, no. 2, pp. 157–176, 2013.

[35] A. Baykasoğlu and L. Özbakir, "MEPAR-miner: Multi-expression programming for classification rule mining," *Eur. J. Oper. Res.*, vol. 183, no. 2, pp. 767–784, 2007.

[36] C. Grosan, A. Abraham, V. Ramos, and S. Y. Han, "Stock market prediction using multi expression programming," in *Proc. Conf. Artif. Intell.*, Dec. 2005, pp. 73–78.

[37] A. Askarzadeh, "A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm," *Comput. Struct.*, vol. 169, pp. 1–12, Jun. 2016.

[38] G. I. Sayed, A. E. Hassanien, and A. T. Azar, "Feature selection via a novel chaotic crow search algorithm," *Neural Comput. Appl.*, vol. 31, no. 1, pp. 171–188, 2017.

[39] R. S. Crowder, III, "Predicting the Mackey-glass timeseries with cascade-correlation learning," in *Connectionist Models*. Pittsburgh, PA, USA: Carnegie Mellon Univ., 1990, pp. 117–123.

[40] Y. Chen, B. Yang, and A. Abraham, "Flexible neural trees ensemble for stock index modeling," *Neurocomputing*, vol. 70, nos. 4–6, pp. 697–703, Jan. 2007.

[41] Q.-F. Meng, Y.-H. Chen, and Y.-H. Peng, "Small-time scale network traffic prediction based on a local support vector machine regression model," *Chin. Phys. B*, vol. 18, no. 6, pp. 2194–2199, 2009.

[42] D. Kim and C. Kim, "Forecasting time series with genetic fuzzy predictor ensemble," *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 4, pp. 523–535, Nov. 1997.

[43] J. F. Qiao, H. G. Han, and J. F. Qiao, "Optimal structure design for RBFNN structure," *Acta Autom. Sinica*, vol. 36, no. 6, pp. 865–872, 2010.

[44] H. G. Han, Q. L. Chen, and J. F. Qiao, "An efficient self-organizing RBF neural network for water quality prediction," *Neural Netw.*, vol. 24, no. 7, pp. 717–725, 2011.

[45] H. G. Han, Y. Li, Y. N. Guo, and J. F. Qiao, "A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network," *Appl. Soft Comput.*, vol. 38, no. 1, pp. 477–486, 2016.

[46] J. Li and D. C. Li, "Wind power time series prediction using optimized kernel extreme learning machine method," *Acta Phys. Sinica*, vol. 65, no. 13, 2016, Art. no. 130501.

[47] C. H. López-Caraballo, I. Salfate, J. A. Lazzús, P. Rojas, M. Rivera, and L. Palma-Chilla, "Mackey-glass noisy chaotic time series prediction by a swarm-optimized neural network," *J. Phys., Conf. Ser.*, vol. 720, no. 1, 2016, Art. no. 012002.

[48] Y. H. Chen and A. Abraham, "Hybrid learning methods for stock index modeling," in *Artificial Neural Networks in Finance, Health and Manufacturing: Potential and Challenges*. J. Kamruzzaman, R. K. Begg, and R. A. Sarker, Eds. Hershey, PA, USA: IdeaGroup, 2006.

[49] Y. Chen, B. Yang, and J. Dong, "Time-series prediction using a local linear wavelet neural network," *Neurocomputing*, vol. 69, nos. 4–6, pp. 449–465, 2006.

[50] W. Zhang and B. Yang, "Stock market forecasting using S-system model," in *Advances in Intelligent Systems and Interactive Applications*. Cham, Switzerland: Springer, 2017, pp. 397–403.

Authors' photographs and biographies not available at the time of publication.