Modeling for chaotic time series based on linear and nonlinear framework: Application to wind speed forecasting

Ping Jiang[a], Biao Wang[a,*] Hongmin Li[a], Haiyan Lu[b]

[a] *School of Statistics, Dongbei University of Finance and Economics, Dalian, China*
[b] *School of Software, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia*
* Corresponding author. Address: School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China
Tel.: +86 15382250021.
E-mail address: wangbiaofugg@163.com.

Abstract:

Wind-speed forecasting plays a crucial part in improving the operational efficiency of wind power generation. However, accurate forecasts are difficult owing to the uncertainty of the wind speed. Although numerous investigations of wind-speed forecasting have been performed, many of the previous studies used wind-speed data directly to make forecasts, which were rarely based on the structural characteristics of the data. Therefore, in this study, a hybrid linear-nonlinear modeling method based on the chaos theory was successfully employed to capture the linear and nonlinear factors hidden in chaotic time series. Before the forecast, the noise in the data was removed using a decomposition algorithm. Then, through the phase-space reconstruction, the one-dimensional time series were extended to the multi-dimensional space to determine the utilization form of the data. Finally, Holt's exponential smoothing based on the firefly optimization algorithm and support vector regression were combined to predict the wind speed. The experimental results show that the proposed model is not only better than the comparison models but also has great application potential in the wind power generation system.

Keywords: Wind speed; Data characteristics; Linear and nonlinear; Chaotic time series; Phase space reconstruction

1. Introduction

With the rapid increase in the demand for energy, the potential of wind power is gaining increased attention. The use of wind energy to generate electricity can reduce the consumption of traditional energy resources as well as environmental pollution. However, the effective use of wind energy is not simple. Owing to its inherent discontinuity and limited predictability, the integration of the electric system poses a profound challenge to operations and planning practices [1]. The power system will benefit from efficient and accurate wind-speed forecasting for mitigating unstable operation, especially when the generated electricity is incorporated into the grid system [2]. Additionally, a correlation study showed that if the wind-speed prediction accuracy is improved by 10%, the expected value of wind power generation can be increased by approximately 30% [3]. Thus, wind-speed forecasting has become a popular research topic owing to its importance for the efficiency of wind power generation [4].

In view of the significant demand, numerous researchers have studied wind-speed predictions both theoretically and practically. Wind-speed forecasting models are diverse and can be classified into three major categories [5]: physical approaches, statistical methods, and artificial intelligence methods. Physical models based on numerical weather prediction (NWP) are often used for long-term wind-speed forecasting [6]. However, the complexity of the calculation process and high costs limit the application of NWP models. In contrast, the statistical model is based on the sample data and involves obtaining the functional relationship among the variables, usually via a conventional time series analysis, e.g., exponential smoothing (ES) [7], autoregressive moving average [8], and autoregressive integrated moving average (ARIMA) [9]. These are all inferred from the historical data to determine the trend of the wind speed. However, the factors affecting the wind-speed series are diverse and require complex functions to capture the functional relationships between variables [10]. Fortunately, with the rapid development of computer technology, various artificial intelligence prediction methods have been applied to wind-speed forecasting, such as the artificial neural network (ANN) [10-12], extreme learning machine (ELM) [13], and support vector machine (SVM) [14]. They can capture hidden nonlinear relationships in data through machine learning and pattern recognition [15]. However, each method has insurmountable disadvantages owing to its inherent properties, and it is difficult for a single prediction model to completely capture hidden information in a time series [16].

Statistical methods based on assumptions tend to exhibit poor prediction performance for predicting highly nonlinear time series. To highlight the advantages of the model and compensate for its shortcomings, it is effective to establish a model structure combining linearity and nonlinearity. In 2003, Zhang et al. [17] proposed a model combining the ARIMA and an ANN to capture linear and nonlinear factors in a time series. Since then, the combination of linear and nonlinear models has been widely used for wind-speed forecasting. In addition, various optimization algorithms are used to optimize the initial weights and thresholds of the ANN, which improves the prediction performance of the model [18]. In 2017, Zhang et al. [19] applied an ELM based on a backtracking search optimization algorithm to wind-speed forecasting and experimentally demonstrated its good prediction performance.

Furthermore, most of the previous studies focused only on improving the accuracy of the model, ignoring the characteristics of the wind-speed time series. Thus, to discover the useful information in the data, it is necessary to preprocess and characterize the data before prediction [20]. Therefore, data denoising methods such as empirical mode decomposition (EMD) aim to reduce random disturbances in the data sequence

and increase the prediction accuracy, and are considered to be applicable to wind-speed forecasting [21]. Compensating for the deficiencies of EMD, ensemble EMD (EEMD) is a noise-assisted data analysis method that has been widely applied to filter out the disturbing factors of wind-speed series [22, 23]. For instance, Wang et al. [24] utilized an EEMD method in which the raw data are decomposed into signals with different frequencies. Although EEMD has been greatly improved and widely applied in many areas, it has the drawback that noise cannot be effectively removed. Thereby, the complete EEMD with adaptive noise (CEEMDAN) was proposed by Torres et al. [25]. Zhang et al. [26] proposed a combined forecasting model. CEEMDAN is employed to divide the original wind-speed data into finite subsequences, and then the full-parameter fraction forecasting model is applied to forecast each decomposition. Ye et al. [27] compared the four decomposition methods (EMD, EEMD, CEEMD, and CEEMDAN) and combined the models to predict the wind speed. The results showed that CEEMDAN-support vector regression (SVR) had good prediction performance and a high operational efficiency.

The foregoing review of the literature indicates the following. (1) Physical forecasting methods need to collect a large amount of data, mainly applied to regional prediction. Thus, they do not have high economic value for wind power companies. (2) The traditional statistical models do not uncover information hidden in data with noise components and cannot effectively improve the forecasting accuracy. (3) ANNs have difficulty determining a suitable network structure, and may easily enter an overfitting state. (4) The artificial intelligence model improves the reliability of the neural network structure by combining optimization algorithms but ignores the importance of the data preprocessing stage.

In light of the aforementioned limitations, this study proposes a novel hybrid model (CEE-CC-FHS) for wind-speed forecasting. First, the chaotic characteristics of wind-speed time series are determined by calculating the maximum Lyapunov exponent (MLYE) [28]. Numerical simulation results show that the time series can be used as a chaotic system for predictive analysis [29,30]. The raw wind-speed sequence is decomposed into a finite component set via the CEEMDAN method to eliminate the interference factor, and then the information contained in the sequence is mined by determining the extended form of the sequence using the C-C method. On this basis, a linear and nonlinear modeling framework and a hybrid model based on firefly optimization were proposed. The simulation results were compared with those for other methods. The experimental results indicate that the hybrid model can has excellent prediction performance.

The novelty and originality of the proposed method are explained as follows.

- A novel combined model based on linear and nonlinear frameworks is constructed without considering the statistical distribution. The experimental simulation results confirm that the proposed model has excellent prediction accuracy.
- An outstanding model input structure can be obtained. Through phase-space theory, the single-dimensional wind-speed time sequences are converted into a matrix time series to improve the generalization ability. The matrix time series can more clearly describe the sequence characteristics, increasing the accuracy of the forecasting results.
- The hybrid "linear and nonlinear" modeling framework based on the firefly optimization algorithm can capture the linear and nonlinear modes contained in the time series well.
- The practicability and effectiveness of the proposed forecasting model were confirmed. A feasible method for accurate wind-speed prediction is provided.

The structure of this paper is as follows. The theory of the combined forecasting models for wind-speed forecasting is introduced in Section 2. The proposed combined forecasting model is described in Section 3. Detailed experimental analysis is presented in Section 4. Analysis and discussions are presented in Section 5. Conclusions are presented in the last section.

2. Methodology

In this section, the methods and principles of CEEMDAN, PSR, the firefly algorithm (FA), Holt, and SVR are detailed.

2.1. CEEMDAN

Huang et al. [31] proposed the application of EMD for analyzing nonlinear data. As an improved method for EMD, EEMD was first proposed in [32]. EEMD divides the sequence into different components when the white noise is uniformly distributed throughout the time–frequency space. The details of EEMD are presented in [33].

Different distributions of white Gaussian noise with zero mean and unit variance are represented as $w^i(n), i = 1, 2, ..., I$. The standard deviation of the white Gaussian noise is ε_0. The decomposition steps of CEEMDAN are as follows [34].

Step 1: The above I signals are decomposed via EMD to obtain their first mode and calculate $x(n) + \varepsilon_0 w^i(n)$. The formula is as follows:

$$\tilde{IMF}_1(n) = \frac{1}{I} \sum_{i=1}^{I} IMF_1^i(n) = \overline{IMF_1}(n). \tag{1}$$

Step 2: The first residue can be obtained:

$$r_1(n) = x(n) - \tilde{IMF}_1(n). \tag{2}$$

Step 3: Decompose the realizations $r_1(n) + \varepsilon_1 E_1(w^i(n)), i = 1, 2, K, I$ up to their first EMD mode. ε_1 is the standard deviation of white Gaussian noise at the first stage. $\tilde{IMF}_2(n)$ is expressed as follows:

$$\tilde{IMF}_2(n) = \frac{1}{I} \sum_{i=1}^{I} E_1(r_1(n) + \varepsilon_0 E_1(w^i(n))). \tag{3}$$

Step 4: With K being the total number of modes, calculate the kth residue as follows:

$$r_k(n) = r_{k-1}(n) - \tilde{IMF}_k(n), \quad k = 2, 3, K, K. \tag{4}$$

Step 5: Decompose the realizations $r_k(n) + \varepsilon_k E_k(w^i(n))$ with $i = 1, 2, ..., I$, up to their first EMD mode. The $(k + 1)^{th}$ mode can be calculated as follows:

$$\tilde{IMF}_{k+1}(n) = \frac{1}{I} \sum_{i=1}^{I} E_1(r_k(n) + \varepsilon_k E_k(w^i(n))). \tag{5}$$

Step 6: Repeat steps 4 to 6 until the residue becomes a monotonic function, so that it is impossible to further extract the IMF. Assuming that k and $r_k(n)$ represent the total number of modes and the final residue, respectively, the input $x(n)$ is expressed as

$$x(n) = \sum_{k=1}^{K} \tilde{IMF}_k(n) + r_k(n). \tag{6}$$

Eq. (6) completes the decomposition of the original data.

2.2. Phase-space reconstruction

The phase-space reconstruction (PSR) theory is considered to be a useful method

for reconstructing chaotic time series [35].

Definition 1. The given time series $x = \{x_1, x_2, \ldots, x_N\}$ can be extended to a matrix time series as follows:

$$X = \begin{bmatrix} x_1 & x_2 & L & x_{N-(m-1)\tau} \\ x_{1+\tau} & x_{2+\tau} & L & x_{N-(m-2)\tau} \\ M & M & M & M \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & L & x_N \end{bmatrix},$$ (7)

where $M = N - (m-1)\tau$. The parameter m represents the embedding dimension, and τ is the delay time.

In the study, the C-C method [36] is employed to reconstruct the phase space by applying correlation integrals to estimate m and τ simultaneously.

Definition 2. The correlation integral for the embedded time series is as follows:

$$C(m, N, r, t) = \frac{2}{M'(M'-1)} \sum_{1 \leq i \leq j \leq M'} \theta(r - d_{ij}), r > 0,$$ (8)

where $d_{ij} = \|X_i - X_J\|$ represents the sup-norm, τ is the index lag, and $\theta(x) = \begin{cases} 0, x < 0 \\ 1, x \geq 0 \end{cases}$. The correlation integral is a cumulative distribution function, and the Euclidean distance between any two points in the phase space is less than r.

Definition 3. The given sequence x is divided into t disjoint subsequences to form the matrix time series. Let us structure a statistic $S(m, N, r, t)$, which can be calculated via the following formula. The correlation integral $S(m, N, r, t)$ can be expressed as follows:

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^{t} \left[C_s \left(m, \frac{N}{t}, r, t \right) - C_s^m \left(1, \frac{N}{t}, r, t \right) \right].$$ (9)

Select two radii r corresponding to the maximum and minimum values, and define the difference as follows:

$$\Delta S(m, t) = \max \left\{ S(m, r_j, t) \right\} - \min \left\{ S(m, r_j, t) \right\},$$ (10)

where $\Delta S(m, t)$ indicates the maximum deviation of the radius r.

2.3 Firefly algorithm

The FA was proposed by Yang as a heuristic algorithm to imitate firefly luminescence behavior in nature and structure [37]. The FA can be used to solve nonlinear optimization problems under various constraints. The algorithm is a special particle swarm optimization (PSO) algorithm that is easy to understand and implement. Fireflies are described by the following three idealization rules.
1. All fireflies are unisexual and can be attracted to each other.
2. The attraction ability of fireflies depends on brightness, i.e., fireflies move in the direction of bright fireflies. At the same time, attractiveness is inversely proportional to distance.
3. The brightness of fireflies is determined by the fitness function value.

Definition 1. The function of the luminance of the firefly is expressed as follows:

$$I = I_0 e^{-\gamma r_{ij}^2},$$ (11)

where $I(r)$ is the brightness of the firefly; r the distance between two fireflies i

and j; I_0 is the nature brightness; and γ is the attractive coefficient of brightness.

Definition 2. The attractiveness of the firefly is as follows:

$$\beta = \beta_0 e^{-\gamma r_{ij}^2}. \tag{12}$$

The distance between any two fireflies i and j is expressed by the Descartes function. The formula is as follows:

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{d}(x_{i,k} - x_{j,k})^2}. \tag{13}$$

Hence, $x_{i,k}$ defines the kth component of the spatial coordinate x_i of the ith firefly.

Definition 3. The trajectory of firefly i attracted by firefly j is expressed as follows:

$$x_i = x_i + \beta \times e^{-\gamma r_{ij}^2} \times (x_j - x_i) + \alpha \times (rand - 1/2), \tag{14}$$

where α is a random parameter.

Algorithm: **Firefly Algorithm**

Parameters:

Gen_{Max}—the maximum iterations	n—the number of fireflies
n—the number of fireflies	f_i—the fitness function of firefly i
x_i—nest i	g—current iteration number
I_i—light intensity of firefly I	d—the number of dimension

/*Set the parameters of FA. */

/*Initialize population of n fireflies x_i (i = 1, 2,..., n) randomly.*/

FOR EACH i: $1 \le i \le P$ **DO**

| Evaluate the corresponding fitness function f_i

END FOR

/*Determine light intensity. */

FOR EACH i: $1 \le i \le P$ **DO**

| Determine light intensity I_i using f_i.

END FOR

WHILE ($g<$ Gen_{Max}) **DO**

    FOR EACH i=1: P **DO**

        FOR EACH j=1: P **DO**

          /*Move firefly i towards j in all d dimensions. */

          IF ($I_j>I_i$) **THEN**

$$x_i = x_i + \beta_0 e^{-\gamma r^2}\left(x_j - x_i\right) + \alpha\left(rand - \frac{1}{2}\right)$$

$$r_{ij} = \left|x_i - x_j\right| = \sqrt{\sum_{k=1}^{d}\left(x_{i,k} - x_{j,k}\right)^2}$$

          END IF

          Attractiveness varies with the distance r via $\beta_0 e^{-\gamma r_{ij}^2}$

          /*Evaluate the new solution and update the new light intensity I_i.*/

          /*Evaluate the new solution and update the new light intensity I_i.*/

        END FOR

    END FOR

END WHILE/*Post process results and visualization*/

2.4 Holt's ES method

The ES model is widely used in business, environmental science, and other fields because of its simple structure and effectiveness. Holt's ES model was proposed in 1957 [38].

Unlike the general ES model, Holt's ES model involves direct smoothing of the trend data and a prediction of the original time series. It assumes that all the known data influence the prediction value. The short-term data have great influence on the prediction value, and the long-term data have little influence on the prediction value.

Definition 1. The specific formulas are as follows:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{15}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}, \tag{16}$$

where $0 < \alpha$ and $\gamma < 1$ are the smoothing parameters. In this model, α determines the average length of the estimate level, and γ determines the smoothing trend.

Definition 2. A straightforward approach for finding the optimal values of both α and β constrained to the range (0, 1) is to search for the parameter combination that minimizes the sum of squared errors of predictions:

$$error(\alpha, \gamma) = \sum_{i=1}^{n}(X_i^{'} - X_i)^2, \tag{17}$$

where $X_{t+1} = S_2 + T_2$. The starting values for S_2 and T_2 are typically taken as X_2 and 0, respectively.

2.5 Support vector regression

Vapnik [39] proposed SVR, which is derived from SVMs. The advantage of SVR is its ability to model nonlinear relations. SVR is based on the principle of structural risk minimization, and the regression task of support vector classification is completed by introducing the ε insensitive loss function. The following is a brief description of SVR.

Definition 1. The decision function of support vector regression is

$$f(x) = \langle w, \varphi(x) \rangle + b. \tag{18}$$

Here, $\varphi(x)$ represents the nonlinear mapping of the input space to the high-dimensional space, w represents the weight, and b is the offset.

Definition 2. To estimate w and b, the following formula is minimized:

$$R(f) = \frac{1}{2}\|w\|^2 + C\frac{1}{n}\sum_{i=1}^{n}L_\varepsilon\left(y_i, f(x_i)\right), \tag{19}$$

where L_ε is a loss function, and C is a penalty factor. A greater C value indicates a greater penalty for data that exceed L_ε. The ε insensitive loss function is used as the problem of structural minimum risk estimation.

Definition 3. The Lagrange equation is introduced, the partial derivative of the relevant parameters is obtained, and the regression estimation function of the SVM is obtained:

$$f(x) = \sum_{i=1}^{N}(\overline{\alpha_i} - \overline{\alpha_i^*})K(x_i, x) + b, \tag{20}$$

where $\overline{\alpha_i}$, $\overline{\alpha_i^*} \geq 0(i = 1, 2, \cdots, N)$ is a Lagrange multiplier, x is a kernel function,

and N is the number of input parameters. Different types of kernel functions can be selected to obtain different nonlinear decision learning machines. Because the radial basis function has good adaptability and a good convergence area in low- and high-dimensional spaces, the Gauss radial basis function is used in this study.

3. Framework of proposed model

The proposed novel model framework aims to improve the forecasting accuracy and the utilization of wind energy. The model considers the chaotic characteristics in the time series. First, the original data are preprocessed, and then the data input form that can provide more real information is determined using PSR technology. Finally, the "linear and nonlinear" framework is constructed for prediction based on the optimization algorithm.

Owing to the irregularity and randomness in the original observation data before model forecasting, the Lyapunov index is employed to quantitatively describe the chaotic phenomenon of wind-speed series, which is an important quantitative measure of system dynamics. To eliminate noise to the greatest extent possible, a noise-assisted data analysis method with CEEMDAN is developed to remove noise and extract valuable information in the first phase of forecasting. After CEEMDAN, a smooth series is obtained by decomposing the original wind speed and removing the high-frequency *IMF*. Furthermore, the rest of the *IMF*s are refactored to obtain preprocessed data. The process is shown in Fig. 1. PSR technology is used to analyze the collected datasets, and the sequence can be regarded as a chaotic system. Therefore, the C-C method determines the optimal input form of the predicted structure of the design, which is a technique for reconstructing the phase space. Through the obtained optimal embedding dimension (m) and optimal delay time (τ), the sequence features can be described more clearly. A schematic of the input structure of PSR is shown in Fig. 1.

The linear and nonlinear forecasting models are combined to obtain an outstanding tool for capturing information hidden in the data and improving the predictive performance [40]. It is advisable to regard the time series $X_t (t = 1,\ldots,n)$ as consisting of two parts: a linear component and a nonlinear component. That is,

$$X_t = L_t + N_t, \tag{21}$$

where L_t denotes the linear components, and N_t denotes the nonlinear components.

Considering the advantages of Holt's ES method for capturing the linear and nonlinear factors of the SVR model, the proposed hybrid model employs the firefly optimization algorithm. The steps of forecasting in the proposed CEE-CC-FHS method are shown in Fig. 1 and described as follows.

Step 1: The linear component of forecasting L_t is determined using Holt's method. The residuals are as follows:

$$e_t = X_t - \hat{L}_t. \tag{22}$$

The residuals e_t contain information, excluding the linear factor. $\hat{L}_t$ represents the prediction results of Holt's method.

Step 2: SVR is used to determine the nonlinear component e_t.

$$e_t = f(e_{t-1}, \text{K}, e_{t-d}) + \varepsilon_t \tag{23}$$

Here, $f(\square)$ represents the nonlinear process, and ε_t represents a stochastic error. We input $e_{t-1},\ldots,e_{t-d}$ into the SVR model to obtain the output result e_t, as described in Section 2.5. The residual predicted by the SVR model contains the nonlinear component.

8

Step 3: Combining prediction results of linear and nonlinear models: $\hat{e}_t$ represents the predicted residuals. Therefore, the results predicted by the proposed CEE-CC-FHS model are given by

$$\hat{X}_t = \hat{L}_t + \hat{e}_t. \tag{24}$$
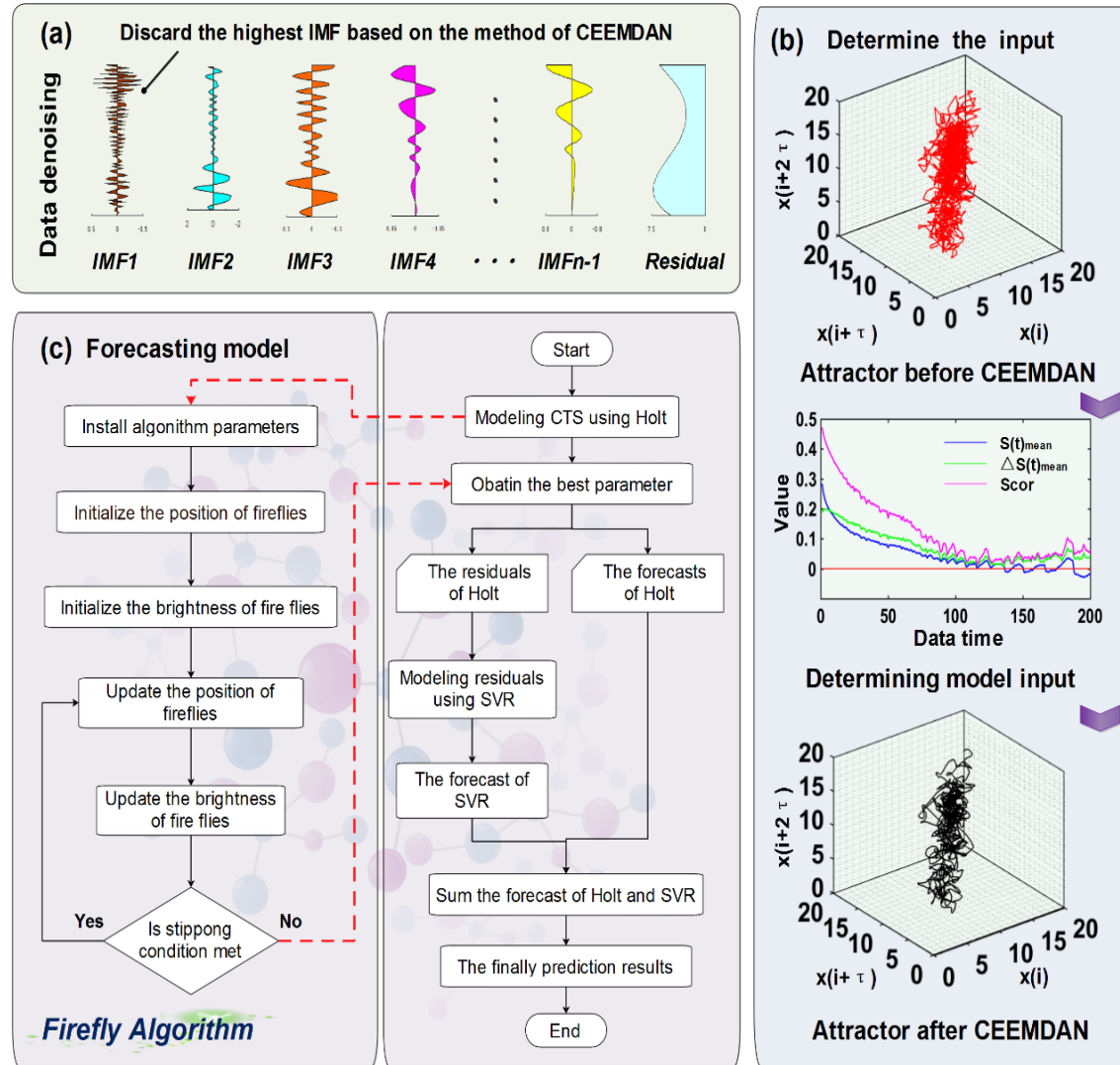
Fig. 1. Flow chart of the proposed combined model.

4. Experimentation and analysis

The experimental analysis is detailed in this section. In Section 4.1, a basic statistical description of the datasets is provided. The statistical criterion of the prediction level is described in Section 4.2. The implementations of numerical modeling for the prediction of wind speed are described in Sections 4.3–4.5. All the numerical simulations were implemented using MATLAB R2017a software.

4.1 Data description

Experimental datasets collected from wind farms in China are used as samples to evaluate the effectiveness of the proposed model. The trends of the wind-speed series for three locations are illustrated in Fig. 2. The descriptive statistics (mean, minimum, maximum, standard deviation, MLYE, and complexity) for the three sites are presented in Table 1. The values of the MLYE, which describe the chaotic characteristics of the time series, are all positive in the experimental datasets.

Table 1
Descriptive statistical of the data set.

| Area	Time	Data	Number	Mean	Min	Max	Std.	MLYE	Complexity
Site 1	10-min	All samples	1500	9.90	2.00	18.70	3.18	0.0122	0.3889
		Training	1300	10.48	2.00	18.70	2.98	0.0223	0.4669
		Testing	200	6.15	3.70	10.40	1.35	0.8029	0.7347
	30-min	All samples	1500	8.78	1.60	17.90	3.20	0.1255	0.5275
		Training	1300	9.04	2.00	17.90	3.25	0.3067	0.5236
		Testing	200	7.09	1.60	13.60	2.18	0.4626	0.8644
Site 2	10-min	All samples	1500	9.15	2.00	18.40	2.94	0.0113	0.4043
		Training	1300	9.68	2.00	18.40	2.77	0.0395	0.4538
		Testing	200	5.67	3.80	8.80	1.10	0.3540	0.7131
	30-min	All samples	1500	7.90	1.50	17.50	3.02	0.3472	0.5082
		Training	1300	8.17	1.60	17.50	3.06	0.1988	0.4887
		Testing	200	6.16	1.50	11.70	2.01	0.6304	0.7996
Site 3	10-min	All samples	1500	7.40	2.10	13.70	2.03	0.0250	0.5929
		Training	1300	7.66	2.10	13.70	1.98	0.0743	0.6588
		Testing	200	5.74	2.30	9.50	1.53	0.1839	0.6483
	30-min	All samples	1500	6.84	0.90	13.90	2.18	0.0141	0.6468
		Training	1300	6.99	2.00	13.90	2.19	1.0104	0.6327
		Testing	200	5.83	0.90	11.30	1.81	0.0539	0.8644
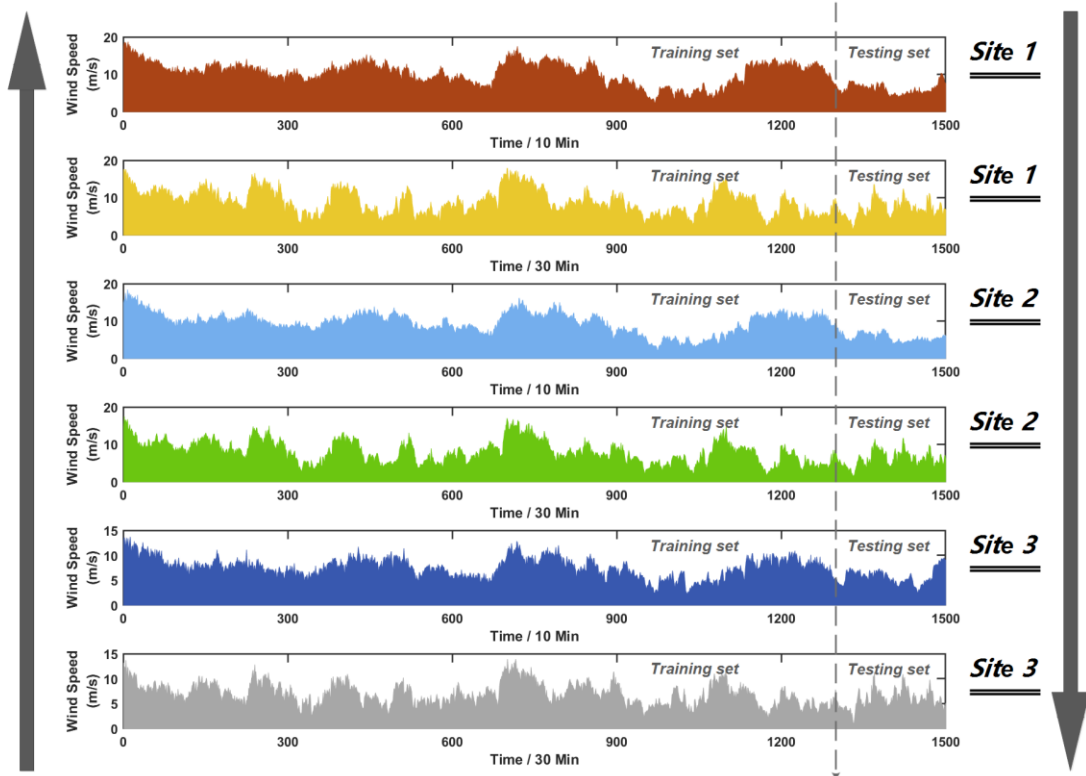
Fig. 2. Test Site: wind speed data.

Table 1 indicates basic characteristics of the experimental datasets. However, it is insufficient to survey the wind-speed patterns in Table 1. Therefore, the frequency distribution of the wind-speed series and the corresponding fit for the distributions (Weibull, gamma, normal) are shown in Fig. 3. In all cases, the change of the data collected from different sites or wind-speed patterns at the same location can be displayed by shape parameters and scale parameters, as well as normal distribution parameters. The results of Table 1 and the distribution parameters of Fig. 3 indicate that the differences between different wind-speed time series are significant. Therefore, it

can be concluded that the three test points of the comparison are meaningful, not invalid.
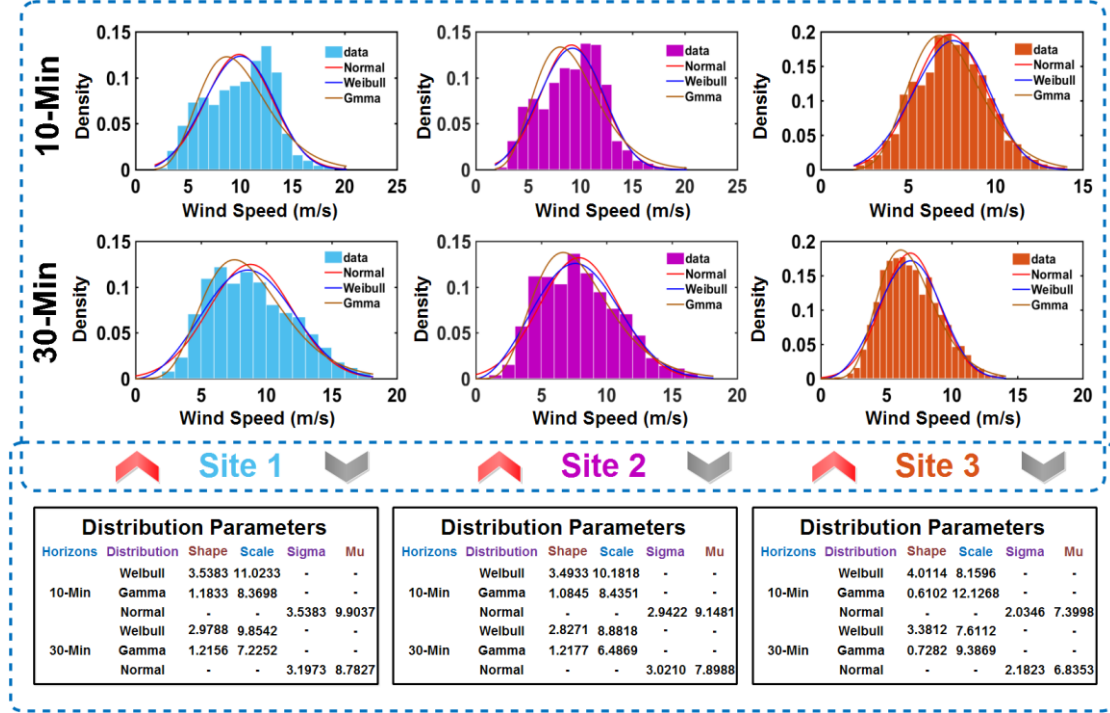

Fig. 3. Distribution fitting of wind speed series.

4.2 The evaluation criteria

The evaluation criterion is used to intuitively measure the accuracy of model forecasting. Although a multitude of evaluation criteria have been applied, no single evaluation criterion has proven to be effective for all situations. Therefore, in this study, three common error indices—the mean absolute error (MAE), mean square error (MSE), and mean absolute percentage error (MAPE)—are applied to evaluate the forecasting performance of the proposed model. The evaluation criteria are given as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i| \tag{25}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \tag{26}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% , \tag{27}$$

where N represents the length of the forecasting series; $\hat{y}_i (i = 1, 2, \ldots, N)$ represents the forecast values, and $y_i (i = 1, 2, \ldots, N)$ represents the observed wind speed during the same period.

The number of predictions within a certain error range is analyzed for comprehensive comparison. Additionally, the maximum permissible error (MPE) is introduced:

$$MPE = \frac{y_i - \hat{y}_i}{y_i} \times 100\% \tag{28}$$

4.3 Data feature processing

As described in this subsection, the disturbance factors of the wind-speed series are removed via CEEMDAN, and then the optimal input of the model is determined by reconstructing the phase space. The process is as follows.

Step 1: Data denoising. The noise component in the wind-speed series greatly affects the model prediction performance. The wind-speed series is treated as a signal and decomposed into ten *IMF*s with different frequency ranges. The original wind-speed time series is decomposed via CEEMDAN into several *IMF*s, such as IMF_1, …, IMF_{n-1} and Res_n. By eliminating the MIF_1 to smooth the original time series, the prediction efficiency and accuracy can be improved.

Step 2: Data expansion. The different input forms have a significant effect on the predictive performance of the model. Therefore, to fully reveal the information in the time series, the time series is expanded to a high-dimensional space via PSR technology. Thus, the optimal embedding dimension and time window can be determined through the C-C method. PSR is used to transform the chaotic wind-speed sequence into a matrix time series, and a wind-speed dynamic model with embedding dimension m, time window ϖ, and time delay τ is constructed. The maximum lag is set as 200 in the C-C method $S(t)_{mean}$, and $\Delta S(t)_{mean}$ and $Scor$ are calculated to determine the embedding dimension m and the time delay τ.

Although there are many observable factors that influence changes in time series, the true regularity of the system can be extracted and recovered from the time series. A trajectory generated by a chaotic system will eventually undergo regular movement after a certain period of change, resulting in a regular, tangible trajectory (chaotic attractor). Therefore, the chaotic time series can reveal the evolution of hidden chaotic attractors through PSR. Furthermore, after the noise is removed via CEEMDAN, the attractor is sufficiently recovered to reveal the true trajectory of the sequence motion, as shown in Fig. 4.
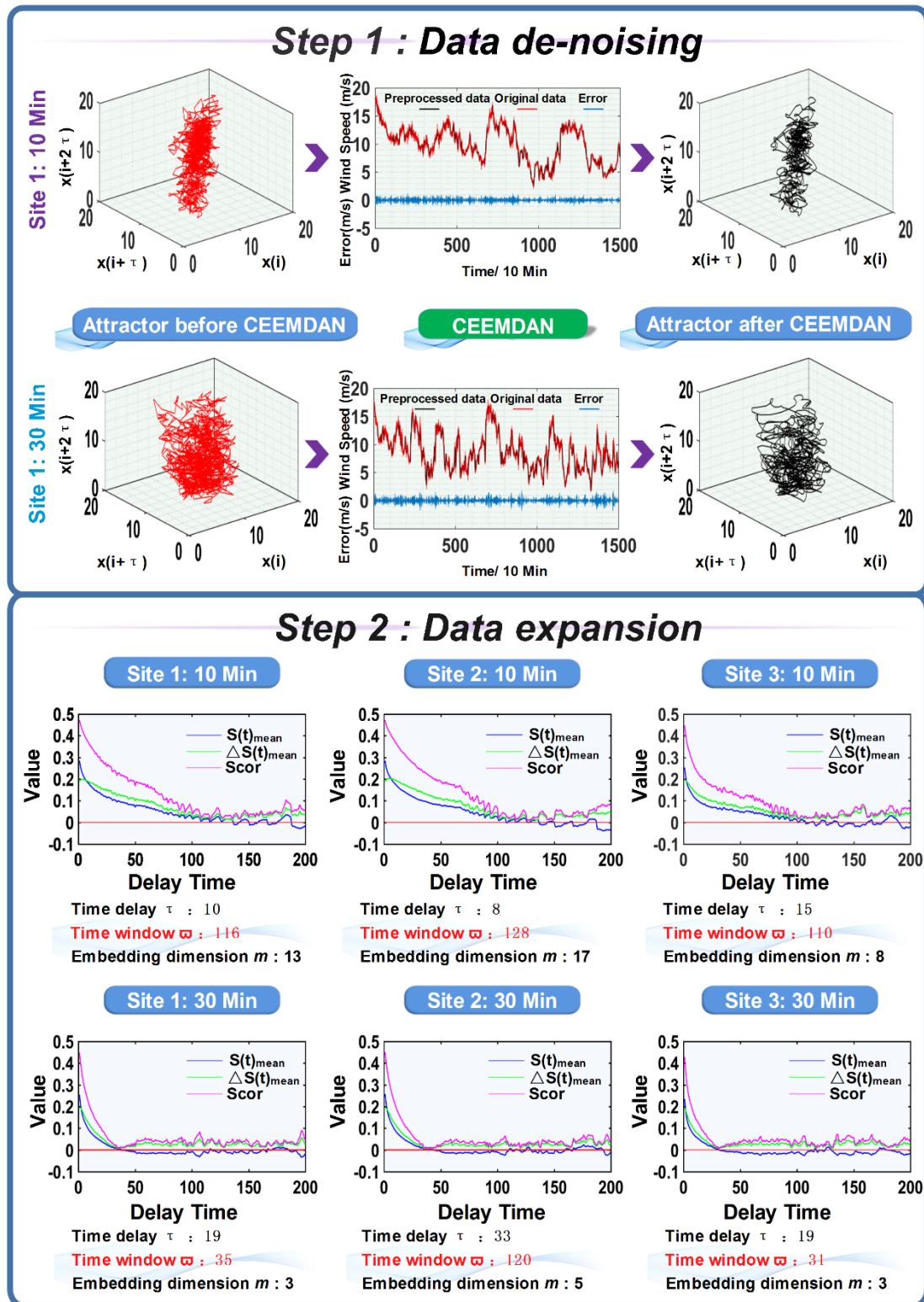
Fig. 4. Processing results of data features

4.4 Experiment I

The prediction module of the designed framework is predicted by combining linear and nonlinear models. To effectively verify the performance of the prediction module, four individual methods—ARIMA, Holt, the general regression neural network (GRNN), and SVR—and two hybrid methods—ARIMA-SVR and Holt-GRNN—are used as benchmark models for comparison. The aforementioned models are all

13

processed using the same data preprocessing method before forecasting, which not only uses CEEMDAN denoising but also obtains the optimal input mode through the C-C method. Further comparisons are made, as follows.

a) From Table 2, the statistical error measures for six comparison models and the proposed models are reported. The MAPE values of the proposed model are smaller than those of the comparative models. For site 1, the MAPE of the CEE-CC-Holt model is 6.8999%, while the MAPE of the CEE-CC-ARIMA model is 7.4737%, indicating a 0.5738% improvement, as shown in Table 2. For the hybrid CEE-CC-ARIMA-SVR and CEE-CC-FHG models, the MAPE values are 4.0905% and 3.5301%, respectively, and produce increases of 1.7675% and 1.2071% compared with the proposed model.

b) For the MPE, a positive value indicates that the predicted value is lower than the actual value, and a negative value indicates that the predicted value is higher than the actual value. The results for the MPE within the error ranges of $\pm2.5\%$, $\pm5\%$, $\pm7.5\%$, and $\pm10\%$ for the three study sites are presented in Table 2. For site 1, 75.76% and 91.41% of the errors are within $\pm5\%$ between the measured values and the values predicted by the CEE-CC-FHG and the proposed models, respectively. The proposed model has more MPE values in the range of $\pm10\%$ than those of the compared models. The MPE values of CEE-CC-ARIMA, CEE-CC-Holt, CEE-CC-GRNN, CEE-CC-SVR, CEE-CC-ARIMA-SVR, and CEE-CC-FHG are 70.00%, 76.38%, 87.94%, 92.96%, 95.96%, and 98.48%, respectively.

c) Fig. 5 shows the error (MPE) and the residual distribution for a 10-min wind speed time series. For site 1, the MPE values of all seven models approximate to a normal distribution. However, the MPE distributions for the proposed model are more concentrated than those for the other models. For site 2, the MPE distribution of the proposed model more closely approximates the normal distribution compared with the other models.

d) The residual distribution in Fig. 5 shows that the points of the proposed model and CEE-CC-FHG are concentrated near the origin. Thus, a higher accuracy of fitting with the model indicates a higher accuracy of fitting with other models.

Remark. According to Table 2 and Fig. 5, the combination of linear and nonlinear models based on firefly optimization can improve the prediction accuracy. That is, the proposed model can accurately and stably forecast changes in future wind speeds.
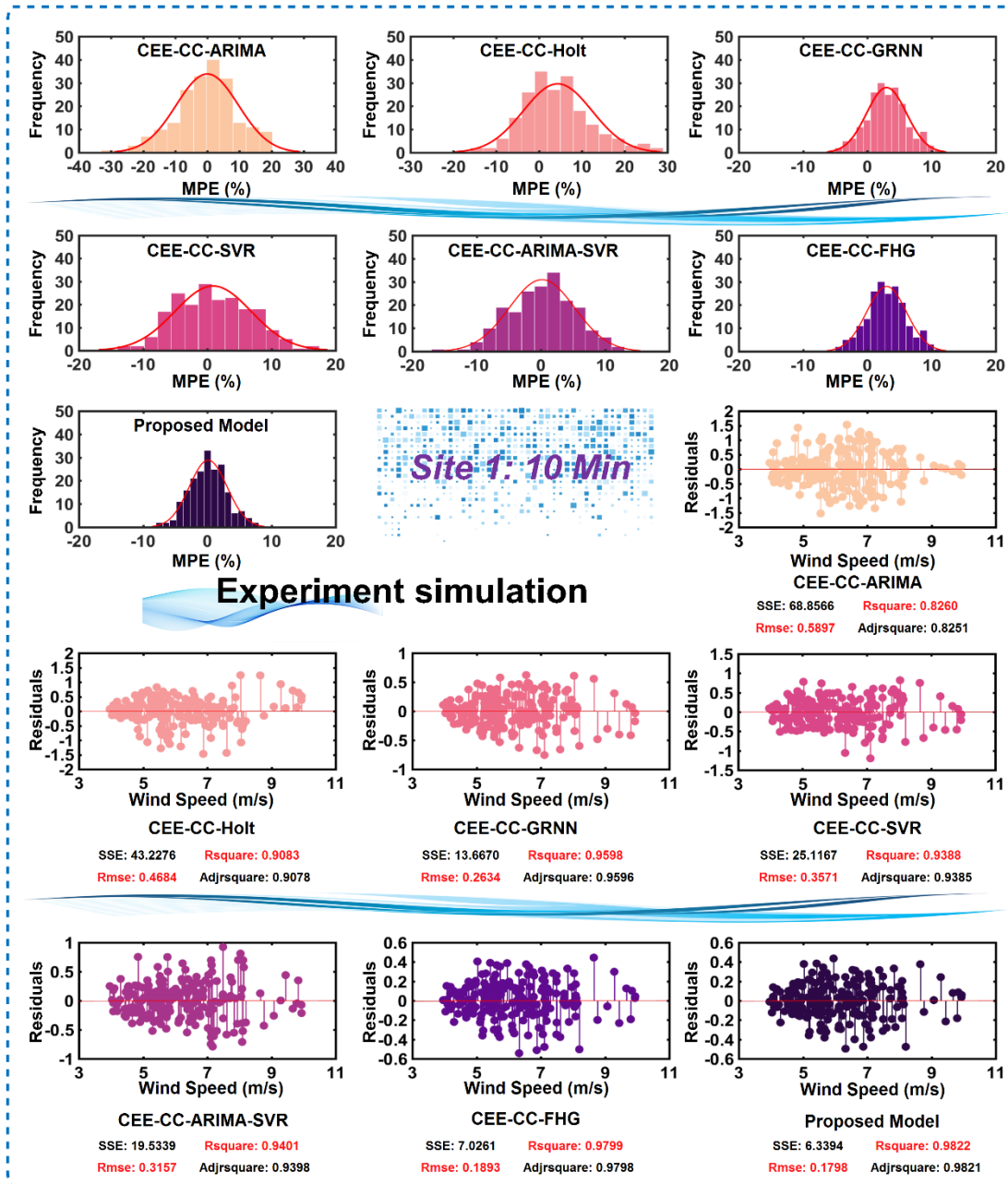
Fig. 5. The distribution of MPE and residual for site

Table 2
Statistical error measures of the seven models for 10-Min wind speed.

Area	Model	MAE (m/s)	MSE (m/s)	MAPE (%)	MPE (% of errors in each margin)			
					±2.5%	±5%	±7.5%	±10%
Site1	CEE-CC-ARIMA	0.4544	0.3457	7.4737	26.00	42.00	61.50	70.00
	CEE-CC-Holt	0.4104	0.2929	6.8999	26.13	47.24	66.83	76.38
	CEE-CC-GRNN	0.3261	0.1480	5.5156	24.62	49.25	70.85	87.94
	CEE-CC-SVR	0.2955	0.1332	4.8748	27.64	56.28	78.89	92.96
	CEE-CC-ARIMA-SVR	0.2540	0.1051	4.0905	38.89	64.14	86.87	95.96
	CEE-CC-FHG	0.2113	0.0671	3.5301	40.40	75.76	91.92	98.48
	Proposed Model	**0.1413**	**0.0321**	**2.3230**	**62.12**	**91.41**	**99.49**	**100.00**
Site2	CEE-CC-ARIMA	0.6937	0.7903	12.3568	14.50	25.50	38.00	47.50
	CEE-CC-Holt	0.3906	0.2417	7.3074	22.11	39.70	63.82	76.38
	CEE-CC-GRNN	0.2526	0.0932	4.4897	31.66	61.81	83.92	93.97
	CEE-CC-SVR	0.1717	0.0483	3.0571	51.26	79.90	94.47	98.99
	CEE-CC-ARIMA-SVR	0.1490	0.0353	2.7020	54.55	85.86	97.47	99.49
	CEE-CC-FHG	0.1619	0.0398	2.9644	45.96	82.32	96.46	99.49
	Proposed Model	**0.1207**	**0.0243**	**2.1587**	**65.66**	**91.41**	**98.48**	**100.00**
Site3	CEE-CC-ARIMA	0.5948	0.6106	10.7482	16.50	35.00	46.50	55.00
	CEE-CC-Holt	0.3777	0.2343	7.1790	25.63	46.73	64.32	77.89
	CEE-CC-GRNN	0.2094	0.0709	4.0289	43.72	68.84	83.92	92.46
	CEE-CC-SVR	0.188	0.0614	3.5643	46.23	76.88	87.94	93.97
	CEE-CC-ARIMA-SVR	0.1598	0.0401	2.9006	51.01	85.86	94.44	100.00
	CEE-CC-FHG	0.1653	0.0398	3.0749	47.98	81.82	95.45	98.99
	Proposed Model	**0.1252**	**0.0249**	**2.2998**	**62.63**	**91.41**	**97.98**	**100.00**

4.5 Experiment II

In this subsection, to further illustrate the generalization ability of the proposed model, different data preprocessing techniques and optimization methods are compared. The importance of data preprocessing, the optimal input structure of the model, and the parameter optimization for prediction accuracy were verified by Experiment II. Five comparison models were used to validate the superiority of model structure. PSO (CEE-CC-PHS) was used to optimize the model parameters, compared with the optimization of the FA. The forecasting results for different models are presented in Table 3 and Fig. 6. The detailed comparison results are as follows.

a) The proposed model is closer to the observed value than all the comparison models, including WT/SSA/EMD CC-FHS, CEE-FHS, and CEE-CC-PHR, for a 10-min wind speed. For site 1, the developed model has higher prediction accuracy than the other compared models (WT/SSA/EMD CC-FHR, CEE-CC-PHS, and CEE-FHS); the MAPE values of the proposed model exhibit improvements of 4.4823%, 4.0087%, 1.5747%, 0.7818%, and 0.6384%, respectively.

b) In Experiment II, the noise-filtering methods, such as WT, SSA, and EMD, were compared. CEEMDAN significantly outperformed the other noise-filtering methods, as it had higher prediction accuracy. For site 2, compared with the models of WT, SSA, and EMD, the MAPE of the proposed model increased by 3.88%, 3.51%, and 1.26%, respectively. Simultaneously, the MPE values in the range of (+2.5%) are more than those of the compared models.

c) By validating the effect of optimizing the input of PSR, CEE-CC-FHS achieved a lower MAPE value than the other models, indicating that the optimal input plays an indispensable role in forecasting chaotic time series. In the experiments for a 10-min wind speed at the three sites, compared with the CEE-FHS model, the MAPE values of the proposed model decreased by 0.7818%, 0.8456%, and 0.9317%.

d) Comparison with the PSO algorithm reveals that the FA can better find the optimal model parameters to improve the accuracy of the prediction. The FA not only yields excellent performance for forecasting the wind speed but also outperforms the benchmark in the proposed model. For the three sites, the MPE of the proposed model exhibit in the range of (+5%) improvements of 8.08%, 9.09%, and 10.61% compared with the CEE-CC-PHS model, within ±5% of the measured values.

e) Fig. 6 shows a comparison of the actual values and the values forecasted in the experiment. The proposed prediction model is more accurate than the reference model because more targets converge near the diagonal.

Remark: According to the foregoing analysis, the proposed model shows better performance than the other benchmark models. Consequently, the design of data denoising, the optimal input structure, and the optimization parameters have greatly improved in model prediction.
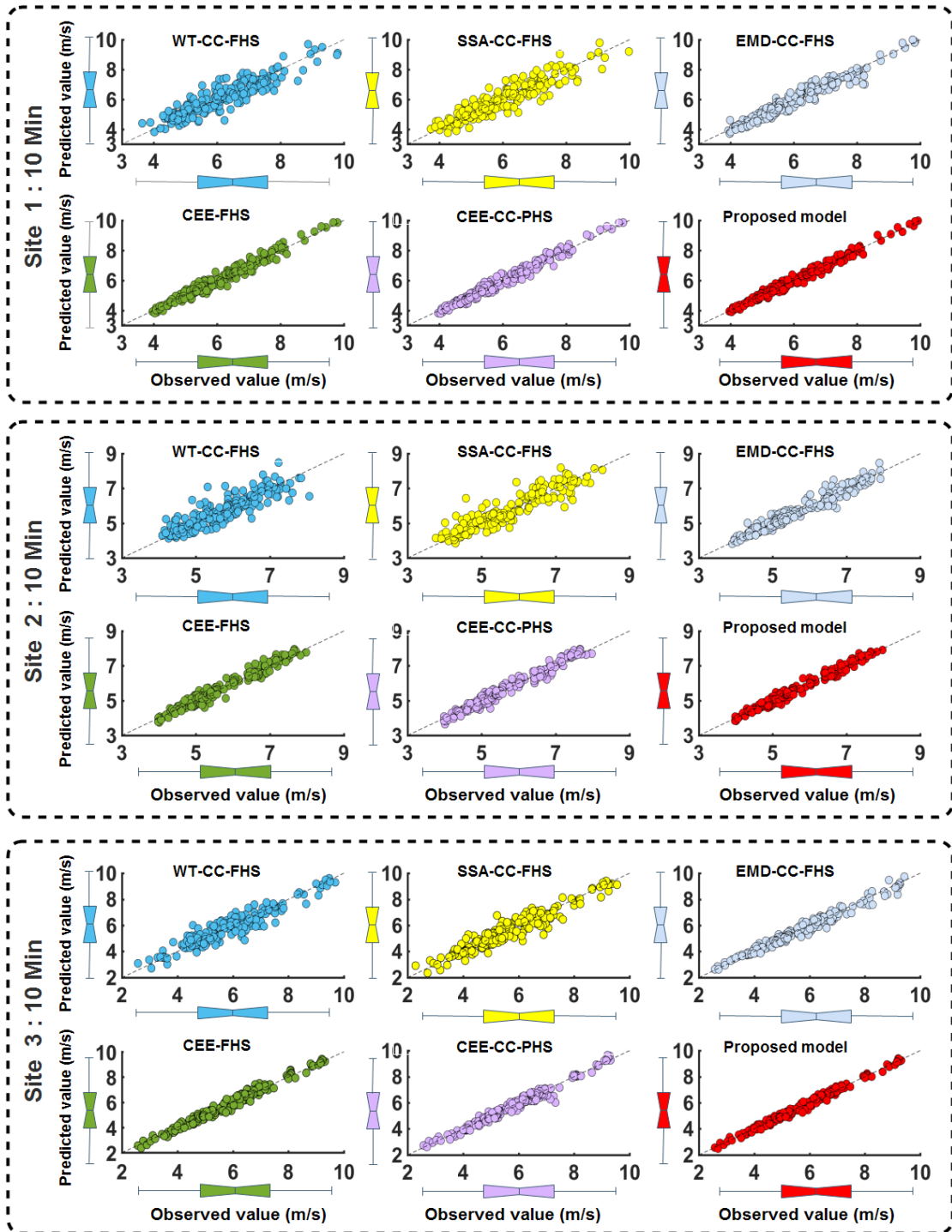
Fig. 6. Comparison of he observed values and the forecasting results for experimental models

Table 3
Statistical error measures of the six models for 10-Min wind speed at three sites.

Area	Model	MAE (m/s)	MSE (m/s)	MAPE (%)	MPE (% of errors in each margin) ±2.5%	±5%	±7.5%	±10%
Site1	WT-CC-FHS	0.4121	0.2711	6.8053	20.60	44.22	66.33	76.88
	SSA-CC-FHS	0.3862	0.2472	6.3317	27.78	46.46	64.65	80.81
	EMD-CC-FHS	0.2348	0.0940	3.8977	41.21	68.84	87.44	96.48
	CEE-FHS	0.1894	0.0572	3.1048	49.49	77.78	93.94	98.99
	CEE-CC-PHS	0.1796	0.0494	2.9614	50.51	83.33	94.95	100.00
	Proposed Model	**0.1413**	**0.0321**	**2.3230**	**62.12**	**91.41**	**99.49**	**100.00**
Site2	WT-CC-FHS	0.3422	0.2162	6.0363	30.65	54.77	70.85	81.41
	SSA-CC-FHS	0.3158	0.1828	5.6639	30.30	56.06	75.76	83.33
	EMD-CC-FHS	0.1930	0.0644	3.4196	44.22	75.88	91.96	96.98
	CEE-FHS	0.1678	0.0445	3.0043	50.51	84.85	94.44	98.99
	CEE-CC-PHS	0.1596	0.0404	2.8769	53.03	82.32	96.97	99.49
	Proposed Model	**0.1207**	**0.0243**	**2.1587**	**65.66**	**91.41**	**98.48**	**100.00**
Site3	WT-CC-FHS	0.3627	0.2220	6.5462	28.64	50.75	62.81	77.39
	SSA-CC-FHS	0.3484	0.2122	6.4013	30.81	48.99	65.66	79.29
	EMD-CC-FHS	0.1869	0.0621	3.2421	46.73	77.39	91.96	97.99
	CEE-FHS	0.1746	0.0505	3.2315	48.48	74.75	91.41	98.48
	CEE-CC-PHS	0.1716	0.0542	3.1316	51.01	80.81	91.92	97.98
	Proposed Model	**0.1252**	**0.0249**	**2.2998**	**62.63**	**91.41**	**97.98**	**100.00**

4.6 Experiment III

To verify the applicability of the proposed model more comprehensively, a 30-min wind-speed time series for the same three sites was used. The results of the experimental simulations are shown in Tables 4 and 5, from which similar conclusions can be drawn to Experiments I and II. Fig. 7 shows the distribution of the MPE and residual for site 1, and a comparison of the actual and the forecasted values is shown in Fig. 8. We draw the following conclusions.

a) From Table 4, compared with the six models, the proposed model has better forecasting performance as the time interval increases. It can be concluded that that the prediction results of the model are more stable. For site 1, the MPE value of the proposed model is more than that of the comparative model in the range of (+2.5%). The MPE values of CEE-CC-ARIMA, CEE-CC-Holt, CEE-CC-GRNN, CEE-CC-SVR, CEE-CC-ARIMA-SVR, and CEE-CC-Holt-GRNN are 13.00%, 19.10%, 30.65%, 31.16%, 31.31%, and 38.38 %, respectively.

b) From Fig. 7, the MPE distribution of the proposed model is more concentrated than those of the other models, and more points are concentrated around the zero point, indicating that more predicted values are close to the actual value. The points in the residual of the proposed model and CEE-CC-FHG are more concentrated in the horizontal zone centered on the origin, indicating that the prediction accuracy is higher than that of the other models.

c) On one hand, CEEMDAN and the C-C method can significantly improve the prediction accuracy of the proposed model compared with the benchmark models. On the other hand, the firefly optimization algorithm can improve the forecasting performance, as shown in Table 5 and Fig. 8. For site 1, the proposed model has higher prediction performance than the compared models (WT/SSA/EMD CC-FHS, CEE-CC-PHS, and CEE-CC-FHS); the MAPE values of the proposed model exhibit improvements of 2.4045%, 7.2375%, 1.4267%, 0.9519%, and 0.7906%, respectively.

d) From Fig. 8, the WT/SSA/EMD CC-FHC points are more scattered around the diagonal compared with the proposed model. The results indicate that the wind-speed time series processed via the CEEMDAN decomposition method combined with PSR can effectively improve the prediction accuracy.

Remark: The test results show that the increase in the wind-speed prediction interval is accompanied by a decrease in accuracy. However, in experimental simulation, the prediction results of the proposed model are better than those of the benchmark model. Therefore, it can be concluded that the proposed model has strong stability and accuracy.
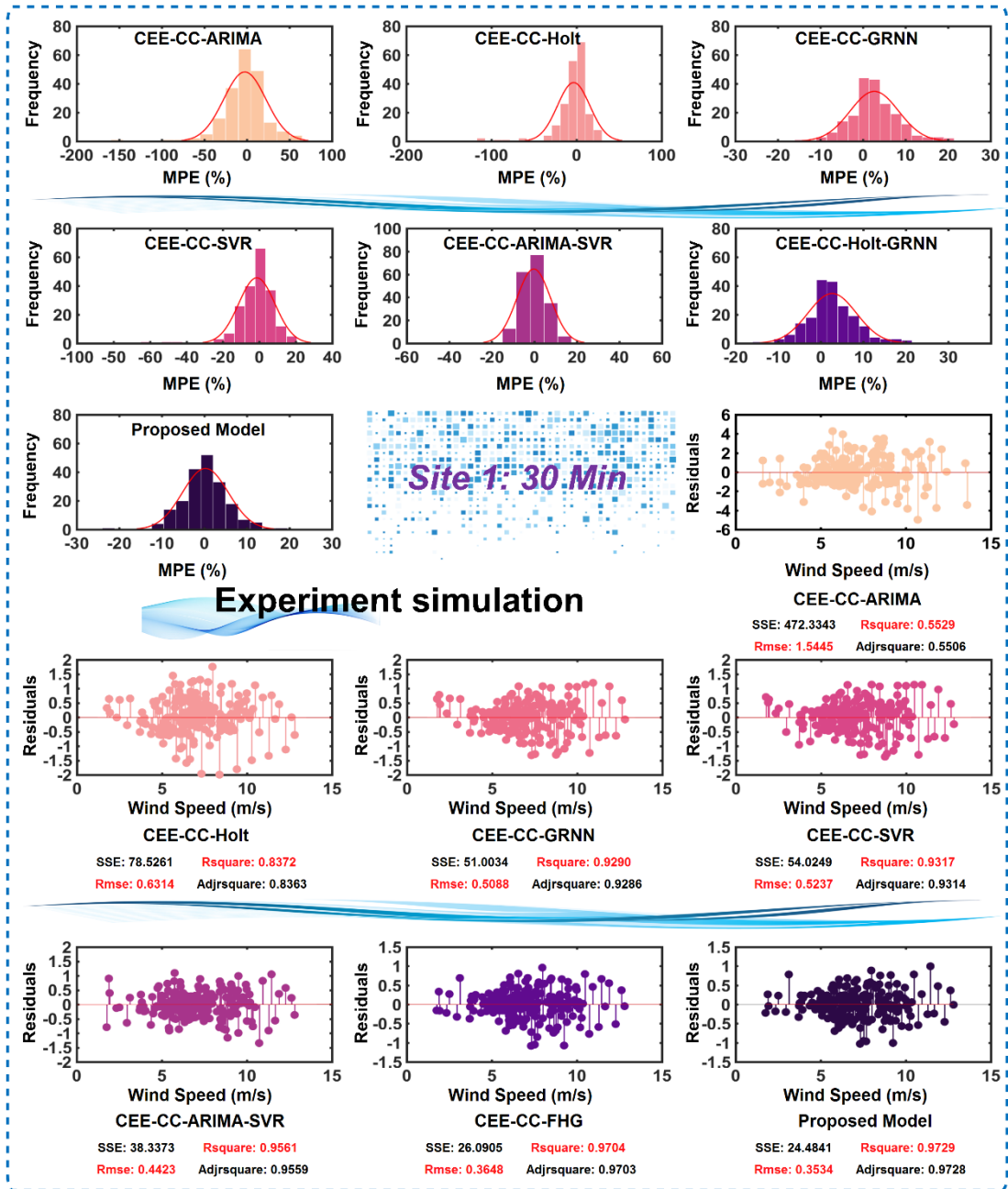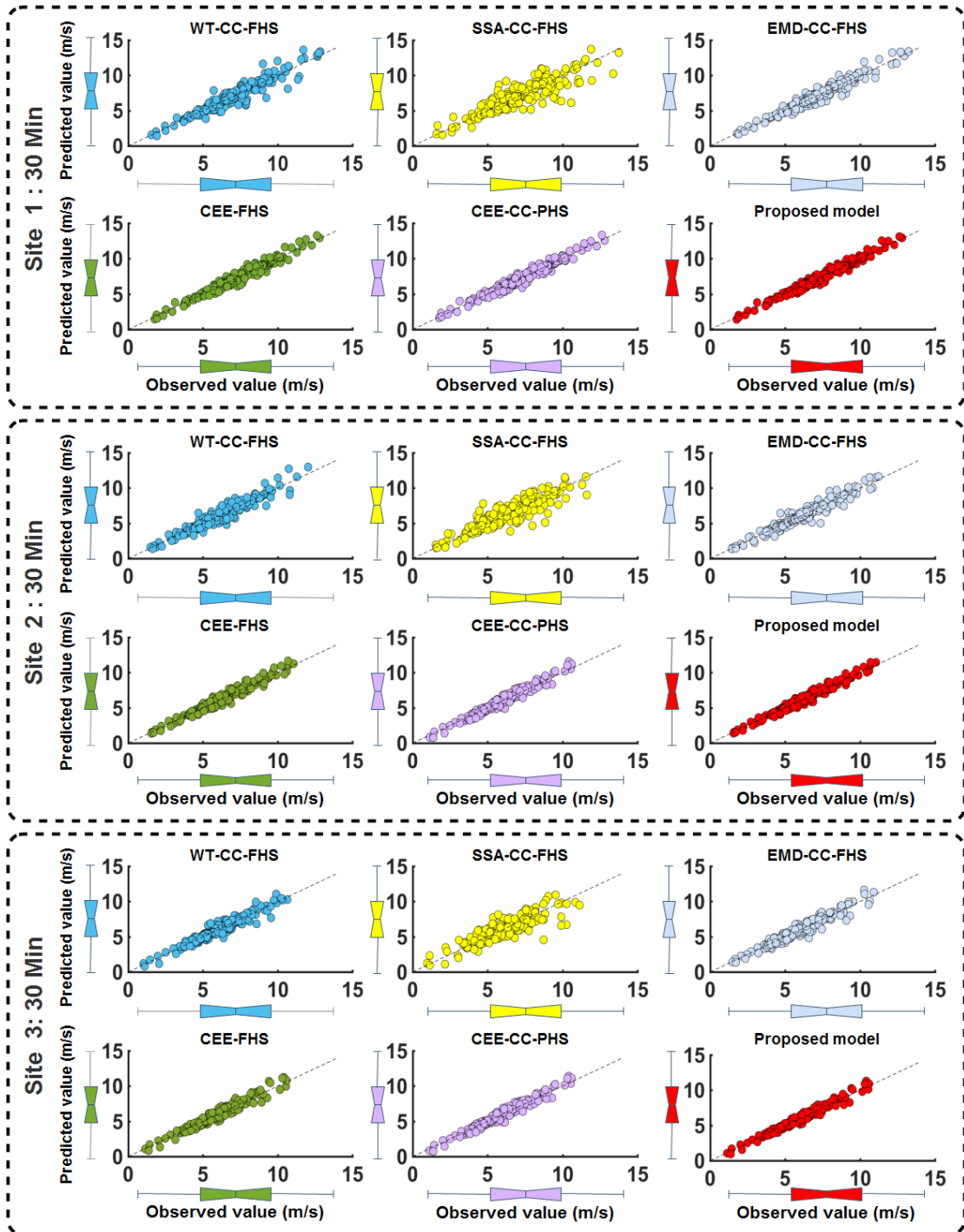
Fig. 7. The distribution of MPE and residual for site 1

Fig. 8. Comparison of he observed values and the forecasting results for experimental models

Table 4
Statistical error measures of the seven models for 30-Min wind speed.

Area	Model	MAE (m/s)	MSE (m/s)	MAPE (%)	MPE (% of errors in each margin)			
					±2.5%	±5%	±7.5%	±10%
	CEE-CC-ARIMA	1.1807	2.5805	17.3199	13.00	23.50	31.50	40.00
	CEE-CC-Holt	0.6757	0.7942	11.3996	19.10	34.67	52.26	63.82
	CEECC-GRNN	0.4223	0.3071	7.0443	30.65	48.74	65.83	80.90
Site1	CEE-CC-SVR	0.4123	0.2893	6.6013	31.16	50.25	64.32	81.41
	CEE-CC-ARIMA-SVR	0.3461	0.1938	5.3930	31.31	57.58	77.27	89.39
	CEE-CC-FHG	0.3086	0.1593	4.7145	38.38	63.13	79.80	90.40
	Proposed Model	**0.2708**	**0.1259**	**4.0564**	**40.40**	**69.70**	**85.35**	**93.43**
	CEE-CC-ARIMA	1.0554	2.0556	17.8843	13.00	20.50	29.00	39.50
	CEE-CC-Holt	0.6299	0.6104	12.1478	14.57	29.15	43.22	57.29
	CEECC-GRNN	0.4219	0.2851	8.1140	21.11	43.22	60.80	75.88
Site2	CEE-CC-SVR	0.4088	0.2618	7.3894	22.11	43.22	63.32	75.38
	CEE-CC-ARIMA-SVR	0.3505	0.1868	6.2453	27.27	52.02	66.67	78.79
	CEE-CC-FHG	0.2961	0.1507	5.1952	36.36	56.06	74.24	85.86
	Proposed Model	**0.2600**	**0.1136**	**4.4386**	**37.37**	**66.67**	**78.79**	**91.92**
	CEE-CC-ARIMA	1.0232	1.9481	18.9920	12.50	23.00	30.50	39.00
	CEE-CC-Holt	0.6345	0.6514	11.9340	12.56	30.65	42.21	53.27
	CEECC-GRNN	0.4161	0.2966	8.8052	23.12	44.22	59.30	69.85
Site3	CEE-CC-SVR	0.4088	0.2618	7.3894	22.11	43.22	63.32	75.38
	CEE-CC-ARIMA-SVR	0.3287	0.1886	5.7871	28.28	52.02	67.68	84.34
	CEE-CC-FHG	0.2873	0.1343	5.1527	31.31	57.58	77.78	91.92
	Proposed Model	**0.2554**	**0.1105**	**4.5598**	**37.37**	**62.63**	**85.35**	**92.42**

Table 5
Statistical error measures of the six models for 30-Min wind speed.

Area	Model	MAE (m/s)	MSE (m/s)	MAPE (%)	MPE (% of errors in each margin)			
					±2.5%	±5%	±7.5%	±10%
Site1	WT-CC-FHS	0.4536	0.4699	6.4609	36.68	54.77	69.85	78.39
	SSA-CC- FHS	0.7804	1.1958	11.2939	20.20	35.35	46.97	57.07
	EMD-CC- FHS	0.3809	0.2671	5.4831	33.17	56.78	73.37	84.42
	CEE- FHS	0.3351	0.1922	5.0083	32.83	60.10	76.26	86.87
	CEE-CC- PHS	0.3161	0.1722	4.8470	37.88	59.09	79.29	86.87
	Proposed Model	**0.2708**	**0.1259**	**4.0564**	**40.40**	**69.70**	**85.35**	**93.43**
Site2	WT-CC-FHS	0.4194	0.3447	7.0936	17.09	33.17	50.25	75.38
	SSA-CC- FHS	0.6921	0.8928	11.7166	13.13	29.29	44.95	55.05
	EMD-CC- FHS	0.3501	0.2520	5.9175	33.17	58.79	72.86	83.42
	CEE- FHS	0.3160	0.1666	5.3843	29.80	58.08	73.23	84.34
	CEE-CC- PHS	0.3046	0.1522	5.6108	29.80	53.03	75.25	85.35
	Proposed Model	**0.2600**	**0.1136**	**4.4386**	**37.37**	**66.67**	**78.79**	**91.92**
Site3	WT-CC-FHS	0.3439	0.2241	6.1986	31.16	53.77	69.35	81.41
	SSA-CC- FHS	0.6892	0.8641	12.5879	14.65	27.27	41.41	50.51
	EMD-CC- FHS	0.3588	0.2218	6.2247	25.63	46.73	67.34	84.42
	CEE- FHS	0.3343	0.1829	5.9659	27.27	53.03	71.21	83.33
	CEE-CC- PHS	0.3126	0.1569	5.7798	25.25	52.02	72.73	85.35
	Proposed Model	**0.2554**	**0.1105**	**4.5598**	**37.37**	**62.63**	**85.35**	**92.42**

5. Analysis and discussions

This section presents a detailed discussion of the proposed model. Additionally, the optimization performance of the firefly algorithm is verified by comparison with other optimization algorithms. Finally, a comparative analysis of the relevant literature is made. The validity of the experimental results was analyzed and verified by a hypothesis test. Finally, this study was compared with related literature.

5.1 Comparative analysis

In the previous three experiments, it was proven that CEE-CC-FHS is far superior to the compared models. To comprehensively test the superiority of the model, one statistical method and three intelligent algorithms are considered to compare the wind-speed forecasting accuracy. These algorithms are the ARIMA, the backpropagation neural network (BPNN), GRNN, and ELM, respectively. According to the results shown in Table 6, the following conclusions are drawn.

Compared with the three intelligent algorithms, the proposed model has superior and stable prediction performance. Thus, the proposed model provides an effective forecasting method for wind speed. For a 10-min wind speed at site 1, the proposed model obtains the best forecasting performance. Compared with the ARIMA, BPNN, GRNN, and ELM, the MAPEs of the proposed model exhibit improvements of 7.4216%, 5.1233%, 6.8001%, and 4.9496%, respectively. Similarly, for a 30-min wind speed at site 1, the MAPEs of the proposed model compared with the ARIMA, BPNN, GRNN, and ELM exhibit improvements of 15.0258%, 8.9732%, 9.2045%, and 8.7387%, respectively. According to the comparison error between the proposed model and the three prediction models, the data preprocessing techniques combined with linear and nonlinear prediction frameworks can significantly improve the prediction accuracy.

Table 6
Comparison of prediction performances of four other models and proposed model.

Model	Indices	10-Min			30-Min		
		Site 1	Site 2	Site 3	Site 1	Site 2	Site 3
ARIMA	MAE(m/s)	0.5997	0.8419	0.6345	1.3042	1.2281	1.1702
	MSE(m/s)	0.5635	1.2116	0.6187	2.8842	2.4948	2.3833
	MAPE (%)	9.7446	15.0181	11.6620	19.0822	20.6847	21.7168
	MPE (±5%)	30.00	23.00	27.50	16.00	13.50	12.00
	MPE (±10%)	59.50	44.00	48.00	33.50	24.00	29.00
BPNN	MAE(m/s)	0.4474	0.3844	0.3772	0.8698	0.8132	0.7511
	MSE(m/s)	0.3112	0.2657	0.2314	1.3879	1.1624	1.0198
	MAPE (%)	7.4463	6.9510	7.0167	13.0296	14.2048	14.2712
	MPE (±5%)	41.50	46.00	44.50	31.00	26.00	26.50
	MPE (±10%)	72.50	80.50	78.00	53.50	48.00	49.50
GRNN	MAE(m/s)	0.5592	0.4793	0.5385	0.8741	0.9217	0.7615
	MSE(m/s)	0.5028	0.3807	0.4662	1.3991	1.4194	1.0402
	MAPE (%)	9.1231	8.7769	10.9159	13.2609	17.0663	15.4501
	MPE (±5%)	33.50	36.50	39.50	31.50	20.50	28.50
	MPE (±10%)	59.50	65.50	58.50	52.00	41.00	48.00
ELM	MAE(m/s)	0.4453	0.3615	0.3685	0.8693	0.7858	0.7413
	MSE(m/s)	0.3163	0.2477	0.2272	1.4046	1.0967	1.0117
	MAPE (%)	7.2726	6.4389	6.8388	12.7951	13.6257	13.8748
	MPE (±5%)	44.00	51.50	49.50	30.50	27.00	26.00
	MPE (±10%)	73.00	79.00	76.00	53.50	49.50	50.00
Proposed Model	MAE(m/s)	0.1413	0.1207	0.1252	0.2708	0.2600	0.2554
	MSE(m/s)	0.0321	0.0243	0.0249	0.1259	0.1136	0.1105
	MAPE (%)	2.3230	2.1587	2.2998	4.0564	4.4386	4.5598
	MPE (±5%)	91.41	91.41	91.41	69.70	66.67	62.63
	MPE (±10%)	100.00	100.00	100.00	93.43	91.92	92.42

5.2 Evaluate performance of optimization algorithm

In the previous experiments, the proposed model was compared with the model (CEE-CC-PHS) optimized by the PSO algorithm. The results show that the FA can improve the predictive performance of the model. To more clearly demonstrate the performance of the FA, three typical test functions—Sphere, Rastrigin, and Ackley— were used for verification. The variable fields and function formulas of the test function are shown in Table 7.

Table 7
Three test functions.

Function	Variable domain	Optimum value	Test function
Sphere	[-1,1]	0	$\sum_{i=1}^{d} x_i^2$
Ackley	[-10,10]	0	$-20\exp\left[-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i}\right] - \exp\left[\frac{1}{d}\sum_{i=1}^{d}\cos\left(2\pi x_i\right)\right] + 20 + e$
Rastrigin	[-5,5]	0	$\sum_{i=1}^{d}\left(x_i^2 - 10\left(2\pi x_i\right) + 10\right)$

Additionally, optimization algorithms—the PSO algorithm, the cuckoo search algorithm, and other heuristics algorithms, i.e., the bat algorithm (BA) and flower

pollination algorithm—are selected for comparative analysis. Fig. 9 shows the variation of the fitness with the increase of the number of iterations for different numbers of dimensions. The test results for 30 runs of the FA and the other algorithms are presented in Table 8.

For the Sphere function, the fitness of the five optimization algorithms tends to zero after 200 iterations, regardless of whether the number of dimensions is 10 or 20. However, as shown in Fig. 9, the convergence speed of the FA is better than that of the other four algorithms. When the dimension is 10, the minimum, average, and standard deviation of the FA test results are 6.65E-08, 1.21E-06, and 1.89E-06, respectively. The optimization results show that the FA has the best search ability among the algorithms. For the Ackley function, the test results of different algorithms differ significantly, as shown in Fig. 9. When the dimension is 10, the search results of the FA are significantly better than those of the other four algorithms, and the minimum value of the fitting in 200 iterations is 0.0049. When the dimension is 20, the search results of the FA and BA are better than those of the other three algorithms. In addition, the minimum value of the FA is 0.0428, and the FA has the lowest fitness value. For the Rastrigin function, the convergence speed of the FA in 200 iterations does not differ significantly from those of the other algorithms. For dimensions 10 and 20, the minimum values of the FA are 3.6810 and 23.2860, respectively. The results show that the FA can find the lowest fitness values.

The optimization results for different test functions in different dimensions indicate that FA has a more accurate and stable search capability than the other optimization algorithms. Specifically, the maximum, minimum, average, and standard deviation are lower for the FA than for the other algorithms. Therefore, the search performance and verification results of the FA algorithm are superior to those of the other algorithms.
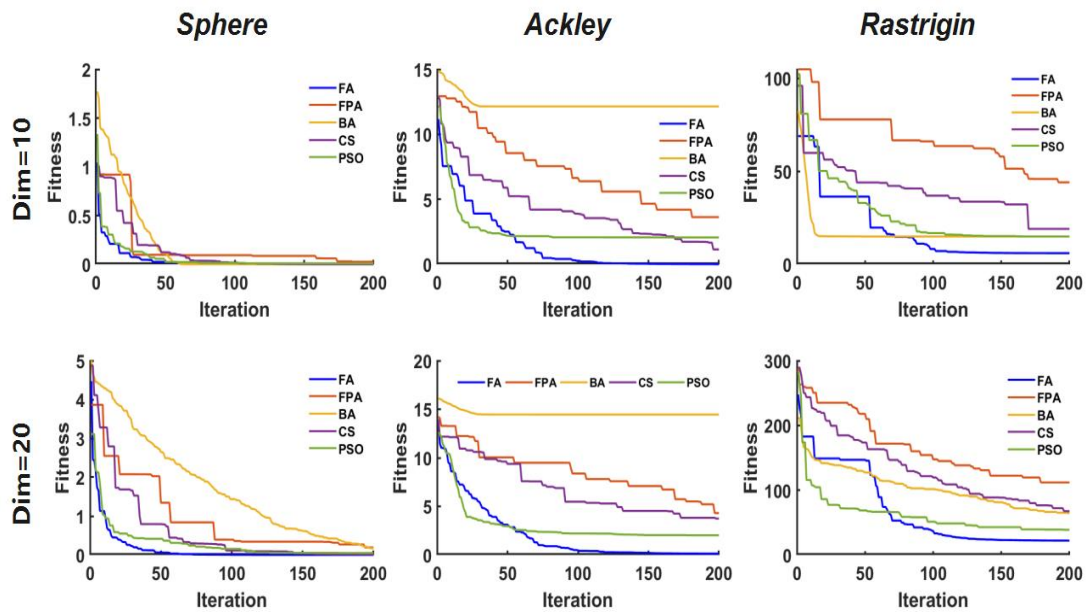

Fig. 9. Test results of FA and other algorithms.

Table 8

Test results of 30 trials of FA and other algorithms.

Function name	Dim	Algorithm	Max value	Min value	Average value	Standard deviation
Sphere	10	PSO	1.64E-02	1.20E-03	6.11E-03	4.33E-03
		CS	3.08E-04	4.67E-05	1.55E-04	7.25E-05
		BA	1.92E-04	1.14E-04	1.44E-04	2.57E-05
		FPA	7.74E-02	2.47E-02	4.03E-02	1.56E-02
		FA	**6.22E-06**	**6.65E-08**	**1.21E-06**	**1.89E-06**
Sphere	20	PSO	0.0851	0.0106	0.0435	0.0310
		CS	0.0385	0.0121	0.0196	0.0083
		BA	0.7408	0.0123	0.1449	0.2235
		FPA	0.3588	0.1485	0.2595	0.0560
		FA	**0.0116**	**0.0019**	**0.0077**	**0.0028**
Ackley	10	PSO	2.6123	1.6681	2.1834	0.3315
		CS	1.7050	0.8431	1.3319	0.2438
		BA	12.5431	10.1768	11.3870	0.7874
		FPA	5.2881	3.4698	4.5021	0.5844
		FA	**0.0154**	**0.0049**	**0.0089**	**0.0037**
Ackley	20	PSO	4.0309	2.1687	2.9099	0.6136
		CS	3.8508	3.3545	3.5989	0.2080
		BA	15.1835	13.6162	14.4169	0.5778
		FPA	7.2390	5.3972	6.2940	0.6090
		FA	**0.1570**	**0.0428**	**0.1075**	**0.0306**
Rastrigin	10	PSO	18.8020	7.5053	12.4701	3.5539
		CS	18.7541	10.5480	15.8923	2.5198
		BA	17.9941	4.1491	11.3233	3.8493
		FPA	40.4292	24.0871	35.5402	4.9954
		FA	**9.3277**	**3.6810**	**7.4864**	**1.8292**
Rastrigin	20	PSO	82.8993	43.7370	60.8321	13.8520
		CS	94.7210	58.4668	76.5177	11.2896
		BA	88.3268	33.9320	67.9996	18.1168
		FPA	135.4307	103.6248	119.0780	9.1622
		FA	**53.9004**	**23.2860**	**42.0311**	**8.4956**

5.3 Test the validity of the model

This section is aimed at comparing the six other models (CEEN-CC-Holt, CEE-CC-ARIMA, CEEN-CC-GRNN, CEE-CC-SVR, CEE-CC-ARIMA-SVR, and CEE-CC-Holt-GRNN) through the error criteria (listed in Table 9). According to the results shown in Table 10, the performance of the forecasting model is compared and analyzed as follows.

With regard to the wind-speed forecasting accuracy, the prediction performance of the hybrid model is better than that of the benchmark model. Table 8 indicates that the value of the percentage increase decreases gradually as the interval of data collection increases. For the proposed model vs. CEE-CC-FHG for a 10-min wind speed, the P_{MAPE} values for sites 1, 2, and 3 are 34.1942%, 27.1784%, and 25.2083%, respectively. According to the comparison order (i.e., compared with CEE-CC-ARIMA, CEE-CC-Holt, CEE-CC-GRNN, CEE-CC-SVR, CEE-ARIMA-SVR, and CEE-CC-FHG), the percentage improvement of the error criterion gradually decreases, which indicates that the prediction performance of the compared model increases. This also signifies that the proposed model has higher prediction accuracy than the other models. The

improvement in the prediction accuracy of the comparison model to a certain level is accompanied by a decrease in the model prediction improvement.

Table 9
Three improvement percentage of criteria.

Metric	Definition	Equation
P_{MAE}	The improvement percentages of MAE	$P_{MAE} = \left\lvert \dfrac{MAE_1 - MAE_2}{MAE_1} \right\rvert$
P_{MSE}	The improvement percentages of MSE	$P_{MSE} = \left\lvert \dfrac{MSE_1 - MSE_2}{MSE_1} \right\rvert$
P_{MAPE}	The improvement percentages of MAPE	$P_{MAPE} = \left\lvert \dfrac{MAPE_1 - MAPE_2}{MAPE_1} \right\rvert$

Table 10
Improvement percentages among the proposed model and the six models.

Model	Indices	10-Min			30-Min		
		site 1	site 2	site 3	site 1	site 2	site 3
Proposed vs. CEE-CC-ARIMA	P_{MAE}	68.9025	82.5996	78.9512	77.0674	75.3659	75.0403
	P_{MSE}	90.7139	96.9254	95.9219	95.1228	94.4731	94.3285
	P_{MAPE}	68.9178	82.5302	78.6028	76.5795	75.1818	75.991
Proposed vs. CEE-CC-Holt	P_{MAE}	65.5717	69.0985	66.8528	59.9277	58.7267	59.7508
	P_{MSE}	89.0418	89.9473	89.3714	84.1537	81.3869	83.0383
	P_{MAPE}	66.3329	70.4587	67.9647	64.4162	63.462	61.7916
Proposed vs. CEE-CC-GRNN	P_{MAE}	56.6762	52.2169	40.2101	35.8763	38.3749	38.6276
	P_{MSE}	78.3059	73.927	64.8801	59.0243	60.1484	62.7546
	P_{MAPE}	57.8832	51.9188	42.9174	42.4161	45.2973	48.2148
Proposed vs. CEE-CC-SVR	P_{MAE}	52.1905	29.703	33.4043	34.3284	36.4048	37.5328
	P_{MSE}	75.8934	49.6894	59.4463	56.4918	56.5986	57.7927
	P_{MAPE}	52.347	29.3873	35.4768	38.5516	39.9337	38.2933
Proposed vs. CEE-CC-ARIMA-SVR	P_{MAE}	44.3642	19.0004	21.6472	21.7603	25.8126	22.2986
	P_{MSE}	69.4511	31.0904	37.8958	35.0727	39.1866	41.4087
	P_{MAPE}	43.2093	20.1063	20.7138	24.7838	28.9297	21.2074
Proposed vs. CEE-CC-FHG	P_{MAE}	33.1186	25.4581	24.2376	12.2575	12.1977	11.0964
	P_{MSE}	52.181	38.9412	37.4163	21.0075	24.6318	17.7311
	P_{MAPE}	34.1942	27.1784	25.2083	13.9598	14.5648	11.5069

5.4 Diebold-Mariano test

The Diebold-Mariano (DM) test was performed to assess the forecasting abilities of the proposed model and the comparison models [41]. The null hypothesis H_0 and the alternative hypothesis H_1 for the DM test are expressed as follows:

$$H_0 : E[L(error_i^1)] = E[L(error_i^2)]] \tag{29}$$

$$H_1 : E[L(error_i^1)] \neq E[L(error_i^2)]] . \tag{30}$$

The DM test statistics can be obtained as

$$DM = \frac{\sum_{i=1}^{n} (L(error_i^1) - L(error_i^2)) / n}{\sqrt{S^2 / n}} s^2 , \tag{31}$$

where s^2 represents an estimation for the variance of $d_i = L(\varepsilon_i^1) - L(\varepsilon_i^2)$. When the DM statistic under the significant level of α falls in the interval $[-Z_{\alpha/2}, Z_{\alpha/2}]$, the null hypothesis H_0 is accepted. This means that the predicted performance is not significantly different between the proposed model and the compared model.

Table 11 compares the five models (WT-CC-FHS, SSA-CC-FHS, EMD-CC- FHS, CEE-FHS, and CEE-CC-PHS) under the confidence level of 1%. The minimum value of |DM| is 2.9036, which is greater than $Z_{0.01/2} = 2.58$. The results show that the probability difference between the proposed model and the compared models is 99%. The DM test proves not only the effectiveness of the experiment but also the superiority of the proposed model.

Table 11

The results of Diebold–Mariano test.

DM test	Average value	Site1	Site2	Site3
	WT-CC- FHS	8.7471[*]	6.6088[*]	8.1942[*]
	SSA-CC- FHS	8.6197[*]	6.0762[*]	7.3828[*]
10-Min	EMD-CC- FHS	4.9286[*]	5.2528[*]	5.1451[*]
	CEE-FHS	7.7355[*]	6.4175[*]	7.3282[*]
	CEE-CC-PHS	4.9232[*]	5.1551[*]	2.9036[*]
	WT-CC- FHS	4.7042[*]	5.7147[*]	4.0590[*]
	SSA-CC- FHS	6.4104[*]	6.3993[*]	5.8269[*]
30-Min	EMD-CC- FHS	5.2752[*]	3.9622[*]	5.1214[*]
	CEE-FHS	6.4708[*]	7.3642[*]	7.0861[*]
	CEE-CC-PHS	4.2017[*]	5.4364[*]	4.3410[*]

Note: [*]1% significance level.

5.5 Comparison of the related literature

Prediction methods based on decomposition techniques have been widely applied to the actual wind speed [42,43]. Wang et al. [44] reported that the model established by the EMD-based prediction algorithm may not be suitable for the updated time series in actual wind-speed prediction and proposed an approximate prediction model based on EMD. With regard to the decomposition method, the processing idea of this study is similar to that of Wang [44]; that is, the most frequent IMF_1 is removed from the original time series, and the approximate time series is obtained for the next analysis. When the original time series is decomposed with the newly obtained data, the proportion of the IMF_1 in the original wind-speed time series is small; thus, the influence of the IMF_1 change on the time series is not significant.

Additionally, in this study, reconstruction technology is used to select features, reconstruct the phase space of the approximate time series to determine the delay time and dimension of the time series, and extend one-dimensional time series to a matrix time series. The matrix time series is used as an input of the model to more clearly show the variation of the time series [45]. According to the literature [46], C-C technology has achieved satisfactory forecasting results in determining the input form of the model. However, in these studies, the model did not capture well the linear and nonlinear factors of the time series from the optimal input determined by the C-C method. Accordingly, in the present study, a "linear and nonlinear" modeling framework was developed to comprehensively excavate the information in the time series. Furthermore, the firefly optimization algorithm is an effective technique that provides optimal model parameters to improve the prediction performance of the model. The performance of the firefly optimization algorithm is described in detail in [47]. Additionally, the experimental results show that the data denoising technique CEEMDAN provides a

significant model prediction performance improvement. A similar conclusion was drawn in [48]; however, that model does not comprehensively capture the information in the time series after CEEMDAN is used to remove the interference factors. Moreover, in this study, CEEMDAN employed PSR to maximize the mining of hidden real-time information in the time series.

6. Conclusions

Accurate wind-speed prediction is of great help for the improvement of wind energy efficiency. However, the wind-speed time series contain many disturbing factors, which pose a severe challenge to accurate wind-speed forecasting. Abundant work has been devoted to improving the performance of models for forecasting wind speed. Therefore, it is very valuable to develop a model with optimized performance for wind-speed prediction.

In this paper, the CEE-CC-FHS model is proposed for chaotic series forecasting. The importance of data preprocessing was verified, which provided as much information as possible for the predictive model. CEEMDAN was developed to eliminate interference components from the original data, and PSR was proven to be a feasible method for capturing the information of time-series fluctuation. The matrix time series can provide more information contained in the time series for the forecasting. Thus, changing wind speeds can be converted into chaotic systems for research.

Experiments indicated that the proposed model removes the interference factors of the chaotic time series and provides valuable predictive information while having the ability to capture linear and nonlinear modes, especially for a chaotic time series characterized by complexity and irregularity. In Experiment I, the average MAPE of the proposed model compared with other models (CEE-CC-ARIMA, CEE-CC-Holt, CEE-CC-GRNN, CEE-CC-SVR, CEE-CC-ARIMA-SVR, and CEE-CC-Holt-GRNN) exhibited improvements of 7.9324%, 4.8683%, 2.4176%, 1.5716%, 0.9705%, and 0.9293%, respectively. In Experiment II, the average MAPE of the designed forecasting architecture exhibited improvements of 4.2021%, 3.8718%, 1.2593%, 0.8530%, and 0.7295% compared with the benchmark models WT/SSA/EMD CC-FHS, CEE-FHS, and CEE-CC-PHS, respectively. In Experiment III, the average MAPE of the proposed model was better than that of the compared models.

The results indicate that the wind speed at different sites can be effectively predicted by the proposed model. Furthermore, the experimental results based on statistical criteria indicate that the proposed model exhibits better forecasting performance than other models. Overall, the model improves the accuracy of wind-speed prediction and provides a new feasible solution for wind-speed prediction and rational power-grid allocation.

Acknowledgements

Appendix A

List of abbreviations

Nomenclature			
		ARIMA	Autoregressive Integrated Moving Average
		PSR	predicted value of the nth datum
CEEMDAN	complete ensemble empirical mode decomposition with adaptive noise	τ	delay time
		m	embedding dimension
EEMD	ensemble empirical mode decomposition	**Holt**	holt's exponential smoothing
EMD	empirical mode decomposition	S_t	preliminary predictive value
WT	wavelet transfer	T_t	local growth value
IMF	intrinsic mode function	L_t	linear components
SVR	support vector regression	N_t	non-linear components
RBF	radial basic function	α	the smoothing parameter
PSO	particle swarm optimization	γ	the smoothing parameter
FA	firefly algorithm	**CTS**	chaotic time series
Gen_{Max}	maximum number of iterations	**GRNN**	general regression neural network
I_0	original light intensity.	**MLYE**	maximum Lyapunov exponent
r_{ij}	distance between any two fireflies, i and j	**ELM**	extreme learning machine
P	number of fireflies	**ANN**	artificial neural network
SSA	Singular Spectrum Analysis	**SVM**	support vector machine

References

[1] Xie L, Carvalho PMS, Ferreira L a FM, Juhua Liu, Krogh BH, Popli N, et al. Wind Integration in Power Systems: Operational Challenges and Possible Solutions. Proc IEEE 2011. http://doi:10.1109/JPROC.2010.2070051.

[2] Khalid M, Savkin AV. A method for short-term wind power prediction with multiple observation points. IEEE Trans Power Syst 2012. http://doi:10.1109/TPWRS.2011.2160295.

[3] Ackermann T, Söder L. Wind energy technology and current status: a review. Renew Sustain Energy Rev 2000. http://doi:10.1016/S1364-0321(00)00004-6.

[4] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 2014. http://doi:10.1016/j.rser.2013.12.054.

[5] Liu H, Tian Hqi, Li Yfei. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Appl Energy 2012. http://doi:10.1016/j.apenergy.2012.04.001.

[6] Zhao J, Guo ZH, Su ZY, Zhao ZY, Xiao X, Liu F. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. Appl Energy 2016. http://doi:10.1016/j.apenergy.2015.10.145.

[7] Yang D, Sharma V, Ye Z, Lim LI, Zhao L, Aryaputera AW. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy 2015. http://doi:10.1016/j.energy.2014.11.082.

[8] Torres JL, García A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Sol Energy 2005. http://doi:10.1016/j.solener.2004.09.013.

[9] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009. http://doi:10.1016/j.renene.2008.09.006.

[10] Koo J, Han GD, Choi HJ, Shim JH. Wind-speed prediction and analysis based on geological and distance variables using an artificial neural network: A case study in South Korea. Energy 2015. http://doi:10.1016/j.energy.2015.10.026.

[11] Fu T, Wang C. A hybrid wind speed forecasting method and wind energy resource analysis based on a swarm intelligence optimization algorithm and an artificial intelligence model. Sustain 2018;10. http://doi:10.3390/su10113913.

[12] Liu D, Wang J, Wang H. Short-term wind speed forecasting based on spectral clustering and optimised echo state networks. Renew Energy 2015;78:599–608. http://doi:10.1016/j.renene.2015.01.022.

[13] Nikolić V, Motamedi S, Shamshirband S, Petković D, Ch S, Arif M. Extreme learning machine approach for sensorless wind speed estimation. Mechatronics 2015. http://doi:10.1016/j.mechatronics.2015.04.007.

[14] Hu J, Wang J, Ma K. A hybrid technique for short-term wind speed prediction. Energy 2015. http://doi:10.1016/j.energy.2014.12.074.

[15] Wang J, Wang J. Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations. Energy 2016. http://doi:10.1016/j.energy.2016.02.098.

[16] Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting. Appl Energy 2017;208:344–60. http://doi:10.1016/j.apenergy.2017.10.031.

[17] Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 2003. http://doi:10.1016/S0925-2312(01)00702-0.

[18] Li H, Wang J, Lu H, Guo Z. Research and application of a combined model based on variable weight for short term wind speed forecasting. Renew Energy 2018. http://doi:10.1016/j.renene.2017.09.089.

[19] Zhang C, Zhou J, Li C, Fu W, Peng T. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. Energy Convers Manag 2017. http://doi:10.1016/j.enconman.2017.04.007.

[20] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manag 2017. http://doi:10.1016/j.enconman.2017.01.022.

[21] Du P, Wang J, Guo Z, Yang W. Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting. Energy Convers Manag 2017. http://doi:10.1016/j.enconman.2017.07.065.

[22] Wang J, Niu T, Lu H, Guo Z, Yang W, Du P. An analysis-forecast system for uncertainty modeling of wind speed: A case study of large-scale wind farms. Appl Energy 2018. http://doi:10.1016/j.apenergy.2017.11.071.

[23] Jiang Y, Huang G. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers Manag 2017;144:340–50. http://doi:10.1016/j.enconman.2017.04.064.

[24] Wang C, Zhang H, Fan W, Ma P. A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction. Energy 2017. http://doi:10.1016/j.energy.2017.07.112.

[25] Torres ME, Colominas MA, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. 2011 IEEE Int Conf Acoust Speech Signal Process 2011. http://doi:10.1109/ICASSP.2011.5947265.

[26] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manag 2017. http://doi:10.1016/j.enconman.2017.01.022.

[27] Ren Y, Suganthan PN, Srikanth N. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. IEEE Trans Sustain Energy 2015;6:236–44. http://doi:10.1109/TSTE.2014.2365580.

[28] Guo Z, Chi D, Wu J, Zhang W. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm. Energy Convers Manag 2014. http://doi:10.1016/j.enconman.2014.04.028.

[29] Lorenz EN. Deterministic Nonperiodic Flow. J Atmos Sci 1963. http://doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

[30] Tennyson J. In the wake of chaos. Unpredictable order in dynamical systems. Endeavour 1995. http://doi:10.1016/0160-9327(95)90056-X.

[31] Huang N, Shen Z, Long S, Wu M, SHIH H, ZHENG Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc A Math Phys Eng Sci 1998. http://doi:10.1098/rspa.1998.0193.

[32] Wu Z, Huang NE. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. Adv Adapt Data Anal 2009. http://doi:10.1142/S1793536909000047.

[33] Wang T, Zhang M, Yu Q, Zhang H. Comparing the applications of EMD and EEMD on time-frequency analysis of seismic signal. J Appl Geophys 2012. http://doi:10.1016/j.jappgeo.2012.05.002.

[34] Jiang P, Li C, Li R, Yang H. An innovative hybrid air pollution early-warning

system based on pollutants forecasting and Extenics evaluation. Knowledge-Based Syst 2018. http://doi:10.1016/j.knosys.2018.10.036.

[35] Takens F. Dynamical systems and turbulence lecture notes in mathematics. Berlin:Springer; 1981. 366e81.

[36] Kim HS, Eykholt R, Salas JD. Nonlinear dynamics, delay times, and embedding windows. Phys D Nonlinear Phenom 1999. http://doi:10.1016/S0167-2789(98)00240-1.

[37] Yang X-S. Nature-Inspired Metaheuristic Algorithms. Nature-Inspired Metaheuristic Algorithms Second Edition 2010. http://doi:10.1016/B978-0-12-416743-8.00005-1.

[38] Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. Int J Forecast 2004. http://doi:10.1016/j.ijforecast.2003.09.015.

[39] Vapnik VN. The nature of statistical learning theory. IEEE Trans Neural Networks 1995;8:1564–1564. http://doi:10.1109/TNN.1997.641482.

[40] Chen K-Y. Combining linear and nonlinear model in forecasting tourism demand. Expert Syst Appl 2011. http://doi:10.1016/j.eswa.2011.02.049.

[41] Wang J, Du P, Lu H, Yang W, Niu T. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput J 2018;72:321–37. http://doi:10.1016/j.asoc.2018.07.022.

[42] Jiang P, Yang H, Heng J. A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. Appl Energy 2019;235:786–801. http://doi:10.1016/j.apenergy.2018.11.012.

[43] Zhao X, Wang C, Su J, Wang J. Research and Application Based on the Swarm Intelligence Algorithm and Artificial Intelligence for Wind Farm Decision System. Renew Energy 2018. http://doi:10.1016/j.renene.2018.11.061.

[44] Wang Y, Wu L. On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. Energy 2016;112:208–20. http://doi:10.1016/j.energy.2016.06.075.

[45] Sun S, Qiao H, Wei Y, Wang S. A new dynamic integrated approach for wind speed forecasting. Appl Energy 2017;197:151–62. http://doi:10.1016/j.apenergy.2017.04.008.

[46] Wang Y, Wang J, Wei X. A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: A case study of wind farms in northwest China. Energy 2015. http://doi:10.1016/j.energy.2015.08.039.

[47] Yang X. Firefly Algorithm. 2010. http://doi:10.1016/B978-0-12-416743-8.00005-1.

[48] Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers Manag 2017. http://doi:10.1016/j.enconman.2017.01.022.