Applied Energy 215 (2018) 643–658

Contents lists available at ScienceDirect

Applied Energy

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

A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting

Jingjing Song[a], Jianzhou Wang[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

HIGHLIGHTS

- A combined model based on advanced optimization algorithm is successfully proposed.
- Design three experiments from the real wind farm to verify the effectiveness.
- The proposed combined model can enhance the forecasting accuracy significantly.
- Experiments demonstrate the availability and reliability of the developed model.

ARTICLE INFO

Keywords:
Wind speed forecasting
Combined model
Data preprocessing technique
Advanced optimization algorithm

ABSTRACT

Short-term wind speed forecasting has a significant influence on enhancing the operation efficiency and increasing the economic benefits of wind power generation systems. A substantial number of wind speed forecasting models, which are aimed at improving the forecasting performance, have been proposed. However, some conventional forecasting models do not consider the necessity and importance of data preprocessing. Moreover, they neglect the limitations of individual forecasting models, leading to poor forecasting accuracy. In this study, a novel model combining a data preprocessing technique, forecasting algorithms, an advanced optimization algorithm, and no negative constraint theory is developed. This combined model successfully overcomes some limitations of the individual forecasting models and effectively improves the forecasting accuracy. To estimate the effectiveness of the proposed combined model, 10-min wind speed data from the wind farm in Peng Lai, China are used as case studies. The experiment results demonstrate that the developed combined model is definitely superior compared to all other conventional models. Furthermore, it can be used as an effective technique for smart grid planning.

1. Introduction

Owing to the deterioration of the environment and the depletion of conventional energy resources, renewable energy has aroused widespread interest and research enthusiasm [1]. Wind energy is of crucial importance among the low-carbon energy technologies. It has the potential to achieve sustainable energy supply and occupies an indispensable position in the global new energy market [2]. Moreover, it has experienced an unexpectedly high growth recently and has received increasing attention globally [3]. In 2016, its annual market was 54.6 GW of wind capacity worldwide with a total installed capacity of almost 487 GW according to the Global Wind Energy Council [4].

However, many problems related to wind power generation that seriously restrict its development have arisen. More specifically, the stochasticity and intermittency of wind power could decrease system forecast reliability and wind power quality [5]. Furthermore, the continuous fluctuation of wind speed makes it difficult to predict how much power will be injected into the distribution network, which may cause energy transportation issues. Generally, wind speed forecasting can effectively reduce the risks associated with wind power generation arising from wind-related uncertainty [6,7]. Therefore, many researchers have given much attention to research methods of wind speed prediction. These methods are categorized into the following four types [8]: (i) physical arithmetic, (ii) conventional statistical arithmetic, (iii) spatial correlation arithmetic, and (iv) artificial intelligence arithmetic.

Physical arithmetic mainly utilizes physical data such as temperature, speed, density, and topography information[9], which are based on numerical weather prediction (NWP) model [10,11], to forecast

* Corresponding author at: School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China.
E-mail address: wjz@lzu.edu.cn (J. Wang).

https://doi.org/10.1016/j.apenergy.2018.02.070
Received 26 October 2017; Received in revised form 11 January 2018; Accepted 8 February 2018
Available online 20 February 2018
0306-2619/

wind speed in the succeeding periods. Cheng et al. [12] assimilated measured wind speed values from wind turbines into a NWP system, which enhanced the prediction accuracy of wind speed and wind power. Cassola et al. [13] improved the forecasting performance of wind speed by using the Kalman filtering technique in correcting the NWP model output. However, physical methods are not capable of dealing with short-term horizons and cost a lot of computing time and resources [14]. Statistical arithmetic, which is more appropriate for short-term wind speed forecasting, merely utilizes historical data to predict wind speed and employs the differences between the actual and forecasted wind speed to adjust the model parameters [15,16]. Examples are the auto-regressive moving average (ARMA) [17], auto-regressive integrated moving average (ARIMA) [18], and fractional ARIMA models [19]. Han et al. [20] proposed two hybrid models based on ARMA model and the non-parametric model. Results showed that non-parametric based hybrid models generally outperformed other models. Maatallah et al. [21] developed a novel wind speed prediction model by combining the Hammerstein model and an autoregressive model, which ultimately obtained a higher forecasting accuracy. However, according to the assumption that there are linear patterns among the time series, nonlinear structures cannot be effectively captured by statistical methods [22]. Moreover, spatial correlation arithmetic considers the spatial relationship of wind speed from different sites. For example, Tascikaraoglu et al. [23] proposed a novel wind speed prediction model using a wavelet transform and a spatio-temporal method, which improved the short-term wind speed forecasting relative to other benchmark models. Nevertheless, this model requires wind speed measurements from multiple spatial correlated sites; thus, it is difficult to implement owing to the stringent measurement requirements and their time delays [24].

In addition, as artificial intelligence arithmetic has rapidly developed and has been widely used, many researchers have effectively utilized artificial intelligence methods to predict wind speed. These methods include artificial neural networks (ANNs) [25,26,27], support vector machines (SVMs) [28,29], and fuzzy logic (FL)methods [30,31]. In the literature, some applications of artificial intelligence methods are as follows. Guo et al. [32] presented a novel hybrid approach based on back propagation neural network (BPNN) and seasonal exponential adjustment, which effectively enhanced the prediction accuracy of wind speed. Liu et al. [33] developed a new forecasting approach using Elman neural network (ENN) and a secondary decomposition algorithm, which exhibited satisfactory performance in multi-step wind speed forecasting. Zhang et al. [34] employed a radial basis function (RBF) neural network and a multi-objective optimization method to perform interval forecasting of wind speed and ultimately achieved a higher forecasting precision. Xiao et al. [35] developed a hybrid wavelet neural network (WNN) model with an improved cuckoo search algorithm. Experiment results demonstrated that the hybrid model significantly reduced the prediction error relative to other comparative models in electrical power system forecasting. Zhou et al. [36] successfully proposed a systematic method to adjust the parameters of SVM for short-term wind speed forecasting resulting in a better precision performance. Owing to the stronger nonlinear forecasting capacity of the model, artificial intelligence arithmetic is generally superior compared to the time-series models [37].

Over the past few decades, numerous wind speed forecasting approaches, which have enhanced the forecasting accuracy of wind speed series to a certain extent, have been presented. However, considering the relatively noisy and unstable characteristics of wind speed data, wind speed prediction by directly using the original data would lead to substantial forecasting errors and poor performance [38]. Therefore, with the aim of decreasing the stochastic disturbance of data sequence and achieving a higher forecasting accuracy, data preprocessing techniques, such as the empirical mode decomposition (EMD) [39] and the ensemble empirical mode decomposition (EEMD), have been considered and applied for wind speed forecasting [40]. Although these two techniques have improved the forecasting performance to a certain extent, they still have some disadvantages, such as the mode mixing problem in EMD and the residual noise in EEMD. Considering the defects of the aforementioned data preprocessing methods, a novel technique called improved complete ensemble empirical mode decomposition adaptive noise (ICEEMDAN), proposed by Marcelo et al. [41], is employed for reducing the noise and uncertainty of wind speed series in this study.

Review of previously literature shows that the forecasting approaches discussed above have some inherent drawbacks. The disadvantages of such methods are summarized as follows:

(1) Physical arithmetic is extremely weak in coping with short-term horizons; therefore, these methods do not have accurate and effective results in short-term forecasting. Moreover, they are fairly sensitive to market information and require a large amount of operation time and computational resources. (2) Conventional statistical arithmetic cannot address forecasting with high noise and fluctuation, irregular and nonlinear trends, or features of wind speed series data, which are mainly limited by the prior assumption of a linear form among time series. Furthermore, in realistic cases, these approaches require numerous historical data for wind speed prediction and have a high dependence on data; thus, once the original data change abruptly due to environmental or social factors, prediction errors will suddenly increase much [42]. (3) Spatial correlation arithmetic makes it relatively difficult to implement perfect wind speed forecasting owing to the vast quantities of information such as wind speed values of many spatially correlated sites that need to be considered and collected [43]. (4) Distinct from the other approaches, artificial intelligence arithmetic, which could successfully capture hidden non-linear relationships among given historical data, has been widely researched and applied to address complicated relationships and effectively perform forecasting [44]. However, there are still many disadvantages and defects with artificial intelligence methods, for example, easily getting into a local optimum, over-fitting, and exhibiting a relatively low convergence rate [45]. (5) The individual forecasting models do not pay attention to the necessity and importance of data preprocessing, and hence, cannot always achieve a high forecasting accuracy and satisfy the requirements of time series forecasting. Therefore, owing to the unavoidable drawbacks of the individual models, the abovementioned forecasting methods cannot always capture the wind speed trend and cannot be applied in all situations. Consequently, a combined model, which is deemed as an excellent method that utilizes the advantages of individual approaches to obtain a higher forecasting accuracy, has often been taken into consideration [46].

Based on the analysis above, a novel combined model is developed in this study. It combines a data preprocessing technique, an advanced optimization algorithm, no negative constraint theory (NNCT) [47], and several forecasting algorithms, namely BPNN [48], ENN [49], WNN [50], and generalized regression neural network (GRNN) [51]. It successfully capitalizes on the merits of the individual forecasting models, resulting in further improvements. More specifically, based on the decomposition and ensemble strategy, the original wind speed series are decomposed and reconstructed into a filtered time series, which ensures that the high frequency noise signal is eliminated effectively. Then, several individual algorithms are used for forecasting the processed wind speed data. Next, a novel deciding weight method based on a swarm intelligence-based evolutionary computation technique and the leave-one-out strategy is successfully developed to integrate the individual models and obtain the final forecasting result. As far as we know, this advanced computation technique has been effectively employed in some hybrid models to enhance the forecasting accuracy, being initially performed to optimize the weight of each model and employed in the combined forecasting model. The primary contributions and novelties of this study are described below:

(1) ***Based on the decomposition and ensemble strategy, a data***

preprocessing technique is employed to eliminate the negative influence of high frequency noise and extract the main characteristics of the data. The original wind speed series are decomposed and reconstructed into a filtered time series, which could decrease the uncertainty and irregularity of wind speed data and effectively enhance the wind speed forecasting performance.

(2) ***A novel deciding weight method based on a swarm intelligence-based evolutionary computation technique and the leave-one-out strategy is successfully developed to integrate the individual models.*** To find the optimal weights of each model, the reserved dataset inspired by the leave-one-out strategy is adopted in the deciding weight stage, which combines the evolutionary computation technique by updating the position under the leadership hierarchy of the grey wolf.

(3) ***Based on four ANNs, the developed novel combined model significantly improves the forecasting accuracy of short-term wind speed.*** The proposed combined model effectively takes advantage of the strengths of its component models and overcomes the limitations of low accuracy and instability of the conventional individual models.

(4) ***A more scientific and comprehensive evaluation is conducted to estimate the forecasting performance of the developed combined model in this study.*** Three experiments using multi-step forecasting, four performance metrics, and five discussions, are employed in this evaluation system, which provides an effective assessment in terms of forecasting accuracy of the model.

(5) ***The developed novel combined model provides a powerful technical support for the scheduling and management of smart grids.*** This model is simulated and tested based on wind speed data of several sites in a large wind farm, which demonstrates that it could effectively improve the accuracy of wind speed prediction compared to the conventional forecasting models.

The structure of this paper is arranged as follows. Section 2 introduces the specific methodology, including the data preprocessing technique, optimization algorithm, and the proposed combined model. To illustrate the performance of the developed model, three different experiments are carried out as described in Section 3. Specifically, the datasets, performance metrics, and testing method applied for this study are presented in Sections 3.1,3.2, and 3.3, while the experiment results of the proposed combined model compared to other models are discussed and analyzed in Section 3.4. To further prove the accuracy and availability of the novel combined model, several concrete discussions are presented in Section 4. Finally, the important results and conclusion are presented in Section 5.

2. Methodology

In this section, the specific methods of the proposed combined model, namely the ICEEMDAN technique and GWO algorithm are discussed in detail. Then, a novel combined model to enhance the wind speed accuracy is presented.

2.1. The ICEEMDAN technique

The EMD technique developed by Huang et al. is an adaptive data-driven method that decomposes the original series into several intrinsic mode functions through a sifting process. To alleviate the mode mixing problem in practical applications, Wu and Huang proposed the EEMD. Even though the EEMD has a wide range of applications, it has a new problem in that it is difficult to completely eliminate residual noise. Subsequently, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was proposed by Torres et al. [52]. It made improvements based on the EEMD and achieved a negligible reconstruction error. In addition, it solved the problem of different number of modes. However, the problems of some residual noise and "spurious" modes still require improvement.

Considering the two previous problems, Marcelo et al. developed a new algorithm, the ICEEMDAN technique. The new optimization algorithm introduces operators $H_k(\cdot)$ and $N(\cdot)$. Let $H_k(\cdot)$ be the operator that produces the k_{th} mode obtained by EMD. Let $N(\cdot)$ be the operator that produces the local mean of the signal, and $w^{(i)}$ be a realization of white Gaussian noise with zero mean and unit variance. The explicit algorithm of ICEEMDAN can be described as follows:

Step 1: Let x be the signal of interest, and then calculate the local means of I realizations using the EMD algorithm:

$$x^{(i)} = x + \beta_0 H_1(w^{(i)}) \tag{1}$$

where $\beta_o = \varepsilon_0 std(x)/std(H_1(w^{(i)}))$, and ε_0 represents the reciprocal of the desired signal-to-noise ratio between the first added noise and the analyzed signal.

Step 2: Calculate the first residue R_1:

$$R_1 = (N(x^{(i)})) \tag{2}$$

Step 3: Compute the first mode at the first stage ($k = 1$) using the following formula:

$$d_1 = x - R_1 \tag{3}$$

Step 4: Estimate the second residue as the average of local means of the realizations $R_1 + \beta_1 \mathrm{H}_2(\mathrm{w}^{(\mathrm{i})})$ and define the second mode:

$$d_2 = R_1 - R_2 = R_1 - (N(R_1 + \beta_1 H_2(w^{(i)}))) \tag{4}$$

Step 5: For $k = 3,\ldots,K$, calculate the k_{th} residue:

$$R_k = (N(R_{k-1} + \beta_{k-1} H_k(w^{(i)}))) \tag{5}$$

$$\beta_k = \varepsilon_0 std(r_k), k \geqslant 1 \tag{6}$$

Step 6: Compute the k_{th} mode: $d_k = R_{k-1} - R_k$

Step 7: Go back to step 4 for the next.

2.2. The GWO algorithm

Meta-heuristic optimization algorithms have received extensive attention in the last two decades. The GWO algorithm, which is a swarm intelligence-based evolutionary computation technique proposed by Mirjalili et al. [53], is a new and effective meta-heuristic algorithm. It was checked based on 29 well-known test functions and was found to achieve very competitive results compared to other meta-heuristics such as gravitational search algorithm, evolutionary programming, and evolution strategy. This algorithm is inspired by grey wolves; it simulates the grey wolf's leadership and hunting mechanism. On one hand, four types of grey wolves are applied to imitate the leadership hierarchy. On the other hand, four major steps of grey wolves, namely, encircling prey, hunting, attacking prey, and searching for prey, are implemented during the process. The GWO algorithm mechanism is introduced as follows:

2.2.1. Social hierarchy

When mathematically simulating the social hierarchy of grey wolves, we consider the first three best solutions as *a*, *b*, and *c* respectively. The rest of alternative solutions are named as *w*. Consequently, the nature of the GWO algorithm is that *a*, *b*, and *c* wolves guide the optimization process, while being followed by *w* wolves.

2.2.2. Encircling prey

As mentioned above, grey wolves encircle their prey during hunting. The equations of encircling behavior in a mathematical model are expressed as follows:

$$\overrightarrow{D} = |\overrightarrow{C} \cdot \overrightarrow{X_p}(t) - \overrightarrow{X}(t)| \tag{7}$$

$$\overrightarrow{X}(t+1) = \overrightarrow{X_p}(t) - \overrightarrow{A} \cdot \overrightarrow{D} \tag{8}$$

where t indicates the current iteration, $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors,

$\overrightarrow{X_p}$ is the position vector of the prey, and $\overrightarrow{X}$ indicates the position vector of a grey wolf.

The vectors $\overrightarrow{A}$ and $\overrightarrow{C}$ are calculated as follows:

$$\overrightarrow{A} = 2\overrightarrow{e}\cdot\overrightarrow{r_1} - \overrightarrow{e} \tag{9}$$

$$\overrightarrow{C} = 2\overrightarrow{r_2} \tag{10}$$

where $\overrightarrow{e}$ components are linearly decreased from 2 to 0 over the course of iterations and r_1 and r_2 are random vectors in [0,1].

2.2.3. Hunting

Grey wolves are able to identify the location of prey and then encircle it. In most cases, the hunting activity is led by *a*, while *b* and *c* wolves may also take part in the hunt sometimes. Nevertheless, we are unaware of the exact location of the prey. For simulating the hunting behavior of grey wolves using a mathematical method, we assume that *a*, *b*, and *c* wolves have a better understanding of the possible location of the prey. Accordingly, the first three fittest solutions obtained are saved and the other search agents, including the *w* wolves, are required to update their positions according to the best search agent. The following mathematical formulas comprise this process:

$$\overrightarrow{D_a} = |\overrightarrow{C_1}\cdot\overrightarrow{X_a} - \overrightarrow{X}|, \overrightarrow{D_b} = |\overrightarrow{C_2}\cdot\overrightarrow{X_b} - \overrightarrow{X}|, \overrightarrow{D_c} = |\overrightarrow{C_3}\cdot\overrightarrow{X_c} - \overrightarrow{X}| \tag{11}$$

$$\overrightarrow{X_1} = \overrightarrow{X_a} - \overrightarrow{A_1}(\overrightarrow{D_a}), \overrightarrow{X_2} = \overrightarrow{X_b} - \overrightarrow{A_2}(\overrightarrow{D_b}), \overrightarrow{X_3} = \overrightarrow{X_c} - \overrightarrow{A_3}(\overrightarrow{D_c}) \tag{12}$$

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{13}$$

2.2.4. Attacking prey

As a stopping move, grey wolves end up hunting by attacking the prey. For approaching the prey in the mathematical model, we reduce the value of $\overrightarrow{e}$. Meanwhile, the fluctuation range of $\overrightarrow{A}$ is also reduced along with $\overrightarrow{A}$. That is, we decrease $\overrightarrow{e}$ from 2 to 0, and then the random value of $\overrightarrow{A}$ is also changed in the interval [−e, e] during the iterations. Moreover, the next position of a search agent could be anywhere between its present position and the prey, when the value of $\overrightarrow{A}$ is in [−1, 1].

2.2.5. Searching for prey

Grey wolves search for prey based on the location of *a*, *b* and *c*. They diverge and gather when they search for prey and attack the prey respectively. In order to mathematically simulate divergence, the search agent is obliged to separate from the prey with random values greater than 1 or less than −1. When mathematically simulating divergence, we take advantage of the random values greater than 1 or less than −1 to oblige the search agent to diverge from the prey, which causes the algorithm to search globally. In addition, $\overrightarrow{C}$ is another component of favoring exploration, which includes random values in the interval [0, 2]. It provides random weights for the prey so that the algorithm could show a more random behavior by optimization, favoring exploration and avoiding local optima. Note that *C* is not linearly decreased compared to *A*. *C* is required to provide random values in order to emphasize exploration during the entire iteration process, which helps to address the local optima stagnation, particularly during the final iterations.

The pseudo code of GWO is described as follows:

Algorithm 1 **: GWO**

Parameters:
$Iter_{\text{Max}}$ - the maximum number of iterations
n - the number of wolves
$\boldsymbol{F}_i$ - the fitness function of i_{th} *wolf*
$\boldsymbol{x}_i$ - the position of i_{th} wolf
t - the current iteration number
d - the number of dimension

```
/*Set the parameters of GWO.*/
/*Initialize population of n wolves x_i (i = 1, 2… n) randomly.*/
FOR EACH i: 1 ≤ i ≤ n DO
    Evaluate the corresponding fitness function F_i
END FOR
/*Determine the best search agent X*.*/
WHILE (t< Iter_Max) DO
    FOR EACH i=1:n DO
        FOR EACH j=1: n DO
            /*Update a, A and C.*/
            IF (|A|<1) THEN
                /*Update the position of the current search agent.*/
                IF (|A|≥1) THEN
                    Select a random search agent (X_rand )
                    /*Update the position of the current search agent.*/
                    X(t+1) = X_p(t) − A·D
                END IF
            END IF
        END FOR
    END FOR
    /*Check if any search agent goes beyond the search space and amend it.*/
    FOR EACH i: 1 ≤ i ≤ n DO
        Calculate the fitness of each search agent F_i
    END FOR
    /*Update the best search agent X*.*/
    t=t+1
END WHILE
RETURN X*
```

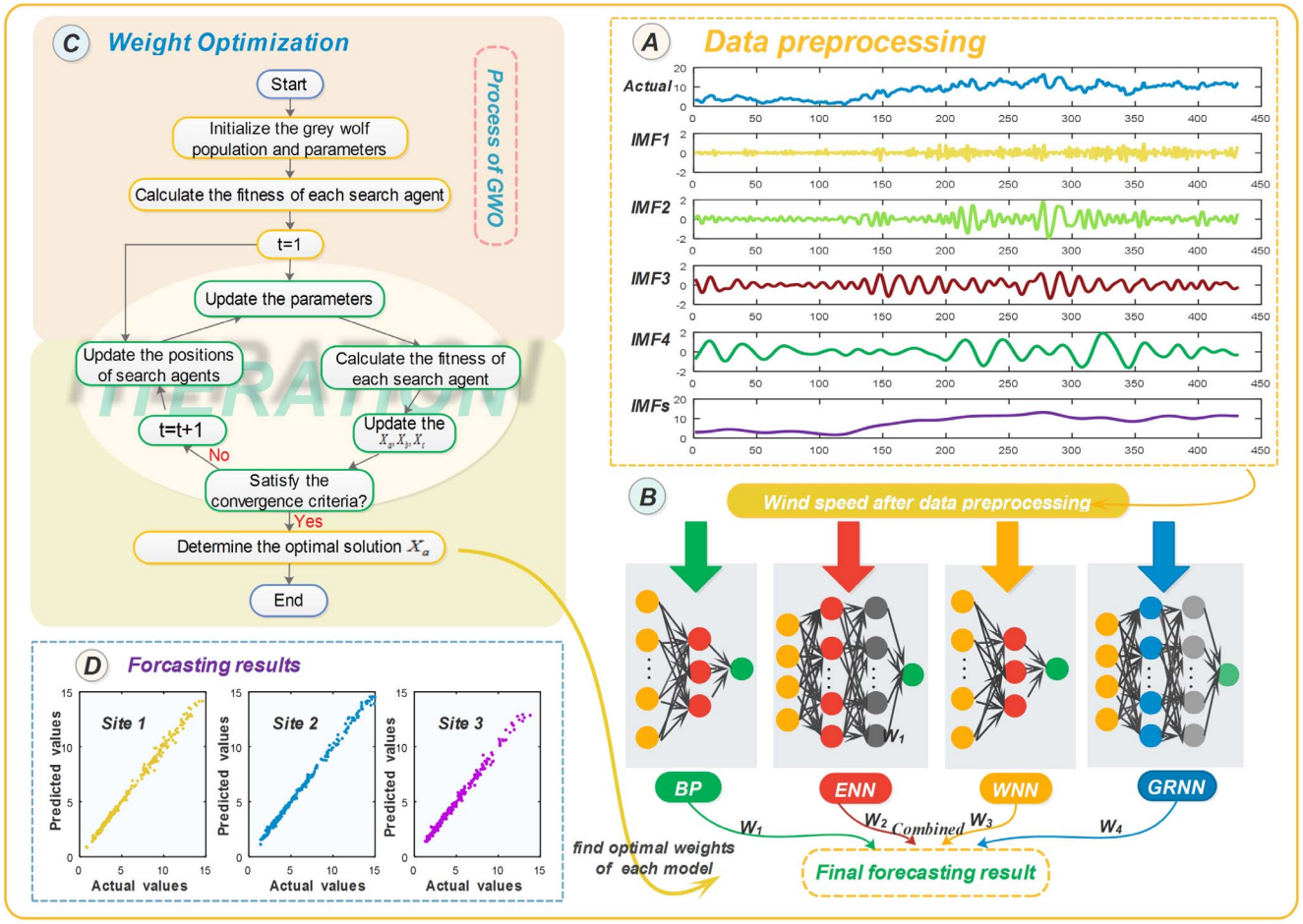

Fig. 1. Structure of the proposed ICEEMDAN-GWO combined model.

2.3. ICEEMDAN-GWO combined model

In this study, with the aforementioned methodology of the components of the model, a novel combined model based on the data preprocessing technique, novel deciding weight method, and several individual forecasting algorithms, is successfully developed to enhance the forecasting effectiveness. The main steps and flowchart of the proposed model are listed below and described in Fig. 1.

2.3.1. Stage 1: Data preprocessing

The ICEEMDAN technique, which is a novel data preprocessing technique, is utilized to decompose the original series and reconstruct the filtered time series so that the noise signals and the stochastic volatility are removed. Moreover, the processed wind speed data will be employed for the subsequent forecasting.

2.3.2. Stage 2: Forecasting of individual models

In this study, BPNN, ENN, WNN, and GRNN, which have high forecasting accuracy, are selected as the individual forecasting models to construct the developed combined model. Therefore, several individual models are used for forecasting the processed wind speed data, which are specifically shown in Fig. 1.

2.3.3. Stage 3: Construction of the combined model

In order to obtain the optimal weight coefficients of the individual models, a novel deciding weight method, based on the GWO algorithm and the leave-one-out strategy, is developed to obtain the best results. More specifically, we leave the final three days' data of the training set to determine the weight values of individual models. It is worth noting that the optimization algorithm would stop when it reaches the maximum number of iterations or the minimum value of the fitness function. Moreover, the iteration number and dimension of GWO is set as 400 and 4 respectively, and the weight coefficient of every model is between -2 and 2. Afterward, based on the obtained weight coefficient of the individual models, the forecasting results of such models are combined together to obtain the final wind speed forecasting result.

2.3.4. Stage 4: Wind speed forecasting

The developed combined model is employed in this study to forecast wind speed based on historical data. Multi-step forecasting is conducted to evaluate the performance of the combined model.

The multi-step ahead forecast is described as follows: define a time index h and a positive integer l as the forecast origin and forecast horizon, respectively. Assume that we are at time index h and are aimed at forecasting $\hat{y}_{h+l}$, where $l \geqslant 1$. Let $\hat{y}_h(l)$ be the forecast value of y_{h+l}; then, we call $\hat{y}_h(l)$ as the l-step ahead forecast of y_t at the forecast origin h. When $l = 1$, we call $\hat{y}_h(1)$ as the one-step ahead forecast of y_t at the forecast origin h [7].

3. Experiments and analysis

To prove the forecasting effectiveness of the developed combined model, experiments using three wind speed datasets collected from wind farm in the Shandong province of China are used as illustrative examples.

3.1. Datasets

In this study, three datasets are collected based on the wind speed data

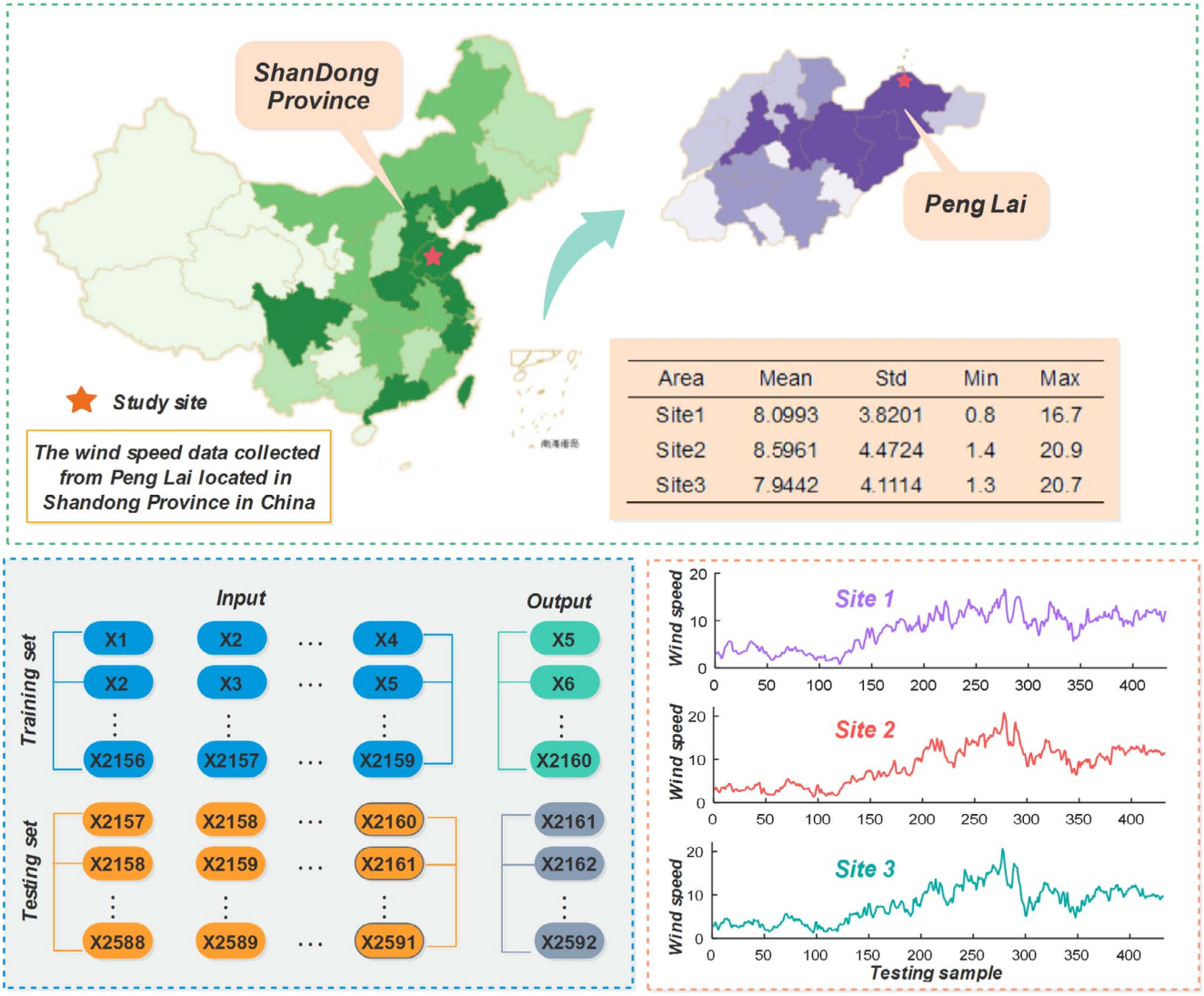

Fig. 2. Location of the Peng Lai wind farm and the data structure.

Table 1
Statistical indicators of the experimental samples for three sites.

Dataset	Samples	Numbers	Statistical Indicator (m/s)			
			Max	Min	Mean	Std.
Site 1	All samples	2592	20.0	0.8	7.0136	3.4131
	Training	2160	20.0	1.0	6.7964	3.2838
	Testing	432	16.7	0.8	8.0993	3.8201
Site 2	All samples	2592	20.9	0.6	7.3321	3.5967
	Training	2160	17.9	0.6	7.0793	3.3388
	Testing	432	20.9	1.4	8.5961	4.4724
Site 3	All samples	2592	20.7	0.6	7.0179	3.3938
	Training	2160	18.6	0.6	6.8327	3.2003
	Testing	432	20.7	1.3	7.9442	4.1114

from Shandong Province, which has one of China's largest wind farms with an installed capacity of 651,000 kW in 2015. Peng Lai (north latitude: 37°48′, east longitude: 120°45′), located in Yantai, Shandong province, possesses rich wind energy resources. A simple map of the study area is shown in Fig. 2. Moreover, some statistical indicators for wind speed data samplings of the three datasets, including minimum, maximum, mean, and standard deviation, are presented in Table 1 and Fig. 2.

To estimate the forecasting effectiveness of the novel proposed model, 10-min wind speed data are chosen from the wind farm in Peng Lai, which are sampled from May 1, 2011 to May 18, 2011, 18 days in all. Each dataset is classified into two sets: the training set and testing set. Specifically, the initial 15 days comprising 2160 data points, serve as the training sample of wind speed forecasting, and we leave the final three days' data of the training set to determine the weight values of the individual models. The remaining three days from May 16, 2011 to May 18, 2011 are the testing sample, which comprise 432 data points. The training set and the testing set both adopt the same rolling prediction mechanism. Moreover, one-step, two-step, three-step forecasting results are outputted, and the rolling prediction mechanism is used for wind speed prediction of three days. Fig. 2 shows the detailed data structure of the developed combined model.

3.2. Performance metrics

Many performance metrics are researched and applied to evaluate the forecasting effectiveness of different models. However, there is no general standard for the error evaluation of prediction models [54]. Therefore, multiple error metrics, namely mean absolute error (MAE), root mean square error (RMSE), mean absolute percent error (MAPE), and sum of squared errors (SSE), as presented in Table 2, are employed to assess the forecasting capacity of the proposed novel combined model in this study. The details of these four performance metrics are as follows:

MAE and RMSE are used to evaluate the average magnitude

Table 2
Four error metrics.

Metric	Definition	Equation
MAE	The mean absolute error of N forecasting results	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert F_i - A_i\rvert$
RMSE	The square root of the average of error squares	$RMSE = \sqrt{\frac{1}{N}\times\sum_{i=1}^{N}(F_i - A_i)^2}$
MAPE	The average of N absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{A_i - F_i}{A_i}\right\rvert \times 100\%$
SSE	The sum of error squares	$SSE = \sum_{i=1}^{N}(A_i - F_i)^2$

between the forecasted and actual value, to avoid the positive and negative forecasting errors from canceling each other out. MAPE, which is the average of absolute error, is the most widely implemented index and is used to reflect the validity and reliability of the proposed novel model. On the other hand, SSE is used to express the total forecasting error of the model. For these four metrics, the smaller the index value, the better the model performance. To explain the performance metrics more clearly, the definition and specific formula of the four error metrics are listed in Table 2. Here A_i and F_i indicate the actual and predicted values respectively, and N is the sample size.

3.3. Testing methods

The Diebold–Mariano (DM) test is conducted and the forecasting effectiveness is calculated to further verify the performance of the proposed combined model from a statistical perspective.

3.3.1. Diebold Mariano test

The DM test, proposed by Diebold and Marino [55], is applied to compare the forecasting efficiency of the developed model with the other models. The principal steps are presented as follows:

$$H_0: E[F(e_t^1)] = E[F(e_t^2)] \tag{14}$$

$$H_1: E[F(e_t^1)] \neq E[F(e_t^2)] \tag{15}$$

The above are the null hypothesis and alternative hypothesis of the DM test, where e_t^1 and e_t^2 are the forecasting errors between actual values and forecasted values of the different models, and the loss function F is the function of the forecasting errors.

$$\overline{d} = \frac{1}{L}\sum_{t=1}^{L}[F(e_t^1) - F(e_t^2)] \tag{16}$$

The sample mean loss differential difference $\overline{d}$, which is involved during the DM test, is the average of the difference of the loss function of two models, where L is the length of forecasting values.

$$DM = \frac{\overline{d}}{\sqrt{2\pi\widehat{f}_d(0)/L}} \rightarrow N(0,1) \tag{17}$$

As shown in the above formula, the DM value based on statistic $\overline{d}$, converges to the normal distribution, where $\widehat{f}_d(0)$ is the zero spectral density, and $2\pi\widehat{f}_d(0)$ is a consistent estimate of the asymptotic variance of $\sqrt{T}\overline{d}$. Based on the above theory, the absolute value of the calculated DM value $|DM|$ is compared with $|Z_{\alpha/2}|$, which we could obtain in the standard normal distribution table. If $|DM|$ is greater than $|Z_{\alpha/2}|$, we can reject the null hypothesis and it is deemed that the difference between the prediction abilities of two models is significant. For example, if $|DM| > 1.96$, the null hypothesis is rejected at a 5% level otherwise, if $|DM| \leqslant 1.96$, we cannot reject the null hypothesis.

3.3.2. Forecasting effectiveness

The forecasting effectiveness is also employed to examine the prediction efficiency of the developed model. This could be calculated not only through the sum of forecasting error squares but also through the mean and mean squared deviation of the forecasting accuracy [56]. The following will provide the principal ideas of forecasting effectiveness.

The k_{th} order forecasting effectiveness unit is calculated by $m^k = \sum_{i=1}^{n} Q_i A_i^k$, where A_i is the forecasting accuracy at time i, Q_i is called the discrete probability distribution, and $\sum_{i=1}^{n} Q_i = 1, Q_i > 0$. An exception is that Q_i is defined as $Q_i = 1/n, i = 1,2,...,n$, when prior information of Q_i cannot be known in some cases.

Afterward, the k_{th} order forecasting effectiveness is measured by $H(m^1,m^2,...,m^k)$, where H is a continuous function of a certain k unit. If $H(x) = x$ is a continuous function of one variable, the first-order forecasting effectiveness is defined as $H(m^1) = m^1$. Then if $H(x,y) = x(1-\sqrt{y-x^2})$ is a continuous function of two variables, $H(m^1,m^2) = m^1(1-\sqrt{m^2-(m^1)^2})$ is called the second-order forecasting effectiveness.

3.4. Different experiments and result analysis

In this section, we present three different experiments, which are conducted to evaluate the forecasting accuracy of the proposed combined model.

3.4.1. Experiment I: Comparison with other ICEEMDAN-based models

This experiment is set to validate the accuracy of the proposed novel model compared with other four ICEEMDAN-based models, namely ICEEMDAN-BPNN, ICEEMDAN-ENN, ICEEMDAN-WNN, and ICEEMDAN-GRNN. The specific forecasting results are indicated in Table 3.

Table 3
Comparison of forecasting performances of the combined model and other ICEEMDAN-based models.

Dataset	Model	MAE			RMSE			MAPE (%)			SSE		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	ICEEMDAN-BP	0.2371	0.2502	0.3221	0.3201	0.3307	0.4184	3.33	3.90	5.42	44.2743	47.2522	75.6808
	ICEEMDAN-ENN	0.3588	0.3741	0.4369	0.4775	0.4959	0.5798	5.48	6.05	7.35	99.4013	107.0033	147.2832
	ICEEMDAN-WNN	0.3598	0.4343	0.5766	0.4694	0.5563	0.7169	5.58	7.46	10.27	99.6408	137.9602	229.1793
	ICEEMDAN-GRNN	0.6217	0.6430	0.6654	0.8011	0.8267	0.8518	10.27	10.82	11.38	277.2242	295.2666	313.4750
	Proposed Model	0.2201	0.2401	0.2886	0.2994	0.3224	0.3849	3.02	3.57	4.43	38.7300	44.9290	64.1497
Site 2	ICEEMDAN-BP	0.2125	0.2498	0.3526	0.2991	0.3748	0.5464	3.04	3.76	5.52	38.6606	61.0356	130.4953
	ICEEMDAN-ENN	0.3370	0.3802	0.4635	0.4765	0.5987	0.7673	5.11	5.98	7.50	100.0171	158.6029	261.6464
	ICEEMDAN-WNN	0.7000	0.8000	1.0000	0.9000	1.1000	1.3000	10.00	10.00	10.00	1064.8000	1101.4000	1259.7000
	ICEEMDAN-GRNN	0.5785	0.5936	0.6155	0.7926	0.8164	0.8479	9.39	9.93	10.52	271.4169	287.9231	310.6125
	Proposed Model	0.1931	0.2382	0.3125	0.2730	0.3729	0.5267	2.71	3.50	4.87	32.2279	60.9766	123.4220
Site 3	ICEEMDAN-BP	0.2526	0.2719	0.3498	0.3481	0.3718	0.4909	3.85	4.27	5.66	52.3479	59.8225	104.5839
	ICEEMDAN-ENN	0.3925	0.4011	0.4524	0.5372	0.5516	0.6327	6.25	6.69	7.79	126.8008	133.2584	175.2341
	ICEEMDAN-WNN	0.3556	0.4700	0.6663	0.5531	0.7174	0.9854	5.13	6.87	9.79	150.7141	243.8239	450.0879
	ICEEMDAN-GRNN	0.6258	0.6461	0.6731	0.8426	0.8708	0.9113	10.52	11.07	11.65	306.7293	327.6156	358.7737
	Proposed Model	0.2409	0.2620	0.3123	0.3288	0.3559	0.4435	3.58	4.12	5.13	46.7293	54.7558	85.3247

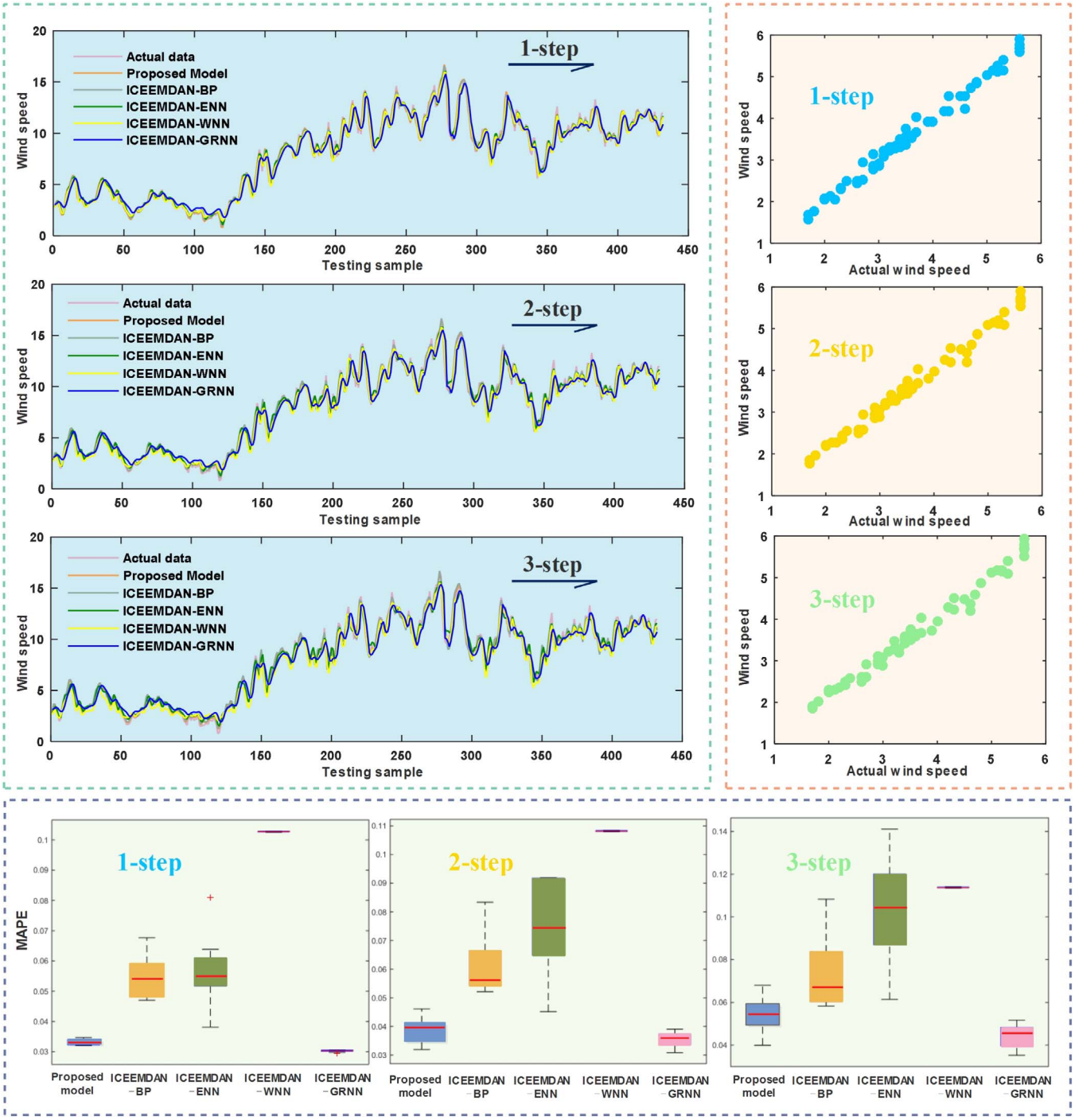

Fig. 3. Comparison of the multi-step forecasting performance of Experiment I for Site1.

More experiment details are described below:

(a) For Site 1, in the one-step forecasting, the proposed model has the best MAE, RMSE, MAPE, and SSE at 0.2201, 0.2994, 3.02%, and 38.7300 respectively. Next, the models from highest to lowest based on the forecasting accuracy are ICEEMDAN-BPNN, ICEEMDAN-ENN, ICEEMDAN-WNN, and ICEEMDAN-GRNN with MAPE values of 3.33%, 5.48%, 5.58%, and 10.27% respectively. When the model forecasting is two-step, the ICEEMDAN-GWO combined model has the best forecasting performance result with a MAPE value of 3.57%. In the three-step forecasting, the developed model with a MAPE value of 4.43% still has the highest prediction accuracy among the four ICEEMDAN-based models. Fig. 3 shows the comparison of the 1-step, 2-step, and 3-step forecasting performances of this Experiment I for Site 1. It could also be observed that the CEEMDAN-GWO combined model is still the most precise and effective forecasting model among the 3-step ahead forecasting models.

(b) For Site 2, when the model forecasting is one-step, the prediction accuracy of the ICEEMDAN-GWO combined model, which has the lowest MAE, RMSE, MAPE, and SSE values of 0.1931, 0.2730, 2.71%, and 32.2279 respectively, is still superior compared to the other models. In addition, based on the model results obtained from two-step forecasting and three-step forecasting, the proposed ICEEMDAN-GWO combined model with MAPE values of 3.50% and 4.87% respectively is more accurate than the other combined

Table 4
Comparison of forecasting performances of the combined model and models using different data preprocessing methods.

Dataset	Model	MAE			RMSE			MAPE (%)			SSE		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	EMD	0.3643	0.3780	0.4197	0.4751	0.494	0.5517	5.14	5.47	6.26	97.5358	105.4688	132.0605
	CEEMDAN	0.2677	0.2855	0.3232	0.3523	0.3761	0.4234	3.87	4.26	5.12	53.6707	61.3601	78.1900
	SSA	0.3812	0.3939	0.4187	0.4736	0.4928	0.5289	6.02	6.41	6.97	96.9230	104.9523	121.0303
	Proposed Model	0.2201	0.2401	0.2886	0.2994	0.3224	0.3849	3.02	3.57	4.43	38.7300	44.9290	64.1497
Site 2	EMD	0.3835	0.4041	0.4423	0.5239	0.5654	0.6301	5.36	5.64	6.39	118.5634	138.2176	172.1287
	CEEMDAN	0.2817	0.2973	0.3753	0.3952	0.4312	0.5865	4.03	4.38	5.78	68.8595	80.9019	149.4016
	SSA	0.3785	0.3988	0.4439	0.4991	0.5355	0.6386	5.83	6.08	6.67	107.9595	124.8311	182.3464
	Proposed Model	0.1931	0.2382	0.3125	0.2730	0.3729	0.5267	2.71	3.50	4.87	32.2279	60.9766	123.4220
Site 3	EMD	0.4474	0.4824	0.5006	0.5923	0.6577	0.6959	6.92	7.39	7.65	151.9644	190.5076	210.2592
	CEEMDAN	0.2993	0.3118	0.3588	0.4027	0.4198	0.4953	4.59	4.86	5.73	70.1346	76.2955	106.5226
	SSA	0.3737	0.3980	0.4464	0.4886	0.5142	0.6007	5.86	6.32	7.13	103.4215	114.3131	156.5117
	Proposed Model	0.2409	0.2620	0.3123	0.3288	0.3559	0.4435	3.58	4.12	5.13	46.7293	54.7558	85.3247

methods.

(c) For Site 3, the ICEEMDAN-GWO combined model with MAPE values of 3.58%, 4.12%, and 5.13% respectively from the one-step forecasting to three-step forecasting is still the most precise and effective forecasting model. Based on these four error indicators, it is evident that the ICEEMDAN-BPNN model is the second most precise model relative to the remaining three combined models. Meanwhile, the ICEEMDAN-GRNN model, with MAPE values of 10.52%, 11.07%, and 11.65% respectively, has the worst forecasting performance among the three-step model predictions.

Remark. Based on this experiment, the proposed ICEEMDAN-GWO combined model achieves the highest forecasting accuracy, with average MAPE values of 3.10%, 3.73%, and 4.81% for the one-step forecasting to the three-step forecasting respectively. Moreover, it could be concluded that the developed model is superior compared to other ICEEMDAN-based models based on the performance metrics, which further validate the efficiency of this model.

3.4.2. Experiment II: Testing with models using different data preprocessing methods

This experiment demonstrates the forecasting performance of the wind speed time series by comparing the ICEEMDAN-GWO combined model with models using different data preprocessing methods, namely EMD, CEEMDAN, and SSA. The comparison results are listed in Table 4. From the table, we can draw the following conclusions:

(a) For the three steps wind speed forecasting of Site 1, the proposed ICEEMDAN-GWO combined model achieves the highest accuracy. In comparison, the model after pretreatment of CEEMDAN is ranked as the second most effective model among the other three data preprocessing methods, with MAPE values of 3.87%, 4.26%, and 5.12% respectively from the one-step forecasting to three-step forecasting. In addition, for the two models after preprocessing by EMD and SSA, the MAPE values of EMD decrease by 0.88%, 0.94%, and 0.71% respectively compared with SSA.

(b) For Site 2, according to the evaluation criteria, the proposed combined model still outperforms the remaining models in the one-step forecasting. Moreover, the MAPE values of the models using EMD, CEEMDAN, and SSA are 2.65%, 1.32%, and 3.12% higher than that of the proposed model, respectively. When the model forecasting is two-step and three-step, among the remaining three methods, the CEEMDAN relatively has a better forecasting performance, with MAPE values of 4.38% and 5.78% respectively. Meanwhile, SSA achieves the worst prediction accuracy of the two-step and three-step forecasting, with a maximum MAPE of 6.08% and 6.67% respectively. Fig. 4 shows a comparison of the 1-step, 2-step, and 3-step forecasting performance of Experiment II for Site 2. It could be concluded that the proposed combined model achieves the highest accuracy compared to the models using other data preprocessing methods in the three steps forecasting.

(c) For Site 3, in all forecasting steps, the performance metric values calculated from the proposed ICEEMDAN-GWO combined model are evidently lower than those calculated from the models using other data preprocessing methods. In addition, the model after pretreatment of CEEMDAN is always superior relative to the models using EMD and SSA, with a forecasting accuracy of 4.59%, 4.86%, and 5.73% from the one-step prediction to three-step prediction respectively. Furthermore, SSA has a better prediction performance compared to EMD, whose MAPE values decrease by 1.06%, 1.07%, and 0.52% respectively in comparison with SSA.

Remark. The ICEEMDAN-GWO combined model still outperforms the models using the other data preprocessing methods, namely EMD, CEEMDAN, and SSA. In addition, the model after pretreatment of CEEMDAN is ranked as the second most effective model among the other models, with average MAPE values of 4.16%, 4.50%, and 5.54% for the one-step forecasting to the three-step forecasting respectively. According to the forecasting results of the three sites, the prediction accuracy of EMD is similar to that of SSA.

3.4.3. Experiment III: Comparison with some classic individual models

This experiment shows a comparison between the proposed model and some classic individual models, namely, BP, ENN, WNN, GRNN, ARIMA, AR, and the naïve benchmark model. The specific forecasting results are listed in Table 5. More experiment details are given below:

(a) For Site 1, when the forecasting is one-step, the ICEEMDAN-GWO combined model achieves the best forecasting performance with a MAPE value of 3.02%. By contrast, the other individual models have fairly lower MAPE values of 9.30%, 9.61%, 9.57%, 11.93%, 8.70%, 8.56%, and 8.56%, respectively. In the two-step and three-step forecasting, the developed combined model is more effective than the other methods for wind speed forecasting. According to the obtained MAE, RMSE, MAPE, and SSE data, the forecasting performances of ENN, BP, and ARIMA are respectively similar to the prediction results for WNN, GRNN, and AR.

(b) For Site 2, in all forecasting steps, the developed combined model outperforms the other individual models based on the four performance metrics, with MAPE values of 2.71%, 3.50%, and 4.87% from the one-step forecasting to three-step forecasting. Based on the MAPE values in Table 5, among the remaining seven models, ARIMA is ranked as the second most effective model in the one-step and two-step forecasting, with lower MAPE values at 8.12% and 8.18% respectively. Meanwhile, the naïve benchmark has the highest MAPE values at 12.37% and 14.69% in the two-step and

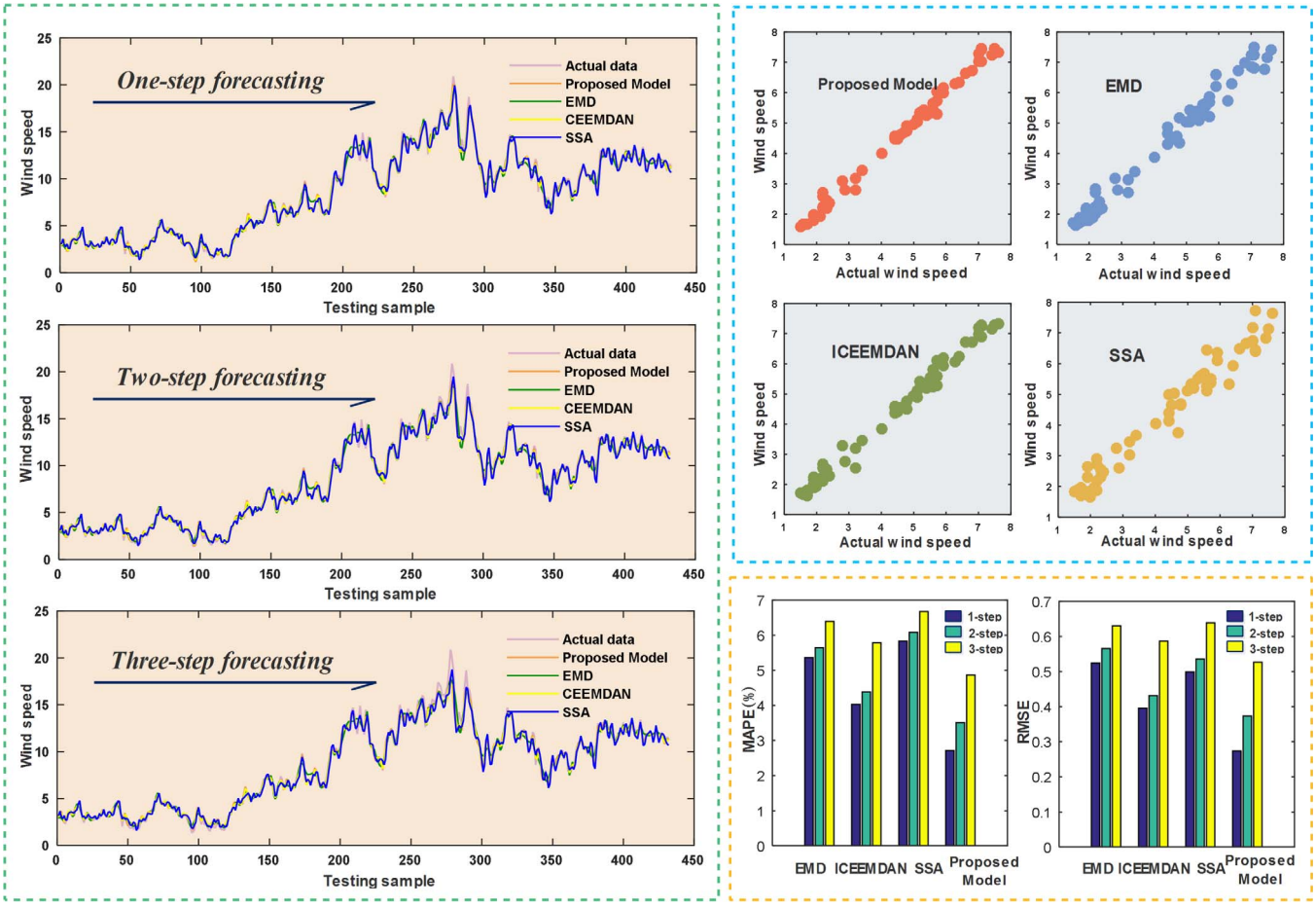

Fig. 4. Comparison of the multi-step forecasting performance of Experiment II for Site 2.

Table 5
Comparison of forecasting performances of the combined model and some classic individual models.

Dataset	Model	MAE			RMSE			MAPE (%)			SSE		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site 1	BP	0.6282	0.7526	0.7716	0.8502	0.9769	0.9966	9.30	12.50	13.00	312.5145	412.3125	429.0577
	ENN	0.6254	0.6506	0.6789	0.8354	0.8560	0.8831	9.61	10.53	11.45	301.5374	316.6648	337.1630
	WNN	0.6353	0.6627	0.6899	0.8516	0.8802	0.9067	9.57	10.37	11.20	313.4019	334.8447	355.3377
	GRNN	0.7275	0.7526	0.7716	0.9459	0.9769	0.9966	11.93	12.50	13.00	386.5572	412.3125	429.0577
	ARIMA	0.6084	0.6154	0.6223	0.8299	0.8365	0.8422	8.70	8.83	8.99	297.5395	302.2797	306.4101
	AR	0.5999	0.6000	0.6000	0.8194	0.8195	0.8195	8.56	8.56	8.56	290.0673	290.1036	290.1186
	Naïve Benchmark	0.5998	0.8889	1.1072	0.8197	1.1977	1.4623	8.56	13.07	16.19	290.2500	619.7400	923.7100
	Proposed Model	0.2201	0.2401	0.2886	0.2994	0.3224	0.3849	3.02	3.57	4.43	38.7300	44.9290	64.1497
Site 2	BP	0.6082	0.6387	0.6654	0.8410	0.8847	0.9215	8.94	9.84	10.66	306.1378	339.4336	369.0452
	ENN	0.6045	0.6308	0.6539	0.8321	0.8714	0.9023	8.93	9.62	10.30	299.6225	329.1604	353.0112
	WNN	0.6252	0.6489	0.6717	0.8661	0.8931	0.9248	9.37	10.18	10.88	324.7802	345.4318	370.7026
	GRNN	0.6859	0.7084	0.7250	0.9312	0.9684	0.9892	10.74	11.31	11.79	374.6226	405.1098	422.7468
	ARIMA	0.5772	0.5834	0.5965	0.7974	0.8087	0.8307	8.12	8.18	8.32	274.6553	282.5585	298.1307
	AR	0.5813	0.5815	0.5819	0.7925	0.7927	0.7934	8.21	8.22	8.22	271.3285	271.4575	271.9668
	Naïve Benchmark	0.5813	0.8701	1.0486	0.7935	1.2045	1.4406	8.20	12.37	14.69	271.9900	626.7500	896.4800
	Proposed Model	0.1931	0.2382	0.3125	0.2730	0.3729	0.5267	2.71	3.50	4.87	32.2279	60.9766	123.4220
Site 3	BP	0.6727	0.6924	0.7119	0.9329	0.9657	0.9987	10.38	10.91	11.41	376.1462	403.3586	432.0070
	ENN	0.6734	0.6931	0.7158	0.9255	0.9549	0.9858	10.59	11.13	11.73	370.2785	394.1883	420.5413
	WNN	0.6713	0.6927	0.7132	0.9261	0.9587	0.9939	10.38	10.90	11.38	370.8403	397.3744	427.2921
	GRNN	0.7513	0.7746	0.7894	1.0372	1.0789	1.1057	12.06	12.58	13.00	464.7722	502.8407	528.1808
	ARIMA	0.6576	0.6566	0.6579	0.8937	0.8904	0.8908	9.95	9.99	10.07	345.0397	342.4899	342.8196
	AR	0.6564	0.6570	0.6576	0.8958	0.8971	0.8987	9.84	9.85	9.85	346.6874	347.6766	348.9429
	Naïve Benchmark	0.6549	0.9748	1.1613	0.8968	1.3358	1.5889	9.82	14.58	17.03	347.4100	770.8100	1090.6900
	Proposed Model	0.2409	0.2620	0.3123	0.3288	0.3559	0.4435	3.58	4.12	5.13	46.7293	54.7558	85.3247

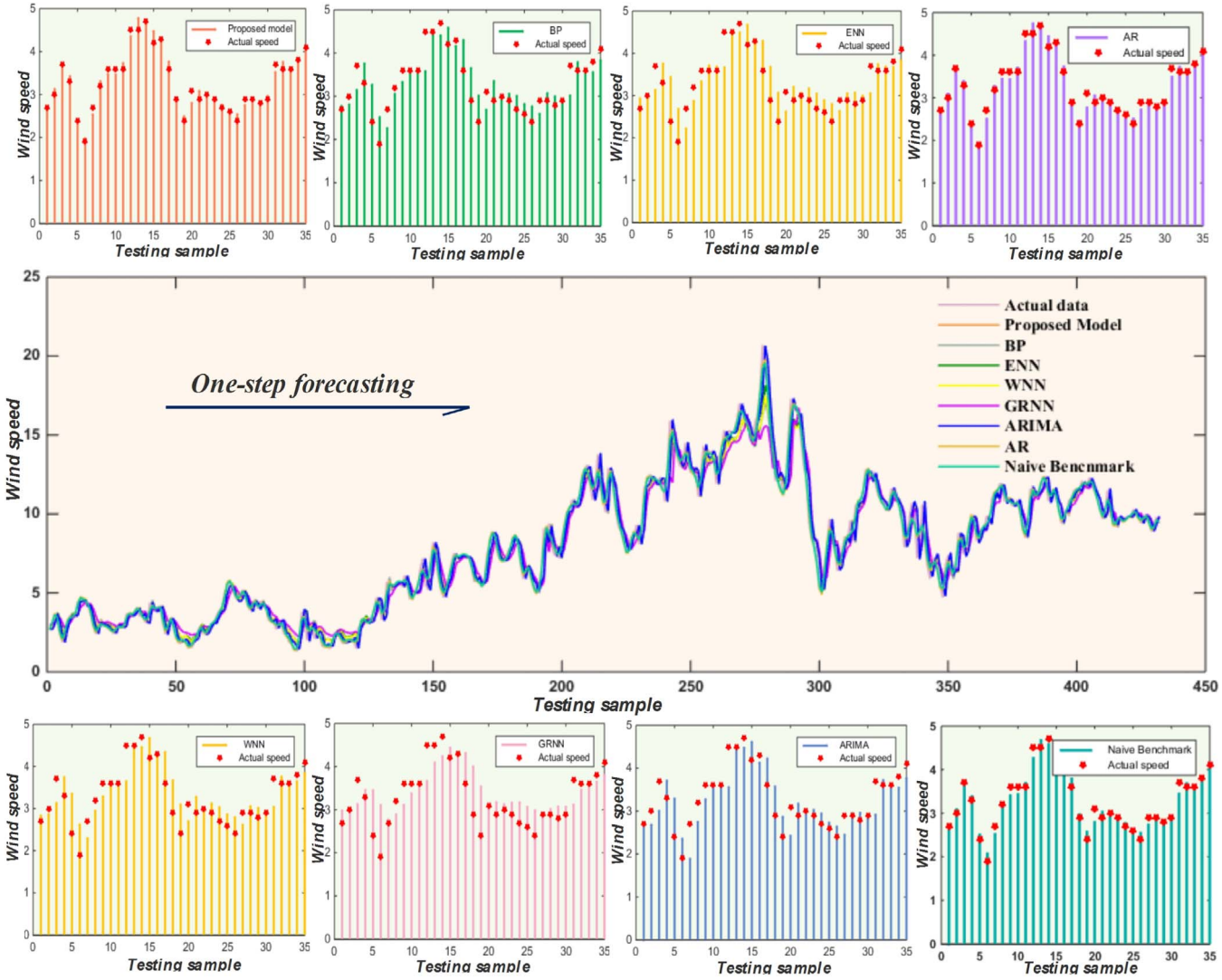

Fig. 5. Comparison of the one-step forecasting performance of Experiment III for Site 3.

three-step forecasting respectively, indicating its poor forecasting accuracy.

(c) For Site 3, when the forecasting is one-step, the developed combined model has the lowest MAE, RMSE, MAPE, and SSE at 0.2409, 0.3288, 3.58%, and 46.7293 respectively, indicating its excellent forecasting accuracy. In the two-step and three-step forecasting, the performance metric values of the proposed model are significantly lower than those of the other individual models; thus, it could be concluded that the prediction accuracy of the developed combined method is still superior compared to the other models. Fig. 5 shows a comparison of the forecasting performance of the proposed combined model in the one-step prediction for Site 3. It could be observed that the novel ICEEMDAN-GWO combined model outperforms the other individual models in the one-step forecasting.

Remark. The evaluation index values calculated from the proposed combined model are significantly lower than those calculated from any of the classic individual models for the one-step forecasting to the three-step forecasting. Therefore, the experiment results demonstrate that the developed ICEEMDAN-GWO combined model is definitely superior compared to all other conventional individual models in multi-step forecasting.

4. Discussion

This section presents an insightful discussion of the experiment results, namely the effectiveness of the proposed model, performance of the employed optimization algorithm, improvements of the proposed combined model, combined strategy, computation time, and practical applications in power systems. The concrete details are as follows.

4.1. Effectiveness of the proposed model

The DM test is first employed to examine the effectiveness of the developed model. All of the other models are compared with the ICEEMDAN-GWO combined model. According to the basic idea of the DM test proposed, the null hypothesis is that there are no significant differences between the forecasting performances of two models, while the alternative hypothesis is that the differences between the forecasting performances of two models are significant. The average values of the DM test values for the three sites are presented in Table 6.

Table 6 indicates that the proposed combined model is different from the other ICEEMDAN-based models at a 1% significance level in multi-step forecasting. Moreover, for the comparison results between the ICEEMDAN-GWO combined model and the models using different data preprocessing methods, the smallest value of $|DM|$ is 2.1017; thus, the null hypothesis could be rejected at a 5% significance level. In

addition, for some classic individual models, all values are much larger than the upper limits at a 1% significance level; consequently, the differences between the proposed model and individual models are significant at the 1% significance level. Therefore, the proposed ICEEMDAN-GWO combined model significantly outperforms the other models.

To further evaluate the developed combined model, the forecasting effectiveness is applied in this study. The greater the forecasting effectiveness, the more effective the forecasting performance of the model. The average values of the three sites for the one-step forecasting to the three-step forecasting are listed in Table 7.

According to the specific details in Table 7, it could be observed that the forecasting effectiveness values of the developed combined model are always higher than those of the other models in the three steps forecasting whether it is 1-order or 2-order. Therefore, the developed ICEEMDAN-GWO combined model is significantly superior compared to all other models. Furthermore, the results in the table indicate that the values of the forecasting effectiveness for the ICEEMDAN-BP model and the model after pretreatment of CEEMDAN are slightly lower than that of the proposed model, and they are ranked as the second most effective models among the other comparative models.

Table 6
DM test of different models.

Model	1-step	2-step	3-step
ICEEMDAN-BP	3.9467*	3.2175*	5.2455*
ICEEMDAN-ENN	7.9542*	6.7756*	5.9841*
ICEEMDAN-WNN	7.4236*	7.6884*	8.7576*
ICEEMDAN-GRNN	9.6320*	9.4442*	9.2623*
EMD	9.9582*	8.7194*	6.3067*
CEEMDAN	5.8313*	3.1556*	2.1017**
SSA	8.8913*	7.9799*	5.9326*
BP	9.4753*	9.2575*	9.0217*
ENN	9.8975*	9.7530*	9.4019*
WNN	9.7969*	9.6998*	9.3123*
GRNN	9.8331*	9.4927*	9.4493*
ARIMA	9.9311*	9.5065*	8.5373*
AR	10.1689*	9.7254*	8.5900*
Naïve Benchmark	10.1406*	10.1860*	10.3263*

* 1% significance level.
** 5% significance level.

Table 7
Forecasting effectiveness of different models.

Model	1-step		2-step		3-step	
	1-order	2-order	1-order	2-order	1-order	2-order
Proposed Model	0.9694	0.9400	0.9638	0.9246	0.9546	0.8992
ICEEMDAN-BP	0.9663	0.9317	0.9611	0.9196	0.9459	0.8853
ICEEMDAN-ENN	0.9465	0.8894	0.9433	0.8754	0.9334	0.8471
ICEEMDAN-WNN	0.9379	0.8719	0.9271	0.8536	0.9040	0.8167
ICEEMDAN-GRNN	0.9002	0.7948	0.8949	0.7811	0.8894	0.7674
EMD	0.9426	0.8909	0.9401	0.8809	0.9359	0.8686
CEEMDAN	0.9636	0.9282	0.9608	0.9198	0.9517	0.8938
SSA	0.9422	0.8852	0.9385	0.8760	0.9330	0.8618
BP	0.9063	0.8173	0.8994	0.8004	0.8928	0.7833
ENN	0.9050	0.8123	0.8984	0.7952	0.8918	0.7787
WNN	0.9044	0.8125	0.8977	0.7954	0.8915	0.7793
GRNN	0.8855	0.7711	0.8802	0.7586	0.8757	0.7473
ARIMA	0.9109	0.8275	0.9101	0.8265	0.9089	0.8244
AR	0.9015	0.8188	0.9015	0.8188	0.9015	0.8187
Naïve Benchmark	0.9169	0.8378	0.8629	0.7528	0.8388	0.7159

4.2. Performance of the employed optimization algorithm

In this study, four typical benchmark functions, namely Sphere, Rosenbrock, Rastrigin, and Ackley, are employed to measure and validate the performance of the GWO algorithm. The variable domain and function formula of the benchmark functions are presented in Table 8.

Furthermore, classic optimization algorithms, i.e., particle swarm optimization (PSO) and bat algorithm (BA), and relatively new optimization algorithms, i.e., dragonfly algorithm (DA) and ant lion optimizer (ALO), are selected to investigate the performance of the GWO optimization algorithm. Fig. 6 shows the evolution of the fitness among GWO and other comparative algorithms with different dimensions. In addition, when the dimension is 30, the test results of GWO and other algorithms of 50 trials are presented in Table 9.

Fig. 6 and Table 9 show the following:

(a) For the Sphere function, regardless of the dimension, GWO, whose optimization results are evidently superior to the other algorithms, has the best searching ability. In particular, when the dimension is 30, the maximum value, minimum value, average value, and standard deviation of the test results of GWO are 2.35E-29, 1.20E-31, 3.21E-30, and 4.11E-30 respectively.
(b) For the Rosenbrock function, although the optimization iteration curves with different dimensions are varied and distinct, GWO has the fastest convergence speed and the best searching result compared to the other algorithms. When the dimension is 30, the maximum value, average value, and standard deviation of 50 trials of GWO have the lowest fitness values, namely 28.8319, 27.7147, and 0.7173 respectively compared to the other optimization algorithms. The minimum value of BA is lower than that of GWO, which also indicates that GWO has a more stable and effective searching result.
(c) For the Rastrigin function, the testing and validation results of 50 trials demonstrate that GWO still obtains the best searching result with different dimensions. With the dimension of 30, the maximum value, minimum value, average value, and standard deviation of the validation results of GWO, namely, 12.6979, 0.0000, 2.8997, and 3.6629 respectively, are all much smaller than those of the other algorithms.
(d) For the Ackley function, the optimization iteration curves with different dimensions make no difference, and GWO has the best searching result compared to the other optimization algorithms. When the dimension is 30, the maximum value, minimum value, average value, and standard deviation of the validation results of GWO, namely, 3.86E-04, 5.25E-05, 1.65E-04, and 7.26E-05, all have the lowest fitness values. Furthermore, the classic

Table 8
Benchmark functions for testing of the GWO algorithm.

Function name	Variable domain	Optimum value	Test function
Sphere	[−5.12, 5.12]	0	$\sum_{i=1}^{d} x_i^2$
Rosenbrock	[−2.084, 2.084]	0	$\sum_{i=1}^{d-1} [100(x_i^2-x_{i+1})^2 + (x_i-1)^2]$
Rastrigin	[−5.12, 5.12]	0	$\sum_{i=1}^{d} (x_i^2-10(2\pi x_i) + 10)$
Ackley	[−32.768, 32.768]	0	$-20\exp\left[-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i}\right]-\exp\left[\frac{1}{d}\sum_{i=1}^{d}\cos(2\pi x_i)\right] + 20 + e$

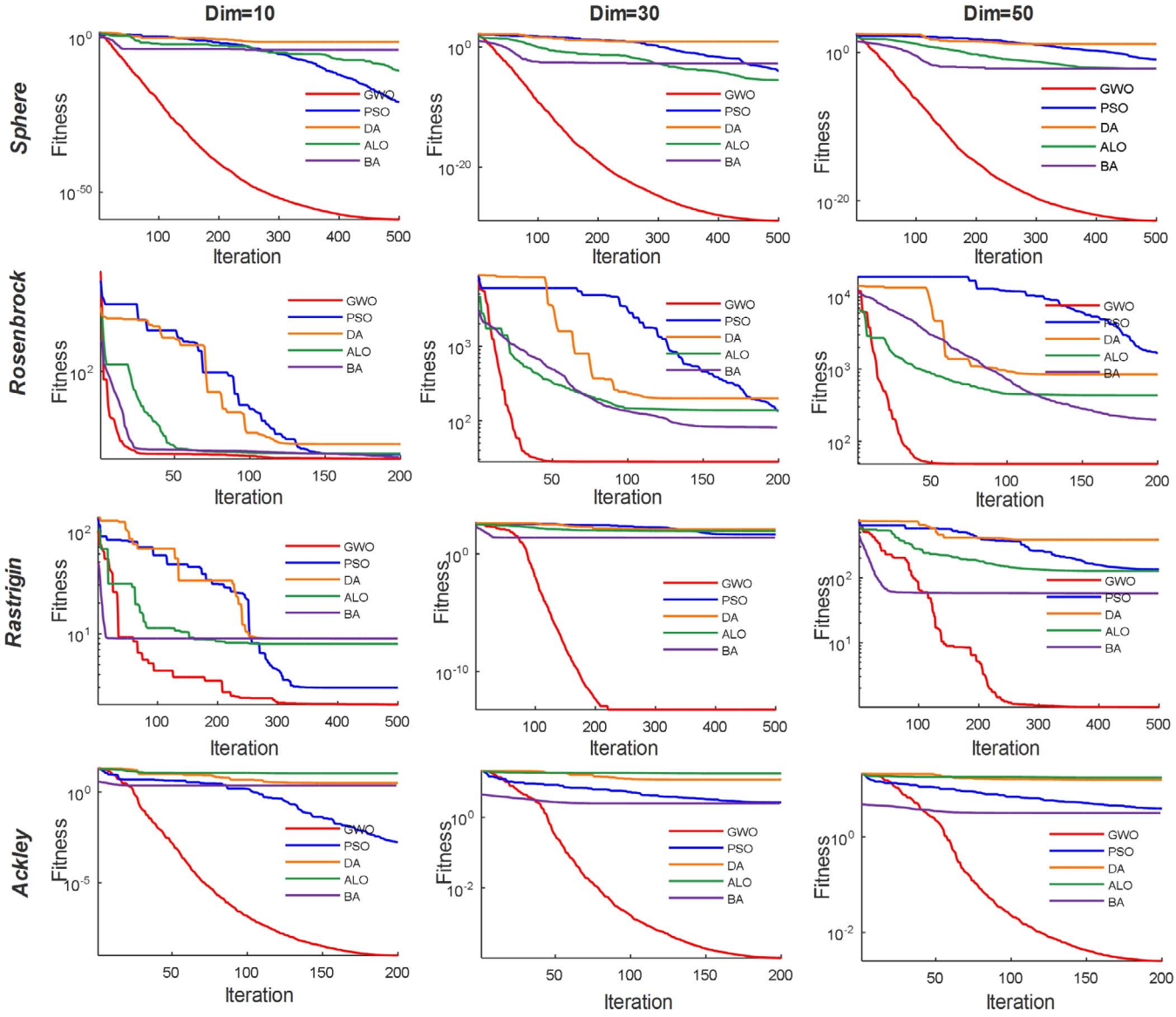

Fig. 6. Test results of GWO and other algorithms.

optimization algorithms obtain better validation results than the relatively new optimization algorithms for this function.

Remark. The test results with different dimensions indicate that GWO has the best searching result compared to the other optimization algorithms. Specifically, the maximum value, minimum value, average value, and standard deviation of the validation results of 50 trials of GWO are almost all much lower than those of the other algorithms. Therefore, it could be concluded that the searching performance and validation results of GWO are superior compared to the other algorithms.

4.3. Improvements of the proposed combined model

In this study, an improvement percentage of the MAPE criteria, which can be used to conduct a comprehensive analysis of the proposed combined model, is adopted. It is defined as

$$P_{MAPE} = \left| \frac{MAPE_1 - MAPE_2}{MAPE_1} \right| \tag{18}$$

To evaluate the effectiveness of the developed model, the forecasting results of the presented combined model are compared to other ICEEMDAN-based models, models using different data preprocessing methods, and some classic individual models. Table 10 presents the improvement percentages of MAPE for the proposed model and other forecasting models.

According to the results indicated in Table 10, some discussion and analysis on the forecasting performance of the proposed combined model are presented as follows:

(a) Compared with other ICEEMDAN-based models, it could be observed that the improvements on the forecasting effectiveness of the proposed model are fairly evident. For instance, the forecasting accuracy of the multi-step forecasting improves on average by 8.95%, 6.23%, and 13.15% for the ICEEMDAN-BP model, which has the best forecasting performance among the four models.
(b) Among the three models using different data preprocessing methods, the maximum improvement percentages of MAPE for the three steps forecasting are 47.39%, 40.51%, and 30.51%, while the minimum are 25.56%, 17.18%, and 13.25%, respectively. Therefore, the forecasting performance of the developed combined model is still superior compared to the other models.
(c) Relative to some classic individual models, the developed combined

Table 9
Test results of 50 trials of GWO and other algorithms.

Function name	Algorithm	Max value	Min value	Average value	Standard deviation
Sphere	PSO	1.08E−03	2.98E−06	1.40E−04	1.84E−04
	BA	2.52E−03	1.17E−03	1.73E−03	2.45E−04
	DA	12.2582	0.4494	5.4116	2.8624
	ALO	2.01E−05	4.49E−07	3.69E−06	3.81E−06
	GWO	**2.35E−29**	**1.20E−31**	**3.21E−30**	**4.11E−30**
Rosenbrock	PSO	669.9580	71.6239	226.5409	105.8788
	BA	98.3310	**25.2559**	34.8920	16.2416
	DA	685.0800	74.5784	334.7048	135.0048
	ALO	281.9166	66.4347	148.8041	57.2403
	GWO	**28.8319**	26.1437	**27.7147**	**0.7173**
Rastrigin	PSO	108.9154	24.7936	53.4764	15.4688
	BA	42.0609	19.2118	29.2445	5.5271
	DA	263.5504	81.7965	159.9952	39.1623
	ALO	138.2993	39.8004	76.8787	20.8561
	GWO	**12.6979**	**0.0000**	**2.8997**	**3.6629**
Ackley	PSO	2.8871	0.8063	1.8959	0.5369
	BA	3.6321	2.0204	2.7663	0.3130
	DA	17.6113	9.3609	14.1071	1.7826
	ALO	17.7727	12.6000	15.2797	1.1493
	GWO	**3.86E−04**	**5.25E−05**	**1.65E−04**	**7.26E−05**

Table 10
Improvement percentage of the proposed model.

Model	Site 1			Site 2			Site 3			Average		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ICEEMDAN-BP	9.10%	8.46%	18.23%	10.79%	6.77%	11.80%	6.96%	3.45%	9.42%	8.95%	6.23%	13.15%
ICEEMDAN-ENN	44.80%	41.05%	39.71%	47.05%	41.46%	35.09%	42.69%	38.38%	34.19%	44.85%	40.29%	36.33%
ICEEMDAN-WNN	45.79%	52.19%	56.85%	72.92%	64.98%	51.32%	30.18%	39.99%	47.63%	49.63%	52.39%	51.94%
ICEEMDAN-GRNN	70.55%	67.04%	61.06%	71.16%	64.73%	53.73%	65.95%	62.76%	55.99%	69.22%	64.84%	56.93%
EMD	41.15%	34.80%	29.21%	49.48%	37.90%	23.82%	48.24%	44.21%	32.98%	46.29%	38.97%	28.67%
CEEMDAN	21.84%	16.28%	13.45%	32.89%	20.09%	15.78%	21.96%	15.17%	10.53%	25.56%	17.18%	13.25%
SSA	49.75%	44.36%	36.42%	53.55%	42.40%	27.02%	38.87%	34.77%	28.10%	47.39%	40.51%	30.51%
BP	67.47%	71.47%	65.91%	69.71%	64.41%	54.34%	65.49%	62.21%	63.87%	67.56%	66.03%	61.37%
ENN	68.52%	66.13%	61.30%	69.68%	63.59%	52.74%	66.18%	62.96%	64.85%	68.13%	64.23%	59.63%
WNN	68.39%	65.61%	60.43%	71.10%	65.60%	55.26%	65.49%	62.18%	63.77%	68.33%	64.46%	59.82%
GRNN	74.64%	71.47%	65.91%	74.79%	69.03%	58.71%	70.30%	67.23%	68.29%	73.24%	69.24%	64.30%
ARIMA	65.24%	59.61%	50.73%	66.64%	57.18%	41.50%	64.02%	58.73%	59.05%	65.30%	58.51%	50.43%
AR	64.72%	58.29%	48.25%	66.99%	57.42%	40.75%	63.62%	58.17%	47.92%	65.11%	57.96%	45.64%
Naïve Benchmark	64.72%	72.69%	72.64%	66.95%	71.71%	66.85%	63.54%	71.74%	69.88%	65.07%	72.04%	69.79%

model shows significant improvements for multi-step forecasting. Specifically, for the one-step forecasting, the proposed model on average leads to 67.56%, 68.13%, 68.33%, 73.24%, 65.30%, 65.11%, and 65.07% reductions, respectively, in comparison with BP, ENN, WNN, GRNN, ARIMA, AR, and the naïve benchmark. Thus, the developed combined model can obtain a satisfactory forecasting effectiveness.

4.4. Combined strategy

To validate the effectiveness of the proposed combined strategy based on the GWO optimization, the simple average strategy, which calculates the average value of forecasting results from all individual models, is selected and applied in this study. Comparison results between the proposed combined model and the simple average strategy are indicated in Table 11.

From Table 11, it is evident that the forecasting performances of the proposed combined model based on the four error metric rules are always more precise than those of the simple average strategy, regardless of the site and the forecasting step. In other words, the developed combined model optimized by GWO, which successfully enhances the forecasting performance of wind speed, significantly outperforms the simple average strategy.

4.5. Computation time

Table 12 presents the average computation time of the proposed model and the other models with multi-step forecasting. As indicated in Table 12, the response of the developed combined model is 239.81 s, which is the longest computation time except that for ARIMA. Meanwhile, the naïve benchmark has the shortest time of only 0.1 s. Specifically, among the three experiments, the average computation time of the four ICEEMDAN-based models are 115.75 s, 114.25 s, 132.23 s, and 118.84 s respectively. Models using different data preprocessing methods spend longer times of 139.39 s, 232.45 s, and 154.12 s. In addition, some classic individual models, i.e., BP, ENN, WNN, and GRNN, spend shorter times of 27.59 s, 15.79 s, 15.78 s, and 20.77 s respectively. Meanwhile, the AR model has a running time of 106.47 s. Although the developed model is the most time-consuming, it has a superior performance compared to all other models, and in practical implementation, the time consumed is acceptable.

4.6. Practical applications in power systems

Wind speed forecasting has a significant influence on enhancing the operation efficiency and increasing the economic benefits of wind power generation systems. Accurate wind speed forecasting could

Table 11
Comparison between the proposed model and the simple average strategy.

Dateset	Multi-step	Model	MAE	RMSE	MAPE	SSE
Site 1	1-step	Simple average strategy	0.3558	0.4691	5.52%	95.0686
		Proposed model	0.2201	0.2994	3.02%	38.7300
	2-step	Simple average strategy	0.3430	0.4441	5.69%	85.2168
		Proposed model	0.2401	0.3224	3.57%	44.9290
	3-step	Simple average strategy	0.3767	0.4813	6.58%	100.0859
		Proposed model	0.2886	0.3849	4.43%	64.1497
Site 2	1-step	Simple average strategy	0.3657	0.5051	5.86%	110.2343
		Proposed model	0.1931	0.2730	2.71%	32.2279
	2-step	Simple average strategy	0.3755	0.5426	6.13%	127.1986
		Proposed model	0.2382	0.3729	3.50%	60.9766
	3-step	Simple average strategy	0.4287	0.6406	7.15%	177.2601
		Proposed model	0.3125	0.5267	4.87%	123.4220
Site 3	1-step	Simple average strategy	0.3596	0.5007	5.75%	108.2861
		Proposed model	0.2409	0.3288	3.58%	46.7293
	2-step	Simple average strategy	0.3503	0.4993	5.73%	107.6769
		Proposed model	0.2620	0.3559	4.12%	54.7558
	3-step	Simple average strategy	0.3919	0.5735	6.47%	142.0896
		Proposed model	0.3123	0.4435	5.13%	85.3247

Table 12
Computation time of the proposed model and other models.

Model	Computation time (s)	Model	Computation time (s)
ICEEMDAN-BP	115.75	BP	27.59
ICEEMDAN-ENN	114.25	ENN	15.79
ICEEMDAN-WNN	132.23	WNN	15.78
ICEEMDAN-GRNN	118.84	GRNN	20.77
EMD	139.39	ARIMA	606.57
CEEMDAN	232.45	AR	106.47
SSA	154.12	Naïve Benchmark	0.1
Proposed Model	239.81		

effectively reduce the risks of wind power generation caused by wind-related uncertainty. Furthermore, effective forecasts are required for various real-time tasks, such as wind power generation scheduling, timely maintenance scheduling, and power system security analysis. In particular, the specific contributions of a robust wind speed forecasting model to research and applications in power systems are as follows:

(1) To ensure the maximum output of wind energy, it is vital to evaluate the wind energy potential of wind farm sites. In fact, wind energy is proportional to the cube of the wind speed; thus, the assessment and evaluation of wind power generation potential could be determined by wind speed forecasting in a way. Therefore, a high-accuracy wind speed forecasting could provide better decisions for wind farms and help in planning smart grids.
(2) A precise wind speed forecasting could provide the necessary reference information for the operation and management of wind turbines. Based on forecasted wind speed values, managers may adjust the wind turbines in time, in order to guarantee the maximum output of wind energy. If the wind turbine capacity is lower than the wind speed values, wind turbines should be shut down to avoid damage and reduce operation cost.
(3) The forecasting performance of wind speed has a significant effect on power system scheduling and management. It is crucial to balance the power demand and supply to maximize economic requirements and sustainable energy management. Power overload will not only affect the quality of power supply and damage the security and stability of the power system, but also increase the cost of wind farms due to electricity storage problems. Hence, an effective prediction model, which could provide accurate forecasting results, could help managers to make right decisions timely so that the abovementioned problems can be avoided.

5. Conclusion

Wind energy, of crucial importance among the low-carbon energy technologies, has aroused widespread interest and research enthusiasm. However, the intermittent characteristics and continuous fluctuation of wind speed series seriously restrict the development of wind power generation. Thus, an effective wind speed forecasting is very vital in enhancing the operation efficiency and increasing the economic benefits of wind power generation systems. In this study, a novel combined model based on a data preprocessing technique, forecasting algorithms, an advanced optimization algorithm, and NNCT, is successfully developed. This new model effectively capitalizes on the merits of individual forecasting models, finally resulting in further improvements. Specifically, the data preprocessing technique is employed to decompose the original series and reconstruct the filtered time series so that the noise signals are eliminated. Then, several individual algorithms are used for forecasting the processed wind speed data. Moreover, a novel deciding weight method based on the GWO algorithm and the leave-one-out strategy is successfully developed to integrate each individual model and obtain the final forecasting result. Experiments using three wind speed datasets collected from the wind farm in the Shandong province of China are used as case studies to estimate the effectiveness of the developed combined model. The experimental results demonstrate that the forecasting performance of the developed model is evidently superior compared to all the other models. Overall, the developed combined model effectively enhances the effectiveness of wind speed prediction and adds a new feasible option for wind speed forecasting and smart grid planning.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant number 71671029).

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Wang J, Zhang W, Wang J, Han T, Kong L. A novel hybrid approach for wind speed prediction. Inf Sci (Ny) 2014;273:304–18. http://dx.doi.org/10.1016/j.ins.2014.02.159.
[2] Georgilakis PS. Technical challenges associated with the integration of wind power into power systems. Renew Sustain Energy Rev 2008;12:852–63. http://dx.doi.org/10.1016/j.rser.2006.10.007.
[3] Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 2016;182:80–93. http://dx.doi.org/10.1016/j.apenergy.2016.08.108.
[4] Global Wind Energy Council. Global wind statistics (2016), p. 2017 < http://www.gwec.net/wp-content/uploads/vip/GWEC_PRstats2016_EN_WEB.pdf > .
[5] Akçay H, Filik T. Short-term wind speed forecasting by spectral analysis from long-term observations with missing values. Appl Energy 2017;191:653–62. http://dx.doi.org/10.1016/j.apenergy.2017.01.063.
[6] Zhang ZS, Sun YZ, Cheng L. Potential of trading wind power as regulation services in the California short-term electricity market. Energy Policy 2013;59:885–97. http://dx.doi.org/10.1016/j.enpol.2013.04.056.
[7] Wang J, Hu J. A robust combination approach for short-term wind speed forecasting and analysis – Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and

LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. Energy 2015;93:41–56. http://dx.doi.org/10.1016/j.energy.2015.08.045.
[8] Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13:915–20. http://dx.doi.org/10.1016/j.rser.2008.02.002.
[9] 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;162:808–26. http://dx.doi.org/10.1016/j.apenergy.2015.10.145.
[10] Landberg L. Short-term prediction of local wind conditions. J Wind Eng Ind Aerodyn 2001;89:235–45. http://dx.doi.org/10.1016/S0167-6105(00)00079-9.
[11] Negnevitsky M, Johnson P, Santoso S. Short term wind power forecasting using hybrid intelligent systems. 2007 IEEE Power Eng Soc Gen Meet 2007:1–4. http://dx.doi.org/10.1109/PES.2007.385453.
[12] Cheng WYY, Liu Y, Bourgeois AJ, Wu Y, Haupt SE. Short-term wind forecast of a data assimilation/weather forecasting system with wind turbine anemometer measurement assimilation. Renew Energy 2017;107:340–51. http://dx.doi.org/10.1016/j.renene.2017.02.014.
[13] Cassola F, Burlando M. Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. Appl Energy 2012;99:154–66. http://dx.doi.org/10.1016/j.apenergy.2012.03.054.
[14] Zhao W, Wei YM, Su Z. One day ahead wind speed forecasting: A resampling-based approach. Appl Energy 2016;178:886–901. http://dx.doi.org/10.1016/j.apenergy.2016.06.098.
[15] Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. North Am Power Symp 2010:1–8. http://dx.doi.org/10.1109/NAPS.2010.5619586.
[16] 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;150:90–107. http://dx.doi.org/10.1016/j.enconman.2017.07.065.
[17] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405–14. http://dx.doi.org/10.1016/j.apenergy.2010.10.031.
[18] Wang Y, Wang J, Zhao G, Dong Y. Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China. Energy Policy 2012;48:284–94. http://dx.doi.org/10.1016/j.enpol.2012.05.026.
[19] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009;34:1388–93. http://dx.doi.org/10.1016/j.renene.2008.09.006.
[20] Han Q, Meng F, Hu T, Chu F. Non-parametric hybrid models for wind speed forecasting. Energy Convers Manag 2017;148:554–68. http://dx.doi.org/10.1016/j.enconman.2017.06.021.
[21] Ait Maatallah O, Achuthan A, Janoyan K, Marzocca P. Recursive wind speed forecasting based on Hammerstein Auto-Regressive model. Appl Energy 2015;145:191–7. http://dx.doi.org/10.1016/j.apenergy.2015.02.032.
[22] 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;143:360–76. http://dx.doi.org/10.1016/j.enconman.2017.04.007.
[23] Tascikaraoglu A, Sanandaji BM, Poolla K, Varaiya P. Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using wavelet transform. Appl Energy 2016;165:735–47. http://dx.doi.org/10.1016/j.apenergy.2015.12.082.
[24] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 2014;31:762–77. http://dx.doi.org/10.1016/j.rser.2013.12.054.
[25] Liu H, Tian H, Pan D, Li Y. Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. Appl Energy 2013;107:191–208. http://dx.doi.org/10.1016/j.apenergy.2013.02.002.
[26] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy 2010;87:2313–20. http://dx.doi.org/10.1016/j.apenergy.2009.12.013.
[27] Du P, Wang J, Yang W, Niu T. Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. Renew Energy 2018;122:533–50. http://dx.doi.org/10.1016/j.renene.2018.01.113.
[28] Cheng CH, Wei LY. A novel time-series model based on empirical mode decomposition for forecasting TAIEX. Econ Model 2014;36:136–41. http://dx.doi.org/10.1016/j.econmod.2013.09.033.
[29] Ramesh Babu N, Jagan Mohan B. Fault classification in power systems using EMD and SVM. Ain Shams Eng J 2015:1–9. http://dx.doi.org/10.1016/j.asej.2015.08.005.
[30] Hong YY, Chang HL, Chiu CS. Hour-ahead wind power and speed forecasting using simultaneous perturbation stochastic approximation (SPSA) algorithm and neural network with fuzzy inputs. Energy 2010;35:3870–6. http://dx.doi.org/10.1016/j.energy.2010.05.041.
[31] Chenthur Pandian S, Duraiswamy K, Rajan CCA, Kanagaraj N. Fuzzy approach for short term load forecasting. Electr Power Syst Res 2006;76:541–8. http://dx.doi.org/10.1016/j.epsr.2005.09.018.
[32] Guo Z, Wu J, Lu H, Wang J. A case study on a hybrid wind speed forecasting method using BP neural network. Knowledge-Based Syst 2011;24:1048–56. http://dx.doi.org/10.1016/j.knosys.2011.04.019.
[33] Liu H, Tian H, Liang X, Li Y. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy 2015;157:183–94. http://dx.doi.org/10.1016/j.apenergy.2015.08.014.
[34] Zhang C, Wei H, Xie L, Shen Y, Zhang K. Direct interval forecasting of wind speed using radial basis function neural networks in a multi-objective optimization framework. Neurocomputing 2016;205:53–63. http://dx.doi.org/10.1016/j.neucom.2016.03.061.
[35] Xiao L, Shao W, Yu M, Ma J, Jin C. Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting. Appl Energy 2017;198:203–22. http://dx.doi.org/10.1016/j.apenergy.2017.04.039.
[36] Zhou J, Shi J, Li G. Fine tuning support vector machines for short-term wind speed forecasting. Energy Convers Manage 2011;52:1990–8. http://dx.doi.org/10.1016/j.enconman.2010.11.007.
[37] Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 2014;34:243–54. http://dx.doi.org/10.1016/j.rser.2014.03.033.
[38] 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 Manage 2017;136:439–51. http://dx.doi.org/10.1016/j.enconman.2017.01.022.
[39] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, 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;454:903–95. http://dx.doi.org/10.1098/rspa.1998.0193.
[40] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;2009:1–41. http://dx.doi.org/10.1142/S1793536909000047.
[41] Colominas MA, Schlotthauer G, Torres ME. Improved complete ensemble EMD: a suitable tool for biomedical signal processing. Biomed Signal Process Control 2014;14:19–29. http://dx.doi.org/10.1016/j.bspc.2014.06.009.
[42] Yang W, Wang J, Wang R. Research and application of a novel hybrid model based on data selection and artificial intelligence algorithm for short term load forecasting. Entropy 2017;19. http://dx.doi.org/10.3390/e19020052.
[43] 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://dx.doi.org/10.1016/j.apenergy.2017.10.031.
[44] Feng C, Cui M, Hodge BM, Zhang J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl Energy 2017;190:1245–57. http://dx.doi.org/10.1016/j.apenergy.2017.01.043.
[45] Iversen EB, Morales JM, Møller JK, Madsen H. Short-term probabilistic forecasting of wind speed using stochastic differential equations. Int J Forecast 2015;32:981–90. http://dx.doi.org/10.1016/j.ijforecast.2015.03.001.
[46] Xiao L, Wang J, Hou R, Wu J. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 2015;82:524–49. http://dx.doi.org/10.1016/j.energy.2015.01.063.
[47] Xiao L, Wang J, Dong Y, Wu J. Combined forecasting models for wind energy forecasting: a case study in China. Renew Sustain Energy Rev 2015;44:271–88. http://dx.doi.org/10.1016/j.rser.2014.12.012.
[48] Ren C, An N, Wang J, Li L, Hu B, Shang D. Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowledge-Based Syst 2014;56:226–39. http://dx.doi.org/10.1016/j.knosys.2013.11.015.
[49] Wang J, Yang W, Du P, Li Y. Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. Energy 2018;148:59–78. http://dx.doi.org/10.1016/j.energy.2018.01.112.
[50] Wang J-J, Zhang W-Y, Liu X, Wang C-Y. Modifying wind speed data observed from manual observation system to automatic observation system using wavelet neural network. Phys Procedia 2012;25:1980–7. http://dx.doi.org/10.1016/j.phpro.2012.03.338.
[51] Marvuglia A, Messineo A. Monitoring of wind farms' power curves using machine learning techniques. Appl Energy 2012;98:574–83. http://dx.doi.org/10.1016/j.apenergy.2012.04.037.
[52] Torres ME, Colominas MA, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. ICASSP IEEE Int Conf Acoust Speech Signal Process – Proc 2011:4144–7. http://dx.doi.org/10.1109/ICASSP.2011.5947265.
[53] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw 2014;69:46–61. http://dx.doi.org/10.1016/j.advengsoft.2013.12.007.
[54] Xu Y, Yang W, Wang J. Air quality early-warning system for cities in China. Atmos Environ 2017;148:239–57. http://dx.doi.org/10.1016/j.atmosenv.2016.10.046.
[55] Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat 1995;13:253–63. http://dx.doi.org/10.1080/07350015.1995.10524599.
[56] Chen H, Hou D. Research on superior combination forecasting model based on forecasting effective measure. J Univ Sci Technol China 2002:172–80.