10.1016/j.renene.2018.05.093

Research on a Combined Model Based on Linear and Nonlinear Features - A Case Study of Wind Speed Forecasting

Kequan Zhang [1], Zongxi Qu [1,*], Yunxuan Dong [2], Haiyan Lu [3], Wennan Leng [1], Jianzhou Wang [4], Wenyu Zhang [1]

[1] A key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China; zhangkq@lzu.edu.cn; quzx14@lzu.edu.cn; lengwn15@lzu.edu.cn; yuzhang@lzu.edu.cn
[2] School of mathematics and statics, Lanzhou University, China; dongyx15@lzu.edu.cn
[3] Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia; haiyan.lu@uts.edu.au
[4] School of Statistics, Dongbei University of Finance and Economics, China; wjz@lzu.edu.cn
* Correspondence: quzx14@lzu.edu.cn;

Abstract: As one of the most promising sustainable energy sources, wind energy is being paid more attention by the researchers. Because of the volatility and instability of wind speed series, wind power integration faces a severe challenge; thus, an accurate wind energy forecasting plays a key role in smart grid planning and management. However, many traditional forecasting models do not consider the necessity and importance of data preprocessing and neglect the limitation of using a single forecasting model, which leads to poor forecasting accuracy. To solve these problems, a novel combined model based on two linear and four nonlinear forecasting algorithms is proposed to adapt both the linear and nonlinear characteristics of the wind energy time series. In addition, a modified Artificial Fish Swarm Algorithm and Ant Colony Optimization (AFSA-ACO) algorithm is proposed and employed to determine the optimal weight coefficients of the combined models. To verify the forecasting performance of the developed combined model, several experiments were implemented by using ten-minute interval wind speed data in Shandong, China. Then, one-step (ten-minute), three-step (thirty-minute) and five-step (fifty-minute) predictions were conducted. The experimental results indicate that the developed combined model is remarkably superior to all benchmark models for the high precision and stability of wind-speed predictions.

Keywords: Wind speed forecasting; Combined model; Artificial fish swarm algorithm; Ant colony optimization.

1. Introduction

Wind energy, as an alternative to fossil fuel-generated electricity, has received increasing attention around the world due to its abundance, wide distribution, and economics as a non-polluting type of renewable energy [1]. The global cumulatively installed wind capacity reached approximately 539.58GW by the end of 2017, which an annual increase of 9.7%. Furthermore, researchers have noted that there will be a new round of installed wind energy in 2018-2021, and that 22% of the world's electricity will be supplied by wind energy by 2030 [2]. However, the intermittent

and stochastic nature of wind speed always poses many technical challenges, such as increasing the cost and reducing the performance and reliability of the wind industry and power systems, especially when wind power is integrated into traditional grid systems. Therefore, to alleviate the abovementioned problems, accurate and reliable information of dynamic wind power is increasingly significant and urgently needed, and it can not only be utilized to control the wind turbine, but also help make decisions on wind power system's management [3].

Over the past several decades, a large number of forecasting models for wind speed have been developed and applied to obtain precise information of wind energy. Generally, based on the time horizons, the research of wind speed prediction can be classified into four types: very short-term forecasting, short-term forecasting, medium-term forecasting and long-term forecasting [4]. According to computational mechanisms, these forecasting technologies can be divided into three general classes: physical methods, statistical approaches and intelligent learning methods [3,5,6].

The physical methods are based mostly on Numeric Weather Prediction (NWP), which usually applies meteorological information such as temperature and pressure, for wind speed forecasting. However, NWP is unsuitable for short-term wind speeds because of its complex calculation processes, high costs and poor performance [7]. Compared with physical methods, statistical and intelligent-learning approaches, which extracts the laws of historical data, can provide more accurate results for short-term wind speed forecasting. Specifically, the statistical methods, also taken as time series models, are widely applied and easily implemented in the forecasting field. However, these algorithms, including the auto-regressive integrated moving average (ARIMA), Markov chain, exponential soothing (ES), and Kalman filtering method, can achieve more accurate short-term wind speed predictions than can physical models, which capture the linear relationship between the past and future performance of a time series [8]**.** In comparison, intelligent learning algorithms such as the Back Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), Generalized Regression Neural Network (GRNN) and Wavelet Neural Network (WNN), which structure a network relationship of input and output information, are widely used in the field of wind speed forecasting and perform better at handling the nonlinear features of wind speed data [9]. For example, Guo et al. [10] developed a forecasting model using BPNN and seasonal exponential adjustments to predict wind speed and eventually obtained good forecasts. Yu et al. [11] successfully exploited an Elman Neural Network (ENN) based on an improved wavelet transform using singular spectrum analysis for wind speed forecasting, resulting in good performance.

Because of the inherent weaknesses of every model and the intermittency, complex fluctuation and the linear and nonlinear characteristics of wind speed series, the abovementioned forecasting approaches cannot always capture the trend of the wind speed, especially in nonlinear time series, which usually leads to poor forecasting performance. Thus, with the aim of achieving higher accuracy levels and

wider forecast horizons, combined model, that combine the advantages of the single forecasting models, have been developed and widely utilized for wind speed prediction [12]. For example, Wang et al. [13] developed a combined model by applying four artificial neural networks (ANNs), the multi-objective bat algorithm and singular spectrum analysis to forecast a 10-min wind speed. The results indicated that the combined model can improve the prediction performance to some extent. Xiao et al. [14] presented a combined model on the basic of the chaotic particle swarm optimization algorithm to optimize the weight coefficients for wind speed prediction.

However, wind speed series have linear and non-linear characteristics. Hence, to improve the forecasting performance, the linear and nonlinear trends that cause the intermittency and complex fluctuations of wind speed should both be considered in the combined models. To capture the linear and nonlinear traits, both linear models, such as ARIMA and ES, and nonlinear methods, such as BPNN, WNN, ENN should be employed in the combined models. Thus, a novel combined forecasting model based on two linear models, i.e., ES and ARIMA, and four nonlinear methods, i.e., BPNN, GRNN, Wavelet Neural Network (WNN) and ENN, is proposed in this paper. To further study the combined models, many evolutional optimization algorithms are applied to help determine the optimal weights of the combination model. For example, Yang et al. [15] used differential evolution (DE) to optimize the weight coefficients of a combination model to improve forecasting accuracy. Xiao et al. [14] applied the cuckoo search algorithm (CSO) to determine the optimal weights of a combined model. Rahmani et al. [16] developed a hybrid algorithm with ant colony and particle swarm optimization for wind energy forecasting. The ant colony optimization (ACO) algorithm, which is based on the behavior of natural ant colonies, has been widely used in many in many engineering optimization problems. Recently, to improve the performance of the ACO with respect to optimized problems, many new modifications have been proposed. Sen et al. [17] developed a hybrid optimization algorithm, ACO–ABC–HS, to solve the problem of economic dispatch (ED) for a multi-generator system. The hybrid algorithm combines the framework of ACO, artificial bee colony (ABC) and harmonic search (HS) algorithms to find the optimized solution for the system. Patel el al. [18] developed a hybrid ant colony optimization (ACO)/particle swarm optimization (PSO) technique to optimize the multicast tree; the results showed that ACO-PSO outperforms both the ACO and the PSO. To enhance the exploration capacities and improve the local searching ability of the ant colony optimization, a modified version of the AFSA and ACO, i.e., the AFSA-ACO, is proposed in this paper. It has been successfully applied to determine the weight coefficients of the combined model. Compared with the artificial fish swarm algorithm (AFSA) and ACO algorithm, the proposed AFSA-ACO, which has been tested by several benchmark functions, not only enhances the performance of the AFSA and ACO with more optimum solutions and higher convergence rates but also has stronger

capabilities to solve more complex problems. To conduct a case study, the original wind speed datasets with 10-min intervals from four sites in the wind farms of Penglai, Shandong Province, China, were used to verify the proposed combined forecasting model. Since the original wind-speed series are highly unstable and noisy, the singular-spectrum analysis (SSA) [19] was used for data pre-processing to remove the noise of the original wind-speed data for easy forecasting.

The rest of the paper is organized as follows: Section 2 introduces the related individual forecasting models. The combined model and combined theory are presented in Section 3, where a novel modified optimization algorithm, AFSA-ACO, is also presented. The forecasting results of the proposed combined model and comparisons are discussed in Section 4. An insightful discussion of the developed model is conducted in Section 5. Finally, Section 6 concludes the paper.

List of abbreviations

AFSA-ACO	Artificial Fish Swarm Algorithm and Ant Colony Optimization	ENN	Elman Neural Network
AFSA	Artificial Fish Swarm Algorithm	ES	Exponential Smoothing
ACO	Ant Colony Optimization	WNN	Wavelet Neural Network
NWP	Numerical Weather-Prediction	GA	Genetic Algorithm
ARIMA	Auto Regressive Integrated Moving Average Model	PSO	Particle Swarm Optimization
BPNN	Back Propagation Neural Network	MSE	Mean Squared Error
GRNN	General Regression Neural Network	MAE MAPE	Square Sum of the Error Average of Absolute Error
RBFNN	Radial Basis Function Neural Network	NNCT	No Negative Constraint Theory
SVM	Support Vector Machine	SSA	Singular Spectrum Analysis
CS	Cuckoo Search		

2. Individual Forecasting Models

There are many individual models that can be used to handle the issue of wind speed forecasting and are considered effective. Considering the inherent linear and nonlinear features contained in wind speed data, two linear algorithms and four nonlinear algorithms, which performs better in wind speed forecasting, are used to develop the combined forecasting model with the non-positive constraint theory.

2.1 Linear algorithms

The Autoregressive Integrated Moving Average (ARIMA) model and Exponential Smoothing(ES) algorithm are chosen as linear models to capture the linear feature of wind speed data in the developed combined model, which is investigates in a series of experiments [20].

2.1.1 Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model, developed by Box and Jenkins [21], is one of the most popular and widely used forecasting methods. It consists of three parts: autoregressive (AR), moving average (MA) and integration (I). The details are briefly described as follows:

The AR (*p*) model is an autoregressive model with order p, which is defined as:

$$y_t = c + a_1 y_{t-1} + \ldots + a_p y_{t-p} + u_t \tag{1}$$

where $a_1, \ldots, a_p$ are parameters, u_t is a random variable as a white noise, and *c* is constant.

The MA(*q*) model is a moving average model with order *q*, which is defined as:

$$y_t = \mu + u_t + m_1 u_{t-1} + \ldots + m_q u_{t-q} \tag{2}$$

where $m_1, \ldots, m_q$ are parameters, μ is the expectation of y_t and $u_t, u_{t-1}, \ldots, u_{t-q}$ are white noise error terms.

The ARMA (*p, q*) model is combined with p autoregressive terms and q moving-average terms, and is defined as:

$$y_t = c + a_1 y_{t-1} + \ldots + a_p y_{t-p} + u_t + m_1 u_{t-1} + \ldots + m_q u_{t-q} \tag{3}$$

When AR(*p*), MA(*q*) and ARMA (*p, q*) are used for data with non-stationarity, the differencing step corresponding to the integration part should be performed to eliminate the non-stationarity of researched data, which is defined as an ARIMA (p, d, q) model.

2.2.2 Exponential Smoothing (ES)

The double exponential smoothing model is defined as:

$$R_t^{(2)} = \alpha R_t^{(1)} + (1-\alpha) R_{t-1}^{(2)} \tag{4}$$

where α is the smoothing factor $R_t^{(2)}, R_{t-1}^{(2)}$ are the second order exponential smoothing values for stages *t* and *t*-1, and $R_t^{(1)}$ is the first order exponential smoothing value.

The initial value is determined by

$$R_0^{(1)} = R_0^{(2)} = X_0 \tag{5}$$

where X_0 is the initial value of the original data.

The forecasting formula of the second exponential smoothing is defined as

$$R_{t+T} = a_t + b_t \cdot T \tag{6}$$

$$\begin{cases} a_t = 2R_t^{(1)} - R_t^{(2)} \\ b_t = \dfrac{\alpha}{1-\alpha}\left(R_t^{(1)} - R_t^{(2)}\right) \end{cases} \tag{7}$$

where R_{t+T} is the forecasting value of stage $t+T$, and *T* denotes the forecasting stage from stage *t*.

2.2 Nonlinear algorithms

There are several nonlinear algorithms employed to conduct wind speed forecasting. The Back Propagation Neural Network (BPNN), Generalized Regression Neural Network (GRNN), Wavelet Neural Network (WNN) and Elman Neural Network (ENN), have been adopted as nonlinear algorithms to capture the nonlinear characteristics inherent in wind speed data, which are known to be better for wind speed forecasting with relative accuracy. The construction of each nonlinear models is shown in **Fig. 2**.

2.2.1 Back Propagation Neural Network (BPNN)

The Back-propagation algorithm, proposed in 1986 by Rumelhart et al [22], is one of the most popular neural networks. BPNN is a type of multilayer feed-forward neural network that is widely used, and it is based on the gradient descent method. As shown in **Fig. 2**, BPNN consists of three types of layers: an input layer, a hidden layer and an output layer.

Definition 1. The process of the back-propagation algorithm is divided into two phases: updating and learning.

$$w_{ij}(t) = w_{ij}(t-1) - \Delta w_{ij}(t) \tag{8}$$

$$\Delta w_{ij}(t) = \eta\, \partial E / \partial w_{ij}(t-1) + \alpha \cdot \Delta w_{ij}(t-1) \tag{9}$$

where w_{ij} is the weight between nodes i and j, η is the learning speed, α is the impulse parameter, t is the current iterative steps, and E is the error super curve face.

Definition 2. During the learning phase, the thresholds and weights of the joints between two layers are updated by subsequently back-propagating the errors, minimizing the mean squared error (MSE).

$$MSE = \frac{1}{2}\sum_{k=1}^{m}(Y_k - O_k)^2 \tag{10}$$

where O_k denotes the output values, and Y_k denotes the desired output.

2.2.2 Generalized Regression Neural Network (GRNN)

The GRNN model, proposed by Specht [23], is a new type of neural network model. There are four types of layers: the input layer, the pattern layer, the summation layer, and the output layer, as presented in **Fig 2.**

Definition 3. There is a nonlinear transformation in the pattern layer, which binds input space with pattern space. The relationship between the input neuron and the appropriate response from the pattern layer can be memorized in each neuron in this layer.

Definition 4. The summation layer has two summations: the simple summation S_s and the weighted summation S_w. Then, the summation layer is transmitted to the output layer.

2.2.3 Wavelet Neural Network (WNN)

The basic structure of WNN is based on wavelet analysis and is presented in **Fig. 2.** It substitutes the conventional sigmoid function with a mother wavelet.

Definition 5. Function $f(x)$ can approximately describe a cluster of mother wavelets:

$$\hat{f}(x)=\sum_{l=0}^{N} w_l \cdot \psi_l\left(\frac{x-b_1}{a_1}\right) \tag{11}$$

where $\hat{f}(x)$ denotes a fitting function, w_l denotes a weighting coefficient, and N denotes the quantity of mother wavelets.

2.2.4 Elman Neural Network (ENN)

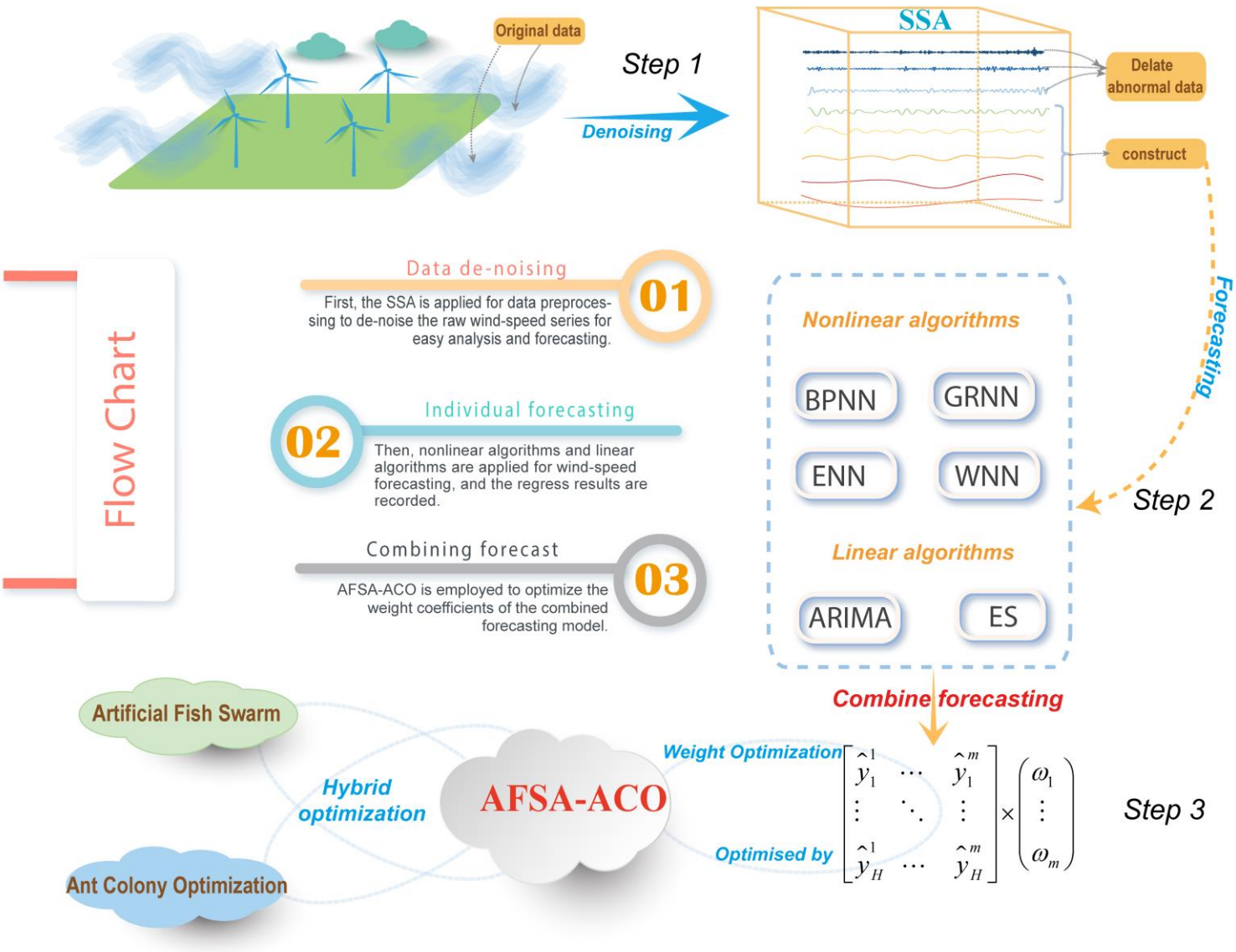
The ENN model, introduced by Elman, consists of three layers: an input layer, a hidden layer, and an output layer. A simple three-layer ENN is presented in **Fig. 2**. The difference between ENN and the three-layer feed-forward neural network is that the ENN model has a context layer to respond to the outputs of the hidden layer. The activations of the context layer is [24]

$$S_i(t)=g\left(\sum_{k=1}^{K} V_{ik} S_k(t-1)+\sum_{j=1}^{j} W_{ij} I_j(t-1)\right) \tag{12}$$

where $S_k(t)$ and $I_j(t)$ are the output of the context layers and input layers, respectively; V_{ik} and W_{ij} are the weights; and $g(x)$ is a sigmoid transfer function.

3. Combined Model

The combined forecasting method is started with J.M. Bates and C.W.J. Granger in the 1960s. They, proved that using two or more models to forecast results is better than using a single model, that is the limitations of the single forecasting model are overcome by the combined forecasting method [25]. Because both linear and nonlinear characteristics appeared in the wind speed time series, a new combined model is developed in this paper that integrates linear and nonlinear models, non-positive constraint theory and a modified optimization algorithm. The main processes of the proposed combined model are illustrated in **Fig. 1**.


Original data
Step 1
Denoising
SSA
Delate abnormal data
construct
Forecasting
Flow Chart
Data de-noising
01
First, the SSA is applied for data preprocessing to de-noise the raw wind-speed series for easy analysis and forecasting.
Individual forecasting
02
Then, nonlinear algorithms and linear algorithms are applied for wind-speed forecasting, and the regress results are recorded.
Combining forecast
03
AFSA-ACO is employed to optimize the weight coefficients of the combined forecasting model.
Nonlinear algorithms
BPNN
GRNN
ENN
WNN
Linear algorithms
ARIMA
ES
Step 2
Combine forecasting
Artificial Fish Swarm
Hybrid optimization
AFSA-ACO
Ant Colony Optimization
Weight Optimization
Optimised by
Step 3

248

Fig. 1. The flowchart of proposed the combined model

3.1 Combined theory

The key of combined forecasting theory is to estimate the weight coefficient. According to an objective function that is contributed by a certain standard, the optimized weight coefficient can be found by minimizing the objective function. The objective function is generally determined by the error value (e.g., Mean Absolute Error (MAE), Mean Square Error (MSE) and Sum of Squares for Error (SSE) et.al) and a series of minimizing standards is explored.

Definition 6. The weight coefficient of the traditional combined forecasting model is calculated by minimizing SSE:

$$\min \boldsymbol{L} = \boldsymbol{W}^T \boldsymbol{E} \boldsymbol{W} = \sum_{t=1}^{T} \sum_{j=1}^{m} \sum_{i=1}^{m} \mathrm{w}_i \mathrm{w}_j e_{it} e_{jt} \begin{cases} \boldsymbol{R}^T \boldsymbol{W} = 1 \\ \boldsymbol{W} \geq 0 \end{cases} \tag{13}$$

where $\boldsymbol{W} = \left(w_1, w_2, ..., w_m\right)^T$ is the weight vector; $\boldsymbol{R} = \left(1, 1, ..., 1\right)^T$ is a column vector and $\boldsymbol{E} = \left(\boldsymbol{E}_{ij}\right)_{m \times m}$ is forecasting error information matrix.

Definition 7. Traditional combination approach based on no negative constraint theory (TCM-NNCT) [14] was developed to improve the traditional method, which is presented as

$$\begin{aligned} &\min \boldsymbol{L} = \boldsymbol{W}^T \boldsymbol{E} \boldsymbol{W} = \sum_{t=1}^{T} \sum_{j=1}^{m} \sum_{i=1}^{m} w_i w_j e_{it} e_{jt} \\ &\text{st} \quad \boldsymbol{R}^T \boldsymbol{W} = 1 \end{aligned} \tag{14}$$

For the improved method, TCM-NNCT, the experimental results demonstrated that the combination forecasting model can achieve desirable results when the range of the weight vector $\boldsymbol{W}$ is [-2, 2] rather than [0, 1]. The provided weight-determined approach was tested by experiment instead of relying on theory.

In this paper, a novel combined forecasting model based on TCM-NNCT and weight-coefficient optimization was developed, and contains two linear algorithms (i.e., ES and ARIMA) and four neural network algorithms (i.e., BPNN, GRNN, WNN and ENN), as shown in **Fig. 2**.

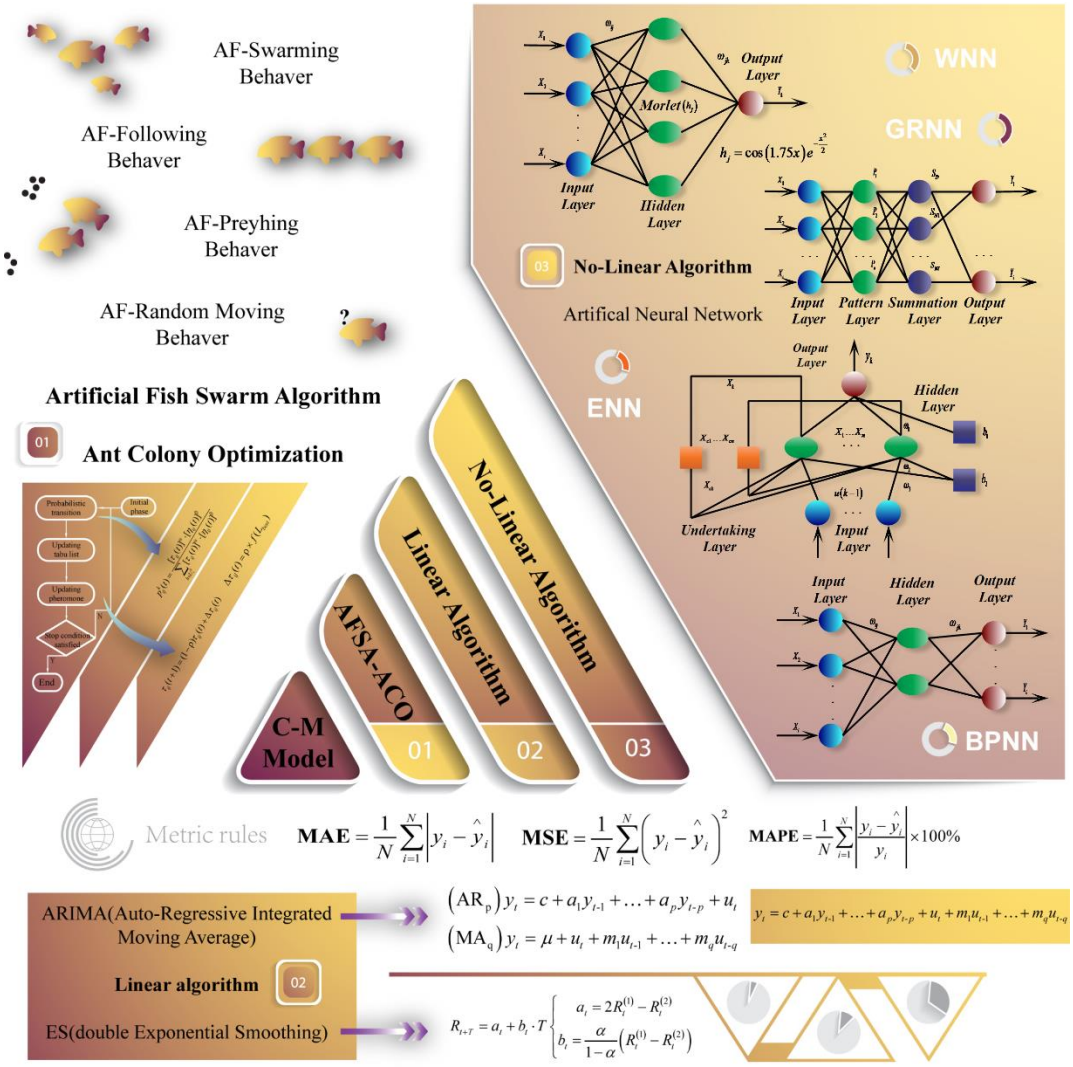


Fig. 2 Structure of the combined models.

3.2 Optimization algorithm-AFSA-ACO

To further improve the forecasting performance of the developed model, considering the limitations of ACO in the later optimization period (i.e., the slow convergence speed and the imprecise convergence accuracy), we modified the ACO with AFSA and then developed a new optimization algorithm named AFSA-ACO. The details are introduced as follows.

3.2.1 ACO

ACO is a heuristics optimization algorithm, proposed by Dorigo, that mimics the cooperative behavior of ant colonies and is has been used to handle the numerous combinatorial optimization problems. It has been proven that the ACO algorithm has robust ability to search near-optimum solutions [26]. When ants search for foods, each ant will try its best to find the optimum routes between the nest and food source. In this process, the ants leave a chemical trail of pheromones on the ground, which guides the behavior of other ants. However, as time passes, the pheromones evaporate. Therefore, the pheromones must be enhanced through more

ants, so the pheromones will not disappear. The main steps of ACO can be summarized as follows:

Step1. Initialization
Generate S_0 initial solution
Step2. Local search
Use the local search procedure to solution S_0.
Repeat
Step3. Perturbation
Perturb solution S_0 and obtain new solution S'.
Step4. Local search
Use the local search procedure to solution S' and obtain solution S".
Step5. Acceptance criterion
Perform acceptance criterion
Until termination condition met

3.2.2 AFSA

In the water, fish can find nutrients alone or follow other fish to find more nutrients. Thus, the highest number of local fish that survive generally is the highest concentration of nutrients in the water. The Artificial fish algorithm is based on this characteristic and constructs artificial fish to simulate fish feed, cluster and collision behavior and thereby realize optimization.

(1) Preying behavior

Assuming the present situation of an artificial fish is X_i, then randomly choose a new state X_j. It is well known that we can convert a maximum problem into the minimum problem. Hence in this case, we just take the maximum problem as an example in the following discussions. The selected fish should move a step in that direction if the maximum problem $Y_i<Y_j$; otherwise, select a state, X_j, randomly one time and determine whether it meets the pre-set forward situation. If it cannot meet the situation, it will move one step at random. The rules of step moving can be expressed by

$$\begin{cases} X_{i+1} = X_i + Step\dfrac{X_j - X_i}{\left\| X_j - X_i \right\|} & \left(Y_j > Y_i\right) \\ X_{i+1} = X_i + Step & \left(Y_j \le Y_i\right) \end{cases} \tag{15}$$

(2) Clustering behavior

Artificial fish at present situation X_i search for the partners' number NF and the central position X, $\left(d_{ij} < Visual\right)$; if $Y_c/NF > \delta Y_i$, it represents that enough food is at the center of the fish colony, where it is not too crowded. The related mathematical simulation formulas of the swarming behavior can be expressed as follows:

$$\begin{cases} X_{i+1} = X_i + Step\dfrac{X_c - X_i}{\left\| X_c - X_i \right\|} & \left(Y_c/NF > \delta Y_i \; and \; NF \ge 1\right) \\ X_{i+1} = \textbf{\textit{Defalut Formula}} & \left(Y_c/NF \le \delta Y_i \; or \; NF = 0\right) \end{cases} \tag{16}$$

(3) Following behavior

Assume that X_i is the present state of AF looking for partner $X_{\max}$ near $Y_{\max}$, if $Y_{\max}/NF > \delta Y_i$, we can see that the present position of partner $X_{\max}$ possesses higher

food consistence and it can be not crowded. The AF will take a step toward partner $X_{\max}$; otherwise, to the process of the searching behavior will be run.

The behavior can be expressed by the following mathematic description:

$$\begin{cases} X_{i+1} = X_i + Step\dfrac{X_c - X_i}{\|X_c - X_i\|} & \left(Y_{\max}/NF > \delta Y_i \text{ and } NF \geq 1\right) \\ X_{i+1} = \textbf{\textit{Defalut Formula}} & \left(Y_{\max}/NF \leq \delta Y_i \text{ or } NF = 0\right) \end{cases} \tag{17}$$

(4) Behavior selection

Based on the problem which should be addressed, we can choose a behavior to simulate after evaluating the current environment of the AFSA. Trial approaches have been continually taken to simulate fish behaviors, and the best solutions can be implemented after evaluation. Here, we call the three biological behaviors of fish swarm as the searching behavior, the following behavior and the clustering behavior.

(5) Bulletin

Bulletin can be utilized to memorize the AF's best situation and the optimal solutions of the given problem. Every AF updates and compares its own state with the bulletin after moving. If the value on the bulletin is worse than the present state of the AF, the bulletin will be replaced. The pseudocode of the AFSA algorithm is as follows:

Algorithm: AFSA

Input:

$X= (x_1, x_2, x_3, \ldots, x_n)$—the individual conditions of artificial fishes.

$Y=$***maxf***(x)—objective function.

Output:

Y_{best}—the global optimal solution of the objective function.

Parameters:

$iter_{max}$—maximum iterations.

d—the distance of each artificial fish.

δ—the factor of crowding degree.

n_f—the number of partners of target fish and target fish swarm.

Visual—the cognitive distance of artificial fishes.

step—the maximum shift step size of artificial fishes.

1: **FOR EACH** $i: 1 \leq i \leq length(X)$ **DO**

2: *X(i)=**rand**(N); Y(i)=**maxf**(X(i));*

3: **END FOR**

4: /*Initialize population with random method*/

5: **WHILE** $(iter \leq iter_{\max})$ **DO**

6: **FOR EACH** $i, j: 1 \leq i, j \leq length(X)$ **DO**

7: **IF** $(Y_j > Y_i) \&\& (d_{i,j} \leq Visual) \&\& ((Y_j / n_f) \geq \delta Y_j)$ **THEN**

8: Individual fish moves a step towards target fish;

9: **ELSEIF** $(Y_i < Y_j) \&\& (d_{i,j} \leq Visual)$ **THEN**

10: Individual fish moves a step towards target fish;

11: **ELSE THEN**

12: Individual fish moves a step in a random direction.;

13: **END IF**

14: /*Execute AF-Follow, elseact AF-Prey, do AF-Move as the default*/

15: **IF** $(Yc > Y_i) \&\& (d_{i,j} \le Visual) \&\& ((Yc / n_f) \ge \delta Y_j)$ **THEN**
16: Individual fish moves random steps towards target fish swarm;
17: **ELSEIF** $(Y_i < Y_j) \&\& (d_{i,j} \le Visual)$ **THEN**
18: Individual fish moves a step towards target fish;
19: **ELSE THEN**
20: Individual fish moves a step in a random direction;
21: **END IF**
22: /*Execute AF-Swarm, elseact AF-Prey, do AF-Move as the default*/
23: Choose the better result of AF-Swarm and AF-Fellow by ***maxf***;
24: Test whether the result is eligibility, break the circulation if the result is eligibility;
25: iter=iter+1;
26: **END WHILE**

3.2.3 AFSA-ACO

ACO possesses strong global exploration and exploitation capabilities. Moreover, the convergence speed of this algorithm is fast in early optimization [27]. Nevertheless, in the later optimization process, its convergence speed and accuracy decrease. To address these shortcomings, a novel modified algorithm on the basis of the AFSA algorithm, named AFSA-ACO, is proposed in this study. A rudimentary AFSA-ACO algorithm is shown as follows:

Algorithm: AFSA-ACO

Input:

$X = (x_1, x_2, x_3, \ldots, x_n)$—the individual conditions.

$Y = \boldsymbol{maxf}(x)$—objective function.

Output:

Ybest—the global optimal solution of the objective function.

Parameters:

g_{ratio}—minimum update rate for AFSA.

N—maximum iterations of ACO.

$iter_{max}$—maximum iterations of AFSA.

τ—the concentration of pheromone.

ρ—pheromone volatilization coefficients.

η—inspiring factor.

1: Initialize initial parameter and generate initial fish school;
2: **WHILE** $(iter \le iter_{\max})$ **DO**
3: Use AFSA to get the optimal solution;
4: **IF** $(g_{end} < g_{ratio})$ **THEN**
5: Break the iteration of AFSA, export the result $Ybest_{AFSA}$;
6: **END IF**
7: **END WHILE**
8: /*Use AFSA to get the fast-optimal solution for the initial ant colony*/
9: Initialize initial pheromone by $\tau_{\max} = \frac{1}{2(1-\rho)} \cdot \frac{1}{L_{best}}$ and the results of AFSA;
10: **WHILE** $(1 \le i, j \le N)$ **DO**

11: $p_{ij}^{k}(t)=\dfrac{[\tau_{ij}(t)]^{\alpha}\cdot[\eta_{ij}(t)]^{\beta}}{\sum_{h\in U_i^k}[\tau_{ij}(t)]^{\alpha}\cdot[\eta_{ij}(t)]^{\beta}}$

12: /* Calculate the transition probability*/

13: $\tau_{ij}(t+1)=(1-\rho)\tau_{ij}(t)+\Delta\tau_{ij}(t);\ \Delta\tau_{ij}(t)=\rho\times f(L_{best})$

14: /*Updating the pheromone and L_{best} is the optimal solution at present*/

15: **END WHILE**

3.2.4 Test functions for validation of ACO and AFSA-ACO

In this section, four test functions, named Sphere, Rosenbrock, Rastrigin and Schaffer, as shown in **Table 1**, are employed to test the optimization performance of ACO and the developed optimized algorithm AFSA-ACO. The test of ACO and AFSA-ACO was performed in MATLAB R2012b on Windows 8 with a 1.80 GHz Intel Core i7 4500U, 64 bits and 8 GB RAM. The ACO and AFSA-ACO parameters are given in **Table 2**. These experimental parameters are widely accepted for solving optimization problems and have usually obtained good results [26].

Table 1 Test functions for validation of ACO and AFSA-ACO

Function name	Test function	Variable domain	Global optimum
Sphere	$f(x)=\sum_{i=1}^{d}x_i^2$	$x_i\in[-5.12,5.12]$	$f_{\min}(0,0,0\cdots0)=0$
Rosenbrock	$f(X)=\sum_{i=1}^{d-1}\left[100\left(x_i^2-x_{i+1}\right)^2+\left(x_i-1\right)^2\right]$	$x_i\in[-2.084,2.084]$	$f_{\min}(1,1,1\cdots1)=0$
Rastrigin	$f(X)=\sum_{i=1}^{d}\left(x_i^2-10\left(2\pi x_i\right)+10\right)$	$x_i\in[-5.12,5.12]$	$f_{\min}(0,0,0\cdots0)=0$
Schaffer	$f(X)=\dfrac{\sin^2\sqrt{\sum_{i=1}^{d}x_i^2}-0.5}{\left[1+0.001\left(\sum_{i=1}^{d}x_i^2\right)\right]^2}+0.5$	$x_i\in[-5.12,5.12]$	$f_{\min}(0,0,0\cdots0)=0$

Table. 2 The experimental parameters of ACO and AFSA-ACO

Experimental parameters	ACO	AFSA-ACO
Maximum generation	10000	10000
Population size	100	100
A	0.1	0.1
β_0	1.0/3.5	1.0/3.5
Γ	0.001	0.001
Convergence tolerance	10^{-5}	10^{-5}

Based on the experimental results (as shown in **Table 3**), we present several conclusions, summarized as follows:

- For the Sphere function, when the dimension is 10, the maximum, minimum and average values of the iteration of ACO and AFSA-ACO are 175 and 16, 154 and 3, and 162 and 7.4, respectively. When the dimension is 20, the maximum, minimum and average values of the iteration of ACO and AFSA-ACO are 242 and 42, 134 and 19, and 175 and 28.6, respectively. When the dimension is 50, the values of the iteration of ACO and AFSA-ACO are 477 and 150, 365 and 72, and 397 and 101.3, respectively.
- For the **Rosenbrock** function, When the dimension is 2, the maximum, minimum and average values of the iterations of AFSA-ACO are 142, 63 and 92, respectively. ACO cannot obtain optimized results.
- Based on the analysis, similar results occur to **Rastrigin** and **Chaffer** functions, which also means that the performance of the AFSA-ACO algorithm is better than that of the ACO method. Details are shown in **Table 3**.

Table. 3 Test results of ACO and AFSA-ACO

Test function	Dimension	Algorithm	Max value of iteration	Min value of iteration	Average value of iteration	Convergence rate
Sphere	10	ACO	175	154	162	1
		AFSA-ACO	16	3	7.4	1
	20	ACO	252	134	175	1
		AFSA-ACO	42	19	28.6	1
	50	ACO	477	365	397	1
		AFSA-ACO	150	72	101.3	1
Rosenbrock	2	ACO	-	-	-	-
		AFSA-ACO	142	63	92	1
Rastrigin	10	ACO	397	303	351	1
		AFSA-ACO	172	119	148	1
	20	ACO	742	651	683	0.84
		AFSA-ACO	477	391	435	1
	50	ACO	-	-	-	-
		AFSA-ACO	1246	847	1037	0.87
Schaffer	2	ACO	1105	891	965	0.77
		AFSA-ACO	71	27	52	1

Remark: Using the above experiments, we compare the maximum, minimum and average iteration values of these two optimization algorithms. The developed AFSA-ACO algorithm performed better under all condition. Therefore, the developed AFSA-ACO algorithm was used to construct the developed combined model.

3.2.5 Experiment: Test the performance of ACO and AFSA-ACO

To evaluate the exploration and exploitation capability of the AFSA-ACO, a metric named excellent rate (ER_{index}) is defined in this paper, which can provide the performance information of the optimization algorithm:

$$ER_{index} = \frac{n_{better}}{N_{test}} \times 100\% \tag{18}$$

where n_{better} denotes the number that AFSA-ACO has better results than ACO. N_{test} denotes the total experiment times.

In the experiment, several cases from original wind speed series are random selected to conduct the comparison test between the ACO and AFSA-ACO, and each case was experimented 100 times. The typical results are shown as **Table 4**. As shown in **Table 4**, the ER_{index} of AFSA-ACO are commonly as 90%-95% and the best result can reach to 97%. Thus, the developed optimization algorithm AFSA-ACO can significantly improve the exploration capability.

Table 4 The test of comparison of ACO and AFSA-ACO

Case	case 1	case 2	case 3	case 4	case 5	case 6	case 7	case 8	case 9	case 10
ER_{index}	92%	90%	89%	92%	97%	88%	94%	90%	92%	93%

4. Results and analysis

To avoid accidental errors and increase the reliability of the models, each experiment was repeated twelve times and the average value was considered as the forecasting performance of one model. All of the simulations operated in the following computing environment as MATLAB R2012b on Windows 8 with a 1.80 GHz(4CPUs) Intel Core i7 4500U, 64 bits and 8 GB RAM.

4.1 Data analysis and preprocessing

The wind speed data indicate a characteristic of indeterminacy. Based on detailed analysis, several characteristics can be summarized:

- An alteration of wind speed led to stronger randomness, with the annual wind speed time series, which showed no significant trend.
- In the wind speed time series, both linear and nonlinear data characteristic appeared; the volatility, intermittency and randomness were all observed.
- The wind speed data presented different change features between morning and night. An obvious variation trend can be observed in the data at night, and the diurnal data are different.

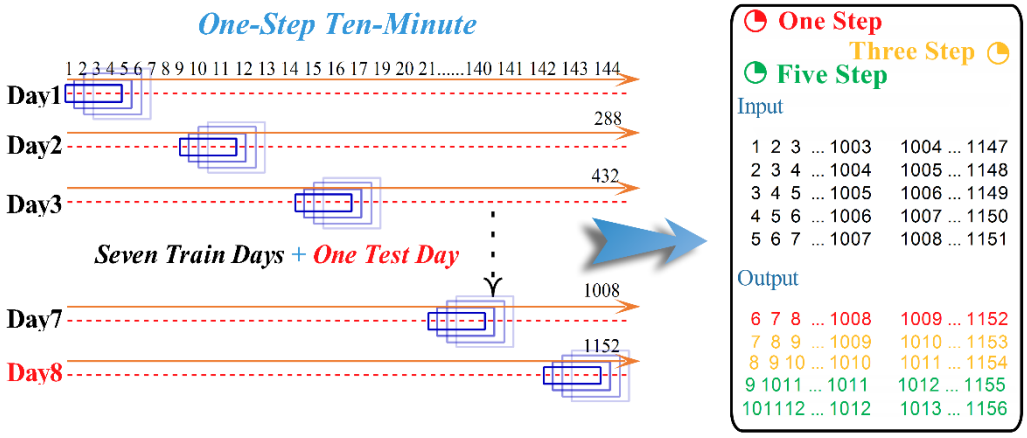


Fig. 3 The process of longitudinal data selection

For this case studies, the data were randomly selected from the ten-minute wind speed data from 2011 in Shandong, China. **Fig. 3** shows the process of the longitudinal data selection procedure for the ten-minute wind speed data as well as the input/output dataset for one-step and multi-step forecasting. In addition, **Fig. 4** provides a brief description of the experiment data used in this paper. For the experiments, ***first***, the wind speed data were divided into several series: a data series included eight days, and some of series were randomly selected to conduct the experiment (the experiment were based on both the linear and nonlinear model). ***Second***, the final results are the average values of these experiments and were divided into the diurnal dataset and nightly dataset for simulation and analysis.

Moreover, to develop an optimized wind speed forecasting model and achieve a better forecasting performance, considering the characteristics of the original wind speed data, some data preprocessing technique that have successfully used in several hybrids forecasting model have been used in the first stage to reduce the negative influence of the feature of random disturbance in the original data. In this paper, the singular spectrum analysis (SSA) is introduced as a decomposing and de-nosing technique to decompose the original data into filtered data in the data preprocessing stage. The details of SSA can be seen in [19].

4.2 The performance metric

Many performance metrics have been employed to verify the forecasting ability of different models. However, there is no universal standard for error evaluation [28]. Therefore, as generally recognized criteria, three evaluation standards were adopted to show and compare the forecasting effectiveness of the models, as shown in **Table 5**.

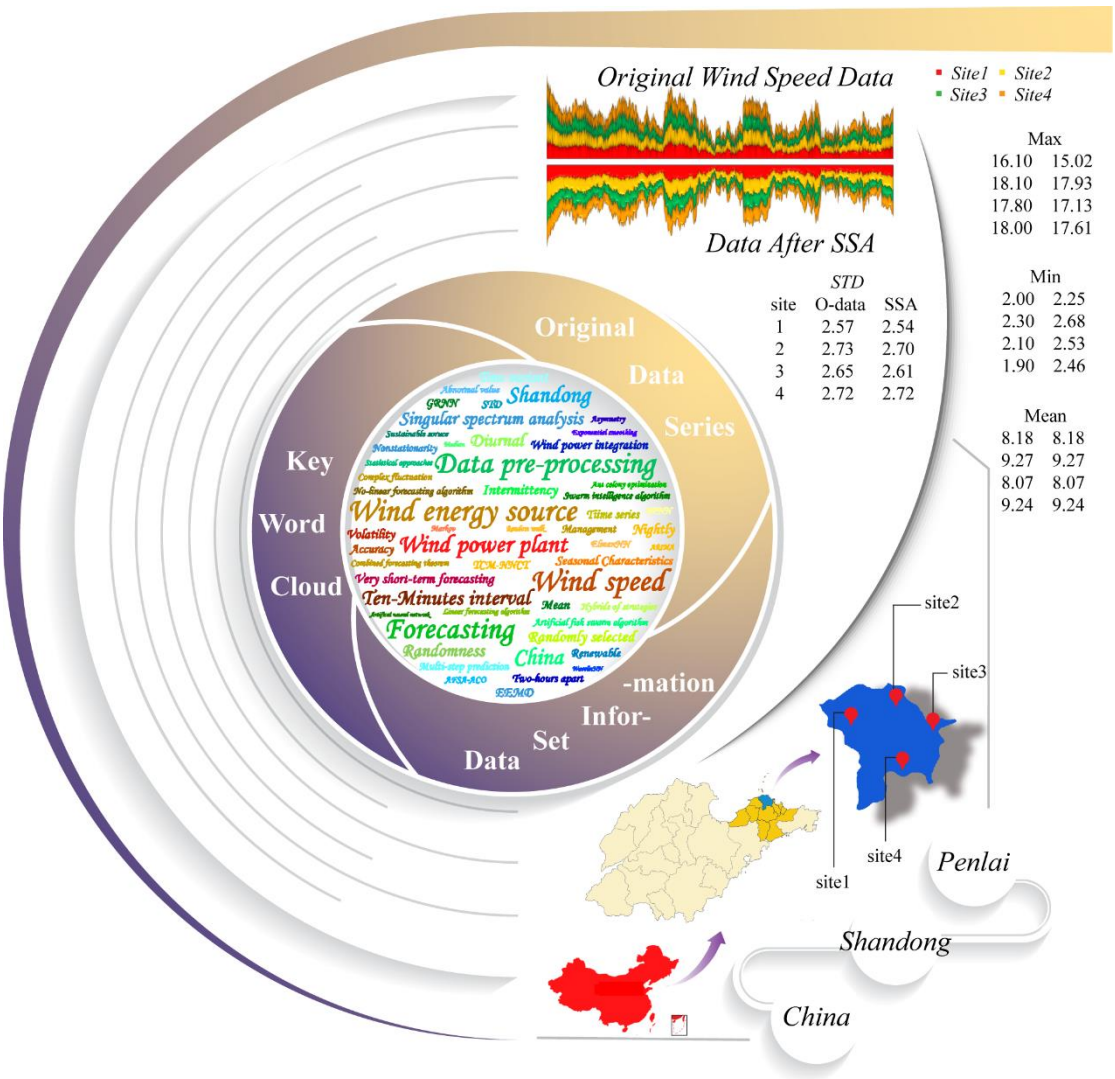


Fig. 4 A brief description of the data used in this paper

Table 5 Three metric rules for model performance

Standard	Definition	Equation
MAE	The mean absolute error of *n* times prediction results	$\mathbf{MAE}=\frac{1}{N}\sum_{i=1}^{N}\left\lvert y_i-\hat{y}_i\right\rvert$
MSE	The mean square error of *n* times prediction results	$\mathbf{MSE}=\frac{1}{N}\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2$
MAPE	The mean absolute percentage error of *n* times prediction results	$\mathbf{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{y_i-\hat{y}_i}{y_i}\right\rvert\times 100\%$

where the observed and forecasting values are denoted by y_i and $\hat{y}_i$, and N is the length of testing data. The MAE shows the level of similarity between the forecasting values and the actual values, the change in the forecasting model is evaluated by MSE. The MAPE is a unit-free standard and is sensitive to small changes in the results. Moreover, a larger value of the above metric indicates worse accuracy.

4.3 The experiment with nonlinear algorithm

To capture the nonlinear characteristic of the wind speed data, four artificial neutral networks models, i.e., BPNN, GRNN, ENN and WNN, are proposed for wind speed forecasting in this section.

4.3.1 The forecasting results of the BPNN model

To obtain the number of artificial neurons, the trial and error method [29] is used in this paper. The experiment parameters of BPNN model are selected through many experiments, which are shown in **Table 6**.

Table 6 The experiment parameters of BPNN

Experimental parameters	Default value
The number of input layer neurons	5
The number of hidden layer neurons	1-11
The number of output layer neurons	1;3;5
The training velocity	0.1
The maximum number of iteration	1000
Object accuracy	0.00004

The typical forecasting results of BPNN model are given in **Table 7**. Several conclusions can be summarized as follows:

- For different time periods in one day, the BPNN model presents different forecasting accuracies, which indicates that wind speed has the feature of volatility, intermittent and instability. Taking one-step forecasting as an example, at site 1, the minimum MAPE value for one day is 1.99% at 22:00-23:50, its maximum is 7.06% at 10:00-11:50, and its daily average the is 4.64%.
- For different sites, the BPNN model shows a relatively smaller difference in terms of forecasting performance, which is within the acceptable range. For example, the average MAPE values of the BPNN model for one-step forecasting are 4.64%, 4.55%, 4.87% and 3.87% corresponding to the MAPE values for three-step forecasting of 5.88%, 5.30%, 6.08% and 5.51% from site 1 to site 4, respectively.
- Based on the daily average forecasting accuracy in different sites, the averaged value of MAPE can be considered the performance of one model. Therefore, the MAPE value are 4.48%, 5.69% and 8.78% for one, three and five-step forecasting, respectively.

Remark: Although the average forecasting accuracy is relative better, the BPNN model presents different forecasting performances in different time periods. Therefore, we can conclude that the BPNN model is unstable for wind speed forecasting.

478 **Table 7** Typical results of the BPNN model for multi-step wind speed forecasting

Time node			0:00	2:00	4:00	6:00	8:00	10:00	12:00	14:00	16:00	18:00	20:00	22:00	Daily
Step	**Metric**	**Site**	**1:50**	**3:50**	**5:50**	**7:50**	**9:50**	**11:50**	**13:50**	**15:50**	**17:50**	**19:50**	**21:50**	**23:50**	**average**
One-step ahead	MAE	site1	0.2895	0.3413	0.1993	0.2115	0.2972	0.3712	0.2847	0.3503	0.2408	0.1475	0.3007	0.2218	0.2713
		site2	0.2841	0.2516	0.206	0.2055	0.3671	0.3322	0.3346	0.3235	0.3425	0.2472	0.4895	0.3034	0.3073
		site3	0.1353	0.1715	0.1707	0.2235	0.2879	0.3919	0.221	0.3272	0.3086	0.1831	0.5613	0.2962	0.2732
		site4	0.2088	0.1362	0.1608	0.1314	0.1881	0.3568	0.2256	0.2564	0.3614	0.24s91	0.4215	0.5211	0.2681
	MAPE	site1	6.51	6.3	2.81	2.81	5.84	7.06	6.56	6.65	3.44	2.23	3.49	1.99	4.64
		site2	5.41	4.6	3.01	2.56	6.41	6.26	6.14	5.13	4.37	2.94	5.31	2.43	4.55
		site3	2.39	4.06	3.35	3.51	5.5	9.68	5.42	6.81	5.31	2.96	6.95	2.49	4.87
		site4	4.45	2.9	2.44	1.73	3.51	7.64	3.86	3.98	4.29	2.89	4.42	4.27	3.87
	MSE	site1	0.1035	0.1989	0.0626	0.0579	0.1608	0.2732	0.1159	0.1852	0.0999	0.0317	0.1424	0.0709	0.1252
		site2	0.1042	0.0779	0.0703	0.0677	0.1907	0.1638	0.1365	0.1633	0.2057	0.0788	0.294	0.1399	0.1411
		site3	0.0295	0.0465	0.0478	0.0914	0.1131	0.2345	0.0741	0.1752	0.1494	0.0502	0.3819	0.1132	0.1256
		site4	0.0679	0.0292	0.0437	0.0232	0.066	0.1836	0.0762	0.0995	0.1783	0.0972	0.2159	0.3806	0.1218
Three-step ahead	MAE	site1	0.3809	0.3341	0.3053	0.1974	0.4523	0.3904	0.3776	0.4909	0.3162	0.1812	0.5198	0.3158	0.3552
		site2	0.3065	0.2767	0.2753	0.3234	0.4251	0.3473	0.4346	0.3784	0.3621	0.3074	0.6098	0.3869	0.3695
		site3	0.2041	0.2238	0.2812	0.2801	0.4297	0.4199	0.2657	0.4078	0.3035	0.2697	0.6514	0.3682	0.3421
		site4	0.2752	0.2342	0.3222	0.2134	0.3115	0.455	0.4968	0.296	0.4084	0.316	0.5948	0.5501	0.3728
	MAPE	site1	7.56	6.35	4.31	2.62	8.87	7.56	8.01	9.41	4.52	2.8	5.78	2.81	5.88
		site2	5.69	4.99	3.82	4.17	7.44	6.47	7.89	5.79	4.46	3.79	6.02	3.1	5.3
		site3	3.63	5.64	5.62	4.33	8.76	10.15	6.54	8.24	5.2	4.48	7.25	3.1	6.08
		site4	5.9	5.11	5	2.89	5.92	9.53	8.5	4.52	4.66	3.79	5.83	4.52	5.51
	MSE	site1	0.2776	0.1699	0.1327	0.0606	0.3764	0.2787	0.2275	0.3815	0.1572	0.0536	0.5783	0.1538	0.2373
		site2	0.1292	0.1065	0.1488	0.1684	0.2799	0.1846	0.2779	0.2211	0.2021	0.1362	0.6202	0.2608	0.228
		site3	0.0747	0.0876	0.1261	0.1589	0.2963	0.2739	0.1151	0.2617	0.1535	0.1026	0.7028	0.1744	0.2106
		site4	0.134	0.104	0.1822	0.0838	0.1721	0.2962	0.4328	0.1491	0.2158	0.1585	0.6475	0.4545	0.2526
Five-step ahead	MAE	site1	0.4723	0.3806	0.4581	0.3309	0.6648	0.4168	0.6778	0.6689	0.416	0.2899	1.1385	0.5154	0.5358
		site2	0.3741	0.4061	0.5466	0.5572	0.6009	0.4022	0.8178	0.5323	0.5608	0.4083	0.9742	0.6084	0.5657
		site3	0.3931	0.417	0.5527	0.3831	0.7186	0.5507	0.4672	0.5683	0.3911	0.3924	1.086	0.5334	0.5378
		site4	0.4388	0.4694	0.5452	0.3527	0.6237	0.6274	0.9474	0.4717	0.5739	0.5226	1.0543	0.7151	0.6118
	MAPE	site1	9.04	7.09	6.39	4.6	13.26	8.59	13.8	12.74	5.9	4.55	11.63	4.53	8.51
		site2	6.84	7.13	7.13	7.66	10.83	7.56	14.74	7.88	6.68	5.24	9.06	4.81	7.96
		site3	6.9	11.17	10.31	6.12	15.68	13.02	11.59	10.88	6.8	6.73	10.97	4.44	9.55
		site4	9.31	9.79	8.18	4.94	11.95	13.25	16.43	6.88	6.31	6.57	9.77	5.8	9.1
	MSE	site1	0.4785	0.2318	0.385	0.1906	0.7594	0.302	0.9534	0.7514	0.2685	0.1613	2.7962	0.4543	0.6444
		site2	0.2166	0.2832	0.553	0.6276	0.6418	0.2648	1.0753	0.485	0.4905	0.281	1.9305	0.7473	0.633
		site3	0.2993	0.3574	0.5667	0.2737	0.9434	0.4791	0.3913	0.5311	0.2961	0.2359	2.6131	0.4638	0.6209

	site4	0.3684	0.4826	0.5327	0.2603	0.6519	0.6308	1.7202	0.4073	0.524	0.4705	2.367	0.8355	0.7709

4.3.2 The forecasting results of the GRNN model

The experiment parameters of the GRNN model are shown in **Table 8**. Moreover, the radial basis function expansion range is [0.1, 2.0], which is performed to establish the superior GRNN model. The one-step and multi-step forecasting results in diurnal and nightly, are shown in **Table 9**. For one-step forecasting, the GRNN model has a minimum MAPE value of 6.42% in the nightly wind speed, compared to the MAPE of 8.62% for the diurnal speed; we can reach the same conclusion for the multi-step forecasting.

Remark: The GRNN model performs better than the BPNN model in terms of forecasting stability, but it is worse than the BPNN model regarding accuracy. Moreover, the forecasting accuracy of the GRNN model for the nightly wind speed forecasting is better than the diurnal, which may account for the diurnal wind speed change being more drastic than the nightly change.

Table 8 Experimental parameters of GRNN

Experimental parameters	Default value
iteration time	100
learning rate	0.1
training requirement accuracy	0.00004

Table 9 Forecasting results of GRNN

Diurnal		MAE	MAPE	MSE	Nightly		MAE	MAPE	MSE
	Site1	0.5014	9.58	0.4125		Site1	0.5033	7.04	0.4620
One	Site2	0.5259	8.62	0.4163	**One**	Site2	0.5157	6.42	0.5021
-step	Site3	0.4935	10.46	0.4120	**-step**	Site3	0.5424	8.25	0.6330
	Site4	0.5943	9.88	0.4903		Site4	0.5695	8.16	0.6221
	Site1	0.6385	12.22	0.6941		Site1	0.6811	9.22	0.9228
Three	Site2	0.6733	10.81	0.6857	**Three**	Site2	0.7689	9.50	1.0908
-step	Site3	0.6676	14.68	0.7487	**-step**	Site3	0.8044	12.58	1.3454
	Site4	0.7849	12.93	0.8520		Site4	0.7314	10.46	1.0706
	Site1	0.7369	13.97	0.9373		Site1	0.8034	10.57	1.2923
Five	Site2	0.8031	12.83	0.9494	**Five**	Site2	0.9386	11.63	1.6297
-step	Site3	0.7691	16.78	0.9395	**-step**	Site3	0.9952	15.25	2.1607
	Site4	0.9664	15.92	1.3716		Site4	0.9057	12.79	1.5458

4.3.3 The forecasting results of the ENN model

In this case, the trial and error method are also used to obtain the best number of neurons. The best results are selected based on many experiments. The experimental parameters of ENN are presented in **Table 10**. To research the performance of the ENN model, the forecasting results (MAE, MAPE and MSE) of the ENN model for wind speed forecasting in diurnal and nightly speeds are shown in **Table 11**.

Remark: The experiment results reveals that the ENN model is also unstable and performs worse than GRNN and BPNN in terms of forecasting accuracy. Similar to BPNN and GRNN, the forecasting accuracy of the ENN model for nightly wind speed forecasting is better than that for diurnal, which may further prove that the diurnal wind speed change is more drastic than the nightly change.

506 **Table 10** The experimental parameters of ENN

Experimental parameters	Default value
The number of input layer neurons	5
The number of hidden layer neurons	1-11
The number of output layer neurons	1;3;5
iteration time	100
learning rate	0.1
training requirement accuracy	0.00004

507 **Table 11** The typical results of ENN

Diurnal		MAE	MAPE	MSE	Nightly		MAE	MAPE	MSE
	Site1	0.4416	8.16	0.3225		Site1	0.4249	6.05	0.3306
One	Site2	0.4516	7.32	0.2962	**One**	Site2	0.4614	5.77	0.3608
-step	Site3	0.3884	8.00	0.2353	**-step**	Site3	0.3717	5.79	0.2599
	Site4	0.4342	7.18	0.2859		Site4	0.4443	6.41	0.3684
	Site1	0.5534	10.43	0.5443		Site1	0.5313	7.32	0.6133
Three	Site2	0.5775	9.27	0.4916	**Three**	Site2	0.6571	7.92	0.8611
-step	Site3	0.5189	11.12	0.4607	**- step**	Site3	0.6494	10.37	0.9331
	Site4	0.6071	9.77	0.5637		Site4	0.5799	7.99	0.7575
	Site1	0.6898	12.93	0.8248		Site1	0.7156	9.38	1.1659
Five	Site2	0.6736	10.65	0.7083	**Five**	Site2	0.8255	9.89	1.4404
-step	Site3	0.6811	14.69	0.7825	**-step**	Site3	0.8536	13.28	1.6686
	Site4	0.7701	12.39	0.9301		Site4	0.7329	9.87	1.2239

508 4.3.4 The forecasting results of WNN

509 To establish the WNN model, the detailed parameters of WNN are set as shown
510 in **Table 12**. The typical results for one-step and multi-step forecasting at four sites
511 are presented in **Table 13**.
512 **Remark**: The forecasting stability of WNN is worse than the other three nonlinear
513 models. The forecasting performance are sometimes unacceptable and sometimes
514 acceptable in certain condition; however, even the best forecasting results are also
515 worse than those of the other neural networks.

516 **Table 12** The experimental parameters of WNN

Experimental parameters	Default value
iteration time	100
learning rate	0.1
training requirement accuracy	0.00004

517 **Table 13** The typical results of WNN

Diurnal		MAE	MAPE	MSE	Nightly		MAE	MAPE	MSE
	Site1	0.5900	12.16	0.5552		Site1	0.5178	8.68	0.4805
One-	Site2	1.0150	16.38	1.5297	**One**	Site2	1.0531	12.72	2.0930
step	Site3	0.4853	10.40	0.3712	**-step**	Site3	0.5138	8.37	0.5242
	Site4	0.8425	13.66	1.0167		Site4	0.8717	12.27	1.4864
Three	Site1	0.6638	12.56	0.7412	**Three**	Site1	0.6509	8.83	0.8444
- step	Site2	0.6565	10.44	0.6525	**-step**	Site2	0.7671	9.35	1.1716

	Site3	0.8274	18.00	1.0716		Site3	1.0144	15.98	2.0813
	Site4	0.7396	11.94	0.8109		Site4	0.7110	9.86	1.0524
	Site1	0.7222	13.60	0.8903		Site1	0.7644	10.06	1.2491
Five	Site2	0.7093	11.28	0.7711	**Five**	Site2	0.8772	10.46	1.6118
-step	Site3	0.7299	15.67	0.8703	**-step**	Site3	0.9132	14.15	1.8394
	Site4	0.8497	13.72	1.0879		Site4	0.8033	10.87	1.3907

4.4 The experiment with the linear algorithm

To forecast the linear characteristic of the wind speed data, two linear algorithms, including ARIMA and ES, are adopted as the linear section in the developed combined forecasting model.

4.4.1 The forecasting results of ARIMA

To obtain the best parameter to adapt to ARIMA, the parameters *p, q* and *r* are used with repeat trials from 1 to 15, and the parameter *r* is cycled from 1-4. The typical results of ARIMA of the wind speed data at diurnal and nightly moments using the one-step and multi-step are presented in **Table 14.** The maximum values of diurnal MAPE and nightly MAPE are 4.24% and 3.36% for one step, and the minimum value are 5.89% and 3.98%, respectively.

4.4.2 The forecasting results of ES

To obtain the best parameter to adapt to ES， the parameter α is used with repeat trials from 0.01 to 0.99. The typical results of the ES of the wind speed data at diurnal and nightly moments on multi-steps are presented in **Table 14**.

Remark: The forecasting stability of ARIMA is more stable than ES but worse than BPNN and GRNN, and the forecasting performance is higher than that of ES and ENN but is lower than BPNN and GRNN.

536 **Table 14** Typical results of ARIMA and ES for wind speed forecasting

		One step ahead				Three-step ahead				Five-step ahead			
		ARIMA		ES		ARIMA		ES		ARIMA		ES	
		Diurnal	Nightly	Diurnal	Nightly	Diurnal	Nightly	Diurnal	Nightly	Diurnal	Nightly	Diurnal	Nightly
	MAE	0.2834	0.2554	0.6046	0.5255	0.3936	0.3563	0.7282	0.7500	0.7459	0.7617	0.8319	0.9036
Site1	MAPE	5.32	3.98	11.18	7.61	7.26	5.04	13.48	10.17	13.85	10.04	15.30	11.92
	MSE	0.1485	0.1001	0.6052	0.5622	0.2357	0.2289	0.9058	1.1791	0.8432	1.1202	1.1480	1.6564
	MAE	0.3251	0.2912	0.6212	0.5901	0.3692	0.4010	0.7479	0.8770	0.7901	0.8675	0.8231	1.0894
Site2	MAPE	5.22	3.91	9.88	7.31	5.77	5.04	11.76	10.62	12.34	10.39	12.75	12.87
	MSE	0.1585	0.1253	0.5821	0.7133	0.2090	0.2946	0.8550	1.5819	0.8962	1.3781	1.0827	2.3521
	MAE	0.2837	0.2550	0.6155	0.5689	0.3944	0.4320	0.7870	0.9528	0.7943	0.9877	0.8452	1.2582
Site3	MAPE	5.89	3.73	12.66	8.55	8.03	6.48	16.30	14.19	16.63	14.77	17.64	18.50
	MSE	0.1319	0.1140	0.6102	0.6742	0.2198	0.3511	0.9661	1.9252	0.8902	1.7768	1.0953	3.1395
	MAE	0.2660	0.2610	0.5595	0.5713	0.4227	0.3779	0.8077	0.7800	0.8431	0.7559	0.9547	0.9285
Site4	MAPE	4.24	3.36	9.01	7.84	6.79	4.94	12.93	10.52	13.61	10.02	15.03	12.14
	MSE	0.1124	0.1243	0.4905	0.6275	0.2568	0.2750	1.0123	1.2949	0.9775	1.1491	1.4453	1.8200

4.5 The forecasting results of the combined model and comparisons

To prove the superiority of this combined model, some experiments were conducted Three performance metrics are used to conduct a comparative study with the individual models contained in the developed combined model were conducted, the bias-variance framework is employed to further evaluate the developed combined model.

4.5.1 Experiment I: Testing the combined model with comparisons

The experiment was performed to further show that the developed forecasting model integrated with the developed AFSA-ACO algorithm could improve the wind speed forecasting performance. The individual models, including two linear algorithms (ES and ARIMA) and four nonlinear algorithms (BPNN, ENN, WNN, GRNN), were employed as comparison models to conduct a comparative study and evaluate the developed combined model. The experiment was designed to perform one-step and multi-step wind speed forecasting, and each experiment and evaluate three times for each step case. The comparative study results are shown in **Table 15 and Fig. 5.**, which shows the forecasting effectiveness for the experimental datasets by BPNN, ENN, GRNN, WNN, ARIMA, ES and the combined model in one, three and five-step forecasting in terms of three typical performance metrics: MAE, MAPE and MSE. According to the average value of the MAPE metric, whether for the one-step or multi-step forecasting, the developed combined model outperforms all benchmark models. For example, for one-step ahead forecasting, the MAPE values obtained by BPNN, ENN, GRNN, WNN, ARIMA, ES and the developed combined model are 4.71%, 7.16%, 6.34%, 9.44%, 5.25%, 10.25% and 4.48%, respectively, while the MAPE value for five-step forecasting are 8.80%, 13.46%, 12.02%, 13.17%, 13.61%, 17.02% and 8.31%, respectively. The BPNN model achieved second place, while the ES model was last for all horizon forecasting. With an increase in the forecasting horizon, the forecasting error became lager. Based on a comparison study of multi-step forecasting performance, the differences between one, three and five-step forecasting are 0.93% and 3.83%, which is within the acceptable range. Thus, we conclude that the developed combined model can be successfully and effectively applied for multi-step forecasting

4.5.2 Experiment II: Testing with bias-variance framework

The bias-variance framework was employed to evaluate the forecasting accuracy and stability of one model, which is important for verifying the model's effectiveness. The absolute value of bias is considered the average difference between the actual value and the forecasting value, and the variance reveals the model's variability. A smaller absolute value of bias reveals a higher forecasting accuracy, while a smaller variance represents better stability. The comparison results of the bias-variance framework are shown in **Table 16**. As shown in **Table 16**, the absolute values of biases and variances are all smaller than other compared models, which proves that the developed model can achieve wind speed forecasting more accurately and provides more stability. Therefore, the developed combined forecasting model, with its higher accuracy and strong stability, is superior to the other compared models and can be an efficient technique in engineering applications.

4.5.3 Experiment III: Test the performance of the combined model through forecasting horizon.

To explore how the performance of the developed model decrease through the forecasting time horizon at the four examined sites, 1-tesp to 24-step ahead forecasting experiments were conducted in this section and the MAPE results are shown in **Table 17**. Two particular times (8:00 and 20:00) were chosen as the beginning of the prediction to

conduct the experiment. In the next 4 hours forecasting horizon, the MAPE value of the Test-1 obtained from 1-step to 24-step are 4.50%, 8.80%, 14.79%, 20.33% and 30.87%, respectively, and the MAPE value of Test-2 are 4.67%, 9.13%, 15.45%, 18.46% and 31.79%, respectively. By making a comparison from 1-step to 24-step, the greater the number of steps of the forecast, the lower the accuracy. The performance of the combined model form 1-step to 12-step can achieve relatively satisfied performance, however, when forecasting horizon beyond 12-step, the MAPE of the proposed model decreasing quickly. Thus, we can affirm that the developed combined model can be effectively applied for multi-step forecasting.

,

Table 17 The MAPE of combined model forecasted through the time horizon

	forecast horizon	stie1	site2	site3	site4	mean
Test-1	8:10(1-step)	4.57	4.32	4.82	4.29	4.50
	9:00(6-step)	8.89	8.17	9.39	8.75	8.80
	10:00(12-step)	14.42	14.94	15.62	14.18	14.79
	11:00(18-step)	20.97	21.08	20.33	18.94	20.33
	12:00(24-step)	31.14	30.56	32.41	29.37	30.87
Test-2	20:10(1-step)	4.68	4.49	4.97	4.54	4.67
	21:00(6-step)	9.37	8.42	10.03	8.71	9.13
	22:00(12-step)	15.56	14.76	15.87	15.62	15.45
	23:00(18-step)	22.87	20.37	20.42	10.17	18.46
	24:00(24-step)	32.64	30.42	33.17	30.93	31.79

4.5.4 Summary: Based on experiments I–III

Based on the above two experiments, we draw the following conclusions:

- The developed combined model based on linear and nonlinear features can provide effective wind speed forecasting results with higher accuracy and strong stability, so it can be widely used in actual application.
- The combined model can take full use of the advantages of each model and make each individual forecasting model perform better in certain cases, which can effectively forecast future wind speed compared with the individual models.
- A modified optimization algorithm, AFSA-ACO, is developed in this paper to optimize the weight of each individual model, which can further improve the forecasting performance of the developed model.
- The combined model including two linear forecasting models and four nonlinear forecasting models, is successfully employed to separately capture the inherent feature of linear and nonlinear characteristics contained in wind speed and ultimately results in success. Thus, it has a superior forecasting ability and will be widely used in the forecasting fields.

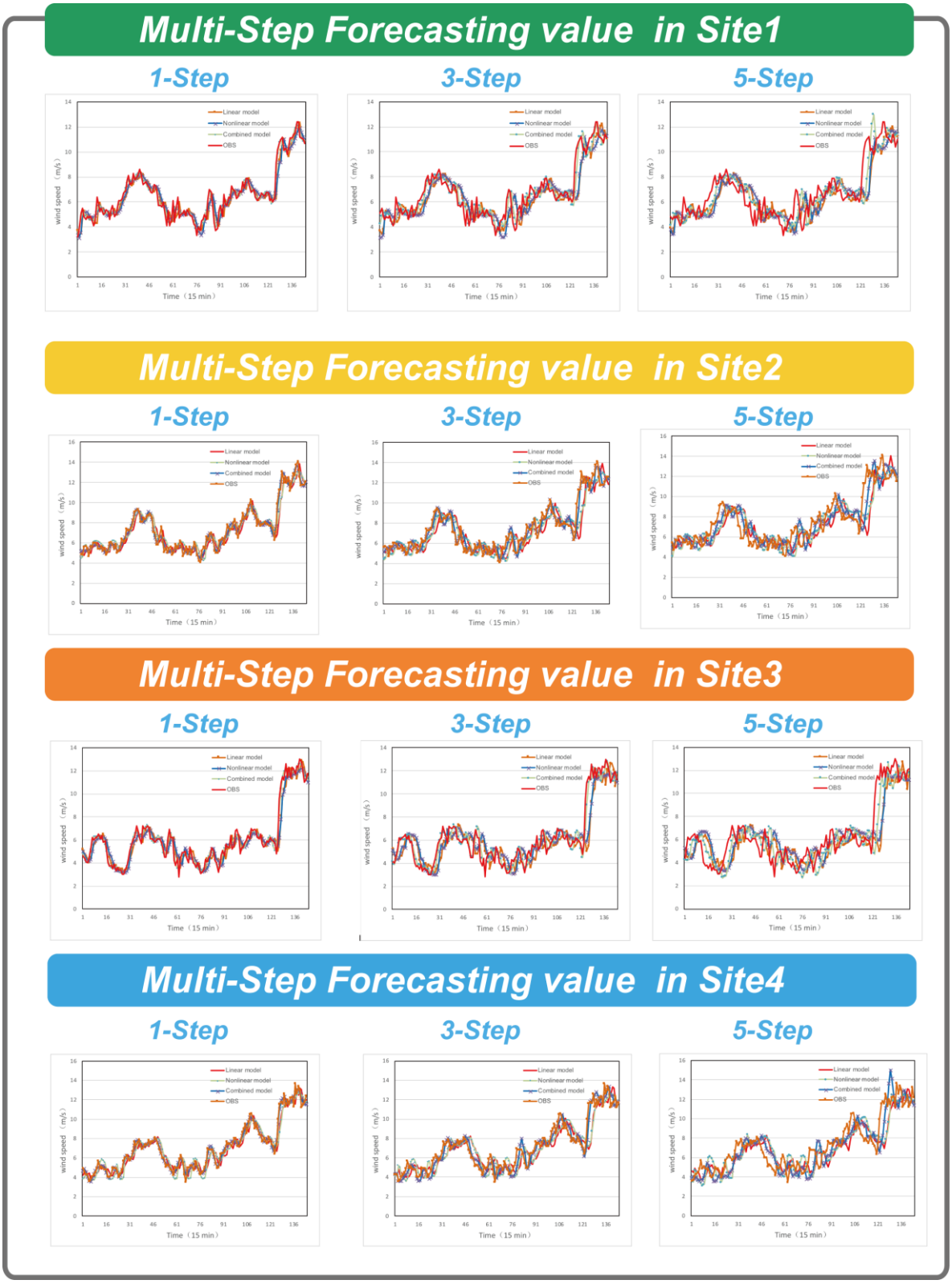


Fig. 5. Forecasted values of four models.

615 **Table 15** Typical results of the combined model and the results of the other models

Model	Test with 1-step			Average	Test with 3-step			Average	Test with 5-step			Average
					MAE							
BP	**0.3073**	**0.3574**	0.3229	**0.3292**	**0.3823**	**0.4117**	0.3756	**0.3899**	**0.5723**	**0.5542**	0.5871	0.5712
ENN	0.5144	0.4874	0.5331	0.5116	0.5830	0.7538	0.6421	0.6596	0.7491	0.8426	0.8857	0.8258
GRNN	0.4611	0.4862	0.4792	0.4755	0.7126	0.6857	0.6643	0.6875	0.8651	0.9311	1.0635	0.9532
WNN	0.5857	0.5671	0.5362	0.5630	0.7634	0.8214	0.7963	0.7937	0.8269	0.8869	1.0326	0.9155
ARIMA	0.2972	0.3642	0.3971	0.3528	0.4286	0.5542	0.4831	0.4886	0.8631	0.9017	0.8517	0.8722
ES	0.6230	0.6541	0.5628	0.6133	0.8542	0.9423	0.7988	0.8651	1.1354	1.0823	1.4529	1.2235
Combined	0.3157	0.3627	**0.3129**	0.3304	0.3942	0.4217	**0.3544**	0.3901	0.6133	0.5648	**0.5219**	**0.5667**
					MAPE							
BP	4.55	**4.86**	4.71	4.71	5.82	6.27	5.64	5.91	8.87	8.41	9.12	8.80
ENN	7.42	6.93	7.12	7.16	8.62	9.74	8.63	9.00	13.42	13.22	13.75	13.46
GRNN	6.24	6.37	6.42	6.34	10..11	9.27	8.94	9.44	11.47	11.86	12.73	12.02
WNN	10.11	9.43	8.79	9.44	10.29	11.57	11.36	11.07	12.57	12.86	14.07	13.17
ARIMA	4.86	5.18	5.72	5.25	7.09	7.24	6.93	7.09	13.57	14.26	13.01	13.61
ES	10.37	10.54	9.84	10.25	12.94	13.54	12.37	12.95	16.89	15.74	18.42	17.02
Combined	**4.39**	4.93	**4.12**	**4.48**	**5.39**	**5.61**	**5.24**	**5.41**	**8.62**	**8.35**	**7.96**	**8.31**
					MSE							
BP	**0.1411**	0.1721	0.1645	0.1592	0.2119	**0.2327**	0.2047	**0.2164**	0.6742	**0.5680**	0.6849	0.6424
ENN	0.4673	0.5132	0.4865	0.4890	0.6714	0.8216	0.7468	0.7466	1.3514	1.1967	1.2532	1.2671
GRNN	0.2762	0.2886	0.2931	0.2860	0.8453	0.7810	0.7246	0.7836	1.2357	1.3392	1.4531	1.3427
WNN	0.9263	0.8446	0.7597	0.8435	0.9582	1.1052	1.0596	1.0410	1.3987	1.3547	1.5017	1.4184
ARIMA	0.1571	0.1669	0.1764	0.1668	0.3129	0.4186	0.3562	0.3626	1.3619	1.4358	1.2956	1.3644
ES	0.8741	0.8452	0.7599	0.8264	1.2576	1.3017	1.2294	1.2629	2.1047	1.9351	2.5033	2.1810
Combined	0.1475	**0.1582**	**0.1463**	**0.1507**	**0.2071**	0.2483	**0.1984**	0.2179	**0.6420**	0.5939	**0.5781**	**0.6047**

616
617

618 **Table 16** Results of the bias-variance framework regarding the combined model and other models

	Site 1						
	Combined model	BPNN	ENN	WNN	GRNN	ARIMA	ES
Bias	**-37.2472**	40.8293	208.24407	154.46983	110.4698	94.2908	156.09009
Variance	**$2.56186*10^6$**	$2.23397*10^6$	$5.57679*10^6$	$4.67358*10^6$	$3.78945*10^6$	$3.65556*10^6$	$3.68538*10^6$
	Site 2						
	Combined model	BPNN	ENN	WNN	GRNN	ARIMA	ES
Bias	**-26.0644**	39.0858	-150.1512	-73.7867	-131.0260	-87.39465	-138.86687
Variance	**$1.78499*10^5$**	$2.94337*10^6$	$6.79984*10^5$	$3.37324*10^5$	$4.49748*10^6$	$3.82120*10^5$	$4.84572*10^5$
	Site 3						
	Combined model	BPNN	ENN	WNN	GRNN	ARIMA	ES
Bias	**-29.8109**	47.5106	209.0871	71.9564	-102.2839	-57.0692	128.0058
Variance	**$2.02388*10^6$**	$2.65826*10^6$	$7.10694*10^6$	$3.87078*10^6$	$4.30894*10^6$	$3.12246*10^6$	$4.14463*10^6$
	Site 4						
	Combined model	BPNN	ENN	WNN	GRNN	ARIMA	ES
Bias	**31.1363**	-48.1986	-117.1073	-117.7205	-100.2213	54.8568	134.0037
Variance	**$2.57589*10^5$**	$3.34525*10^6$	$6.06825*10^5$	$4.57947*10^5$	$3.71257*10^5$	$3.19408*10^5$	$3.72514*10^5$

5. Discussion

An insightful discussion about the developed model is conducted in this section, including the improvements of the developed model, the significance of the developed model, the effectiveness of the combined strategy and the computation time.

5.1 The improvements of the developed combined model

To conduct a comprehensive analysis for the developed model, a metric, the decreased relative error of index (DRE_{index}) [30]**,** is employed in this paper to provide the improvements of information of the developed combined model, as how much the DRE_{index} is means how much promotion between the two compared model, which is defined as

$$DRE_{index} = \frac{index_i - index_j}{index_i} \tag{18}$$

In this paper, we take the MAPE index as an illustration. The detailed results are presented in **Table 18**. According to **Table 18**, the developed combined model shows a great improvement compared with the other models. For example, for one-step forecasting, the developed combined model leads to improvements of 4.8832%, 37.4302%, 29.3375%, 52.5424%, 14.6667% and 56.2927% in terms of the average MAPE compare with BPNN, ENN, GRNN, WNN, ARIMA and ES, respectively. The maximum decreased error of the average MAPE is 58.2239%, and the minimum is 4.8832%. Thus, the developed combined forecasting model is superior to other compared models and can achieve a satisfactory and effective forecasting performance when it is applied in engineering applications.

Table 18 Improvements of the developed combined model for multi-step forecasting

Times	**Combined model vs. BPNN**			**Combined model vs. ENN**			**Combined model vs. GRNN**		
	1-step	**3-step**	**5-step**	**1-step**	**3-step**	**5-step**	**1-step**	**3-step**	**5-step**
1	3.5165	7.3883	2.8185	40.8356	37.4710	35.7675	29.6474	46.6864	24.8474
2	-1.4403	10.5263	0.7134	28.8600	42.4025	36.8381	22.6060	39.4822	29.5953
3	12.5265	7.0922	12.7193	42.1348	39.2816	42.1091	35.8255	41.3870	37.4705
Average	4.8832	8.4602	5.5682	37.4302	39.8889	38.2615	29.3375	42.6907	30.8652
Times	**Combined model vs. WNN**			**Combined model vs. ARIMA**			**Combined model vs. ES**		
	1-step	**3-step**	**5-step**	**1-step**	**3-step**	**5-step**	**1-step**	**3-step**	**5-step**
1	56.5776	47.6190	31.4240	9.6708	23.9774	36.4775	57.6663	58.3462	48.9639
2	47.7200	51.5125	35.0700	4.8263	22.5138	41.4446	53.2258	58.5672	46.9504
3	53.1286	53.8732	43.4257	27.9720	24.3867	38.8163	58.1301	57.6395	56.7861
Average	52.5424	51.1292	36.9021	14.6667	23.6953	38.9420	56.2927	58.2239	51.1751

5.2 The significance of the developed combined model

In this section, a widely employed hypothesis testing approach, the Diebold-Mariano Test [31], is applied to evaluate the significance of the developed combined model with other comparison models. The specific theory of this method can be observed as follows.

Under a given a significance level, α , the null hypothesis, H_0 represents that there is no significant difference between the proposed model and the compared

model in terms of their forecasting performance. In contrast, H_1 is different from H_0. The related formulas can be written as

$$\begin{cases} H_0 : E\left[l\left(e_i^1\right)\right] = E\left[l\left(e_i^2\right)\right] \\ H_1 : E\left[l\left(e_i^1\right)\right] \neq E\left[l\left(e_i^2\right)\right] \end{cases} \tag{19}$$

wherein l is the loss function of the forecasting error and e_i^j, $j = 1, 2$ indicates the prediction errors of the comparison models and the proposed model.

Moreover, the DM test statistics can be described as:

$$DM = \frac{\sum_{i=1}^{n}\left(l\left(e_i^1\right) - l\left(e_i^2\right)\right)/n}{\sqrt{S^2/n}} s^2 \tag{20}$$

where s^2 is an estimation for the variance of $D_i = l\left(e_i^1\right) - l\left(e_i^2\right)$.

Finally, under the given the significance level α, are employed to compared with the, if the DM statistic

If the computed values of DM fall outside the interval $[-Z_{\alpha/2}, Z_{\alpha/2}]$, the null hypothesis H_0 will be rejected. This case means that there is no significant difference in the forecasting performance between the proposed model and the compared model. Otherwise, the null hypothesis will be accepted.

Table 19 Diebold-Mariano test between the comparison models and the proposed model

DM test	1-step	3-step	5-step
BPNN	1.8236541***	1.7334923***	1.9230125***
ENN	2.2239706**	2.2854941**.	2.3680649**
GRNN	1.7582973***	1.6247201	1.8496255***
WNN	2.1593657**	2.2064192**	2.4723091**
ARIMA	2.2453762**	2.3981643**	2.5491064**
ES	6.8723623*	7.12681234*	7.2790154*

* is the 1% significance level; ** is the 5% significance level; *** is the 10% significance level;

From **Table 19,** it can be easily observed that the values of |DM| are all much greater than $Z_{0.05} = 1.645$. Finally, we conclude that the null hypothesis can be accepted T the 10% significance level, which also means there is a 90% probability that the alternative hypothesis will be accepted. This case shows that the proposed combined method has a significant difference with the comparison approaches with the significance level 90%. In other words, our proposed combined model exhibits a significant improvement in performance compare with other models.

5.3 The effectiveness of the combined strategy of the developed combined model

To verify the effectiveness of the combined strategy based on TCM-NNCT and weight-coefficient optimization, a comparison between the developed combined model and the combined model based on the simple average strategy is conducted in this section. The comparison results are shown in **Table 20,** which shows that the developed combined model shows excellent performance compared with the combined model with the simple average strategy. Therefore, we can draw the conclusion that the developed combined model, which is based on TCM-NNCT and

the modified optimization algorithm, AFSA-ACO, effectively improves the performance for wind speed forecasting.

To further evaluate the optimization capacity of the proposed combined model, several common combined models, which based on different weight-coefficient optimization algorithms (BFGS, PSO and ACO), are used to compare with the developed combined model. **Table 21** lists the MAE, MAPE and MSE values of four combined models, which are optimized by the BFGS, PSO, ACO and AFSA-ACO algorithms form four sites. The MAE, MAPE and MSE values of the developed model are less than the other combined models in most cases, which indicates that the combined model optimized by the AFSA-ACO algorithm is relatively more accurate than the other combined models that optimized by BFGS, PSO and ACO, respectively. Thus, it can be concluded that compared with other combined models, the combined modes based on AFSA-ACO can provide more accurate wind-speed forecasting.

Table 20 Comparison results of different combined strategies

Sites	Combined Strategy	MAE			MAPE			MSE		
		1-step	3-step	5-step	1-step	3-step	5-step	1-step	3-step	5-step
Site 1	Simple average strategy	0.4648	0.6207	0.8353	7.26	8.95	12.80	0.4737	0.7096	1.3544
	Developed model	**0.3157**	**0.3942**	**0.6133**	**4.39**	**5.39**	**8.62**	**0.1475**	**0.2071**	**0.6420**
Site 2	Simple average strategy	0.4816	0.6949	0.8665	7.22	9.61	12.73	0.4718	0.7768	1.3049
	Developed model	**0.3627**	**0.4217**	**0.5648**	**4.93**	**5.61**	**8.35**	**0.1582**	**0.2483**	**0.5939**
Site 3	Simple average strategy	0.4719	0.6267	0.9789	7.10	8.98	13.52	0.4400	0.7202	1.4486
	Developed model	**0.3129**	**0.3544**	**0.5219**	**4.12**	**5.24**	**7.96**	**0.1463**	**0.1984**	**0.5781**
Site 4	Simple average strategy	0.4931	0.6832	0.9231	7.28	9.35	12.91	0.4824	0.7548	1.4275
	Developed model	**0.3276**	**0.4157**	**0.6014**	**4.97**	**5.51**	**8.27**	**0.1607**	**0.2293**	**0.6371**

703 **Table 21** Comparison results of different combined models

Model	Test with 1-step			Average	Test with 3-step			Average	Test with 5-step			Average
	MAE											
BFGS-CM	0.3624	0.3953	0.3342	0.3640	0.4298	0.4679	0.3905	0.4294	0.7231	0.6874	0.6469	0.6858
PSO-CM	0.3871	0.4102	0.3571	0.3848	0.4531	0.5016	0.4176	0.4574	0.6574	0.6152	0.5826	0.6184
ACO-CM	0.4239	0.4436	0.4258	0.4311	0.5182	0.5583	0.5247	0.5337	0.8018	0.7513	0.7041	0.7524
Pro. CM	**0.3157**	**0.3627**	**0.3129**	**0.3304**	**0.3942**	**0.4217**	**0.3544**	**0.3901**	**0.6133**	**0.5648**	**0.5219**	**0.5667**
	MAPE											
BFGS-CM	4.68	**4.72**	4.33	4.58	5.77	6.13	5.69	5.86	9.14	8.83	8.24	8.74
PSO-CM	4.86	5.26	4.51	4.88	5.92	6.35	6.02	6.10	9.48	9.16	8.87	9.17
ACO-CM	5.10	5.39	4.87	5.12	6.58	7.18	6.24	6.67	9.72	9.33	9.04	9.36
Pro. CM	**4.39**	4.93	**4.12**	**4.48**	**5.39**	**5.61**	**5.24**	**5.41**	**8.62**	**8.35**	**7.96**	**8.31**
	MSE											
BFGS-CM	0.1533	**0.1562**	0.1497	0.1531	0.2971	0.3416	0.2691	0.3026	0.6839	0.5847	0.6012	0.6233
PSO-CM	0.1569	0.1613	0.1524	0.1569	0.3479	0.4017	0.2903	0.3466	0.7103	0.7482	0.6861	0.7149
ACO-CM	0.1822	0.1884	0.1763	0.1823	0.4107	0.4425	0.3876	0.4136	0.8482	0.7930	0.7425	0.7946
Pro. CM	**0.1475**	0.1582	**0.1463**	0.1507	**0.2071**	0.2483	**0.1984**	0.2179	**0.6420**	0.5939	**0.5781**	0.6047

5.4 The computation time of the developed combined model and comparisons

Table 22 shows the average computational time for each model at different sites. The responses of BPNN, ENN, GRNN, WNN, ARIMA, ES and the developed combined model are 5.56 s, 6.42 s, 7.13 s, 6.42 s, 312.46 s, 3.17 s and 17.75 s, respectively. Among the models, the developed combined model has the second longest computation time, and the ES model has the shortest time. The computation time of each model is within the acceptable range. Although the combined model has the longest computation time, it is superior to all of the compared models. Therefore, the developed combined model is both short and applicable for predicting the wind speed in wind power plants, it can be used in engineering application.

Table 22 The computation time of each model

Model	Computation time (s)	Model	Computation time (s)
BPNN	5.56	ARIMA	312.46
ENN	6.42	ES	3.17
GRNN	7.13	Combined model	17.75
WNN	6.42		

6. Conclusion

Accurate wind speed forecasting plays a vital role in the operation and management of wind power grids, which can bring considerable benefits and increase the utilization rate of wind energy. Thus, it is highly desirable to develop an optimized wind speed forecasting model with a high forecasting accuracy. However, individual forecasting models cannot always meet the increasing requirements. More specifically, individual forecasting models have some limitations and cannot perform forecasting effectively under certain conditions. The combined model can make full use of the merits of each model and can effectively forecast future wind speeds compared with the individual models. In this paper, a novel combined forecasting model based on a modified version of the AFSA and ACO, the AFSA-ACO, with two linear models, i.e., ES and ARIMA, and four nonlinear methods, i.e., BPNN, GRNN, WNN and ENN, was successfully proposed. More specifically, to enhance the forecasting capacity of the proposed combined model, the modified version of the AFSA and ACO, the AFSA-ACO, was integrated into this model and employed to optimize the weight of each individual model. There are two forecasting parts, the linear and nonlinear parts, for enhancing the forecasting performance, and they are used to separately capture the inherent linear and nonlinear features contained in wind speed. The case studies reveal that the developed combined model performs better than the other compared models, is an effective and promising technique for wind speed forecasting and can improve the forecasting effectiveness by combining these two linear and four nonlinear algorithms. Furthermore, based on a series of insightful discussions of our case studies, the effectiveness of the developed combined model is proven to be significantly superior to the compared models. For example, the DM test results reveals that the developed combined model has a higher degree accuracy than and is a significantly different from the other benchmark models. Despite the effectiveness of using the novel combined model illustrated above, some limitations certainly exist. For example, only wind speed series with 10-min intervals were employed to assess the performances of the developed model in this paper. Further research could make forecast using hourly, monthly and quarterly wind-speed data. Moreover, for future

research, the proposed combined model can be employed in other forecasting fields, such as electricity price, air quality, traffic flow and stock index forecasting.

Acknowledgements: This work was supported by the National Natural Science Foundation of China 41630421.

References

[1] L. Martin, Wind Energy – The Facts: A Guide to the Technology, Economics and Future of Wind Power, J. Clean. Prod. 18 (2010) 1122–1123. doi:10.1016/j.jclepro.2010.02.016.

[2] A. Kumar, K. Kumar, N. Kaushik, S. Sharma, S. Mishra, Renewable energy in India: Current status and future potentials, Renew. Sustain. Energy Rev. 14 (2010) 2434–2442. doi:10.1016/j.rser.2010.04.003.

[3] C. Monteiro, H. Keko, R. Bessa, V. Miranda, A quick guide to wind power forecating: state-of-the-art 2009., (2009).

[4] J. Yan, Y. Liu, S. Han, Y. Wang, S. Feng, Reviews on uncertainty analysis of wind power forecasting, Renew. Sustain. Energy Rev. 52 (2015) 1322–1330. doi:10.1016/j.rser.2015.07.197.

[5] S. Qin, F. Liu, J. Wang, Y. Song, Interval forecasts of a novelty hybrid model for wind speeds, Energy Reports. 1 (2015) 8–16. doi:10.1016/j.egyr.2014.11.003.

[6] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, C. Draxl, The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview, Tech. Report, ANEMOS.plus. (2011) 1–109. papers://6470de79-5287-45a9-8e4f-b629919aff7a/Paper/p5443.

[7] S. Al-Yahyai, Y. Charabi, A. Gastli, Review of the use of numerical weather prediction (NWP) models for wind energy assessment, Renew. Sustain. Energy Rev. 14 (2010) 3192–3198. doi:10.1016/j.rser.2010.07.001.

[8] A.K. Rajeevan, P.. Shouri, U. Nair, ARIMA Based Wind Speed Modeling for Wind Farm Reliability Analysis and Cost Estimation, J. Electr. Eng. Technol. 11 (2016) 869–877. doi:10.5370/JEET.2016.11.4.869.

[9] M.G. De Giorgi, A. Ficarella, M.G. Russo, Short-term wind forecasting using artificial neural networks (ANNs), in: Energy Sustain. Ii, 2009: pp. 197–208. doi:10.2495/esu090181.

[10] Z. Guo, J. Wu, H. Lu, J. Wang, A case study on a hybrid wind speed forecasting method using BP neural network, Knowledge-Based Syst. 24 (2011) 1048–1056. doi:10.1016/j.knosys.2011.04.019.

[11] C. Yu, Y. Li, M. Zhang, Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting, Energy Convers. Manag. 147 (2017) 75–85. doi:10.1016/j.enconman.2017.05.008.

[12] W. Zhang, Z. Qu, K. Zhang, W. Mao, Y. Ma, X. Fan, A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting, Energy Convers. Manag. 136 (2017) 439–451. doi:10.1016/j.enconman.2017.01.022.

[13] J. Wang, J. Heng, L. Xiao, C. Wang, Research and application of a combined model based on multi- objective optimization for multi-step ahead wind speed forecasting, Energy. 125 (2017) 591–613. doi:10.1016/j.energy.2017.02.150.

[14] L. Xiao, J. Wang, Y. Dong, J. Wu, Combined forecasting models for wind energy forecasting: A case study in China, Renew. Sustain. Energy Rev. 44 (2015) 271–288. doi:10.1016/j.rser.2014.12.012.

[15] Y. Yang, Y. Chen, Y. Wang, C. Li, L. Li, Modelling a combined method based on ANFIS and neural network improved by DE algorithm: A case study for short-term electricity demand forecasting, Appl. Soft Comput. (2016). doi:10.1016/j.asoc.2016.07.053.

[16] R. Rahmani, R. Yusof, M. Seyedmahmoudian, S. Mekhilef, Hybrid technique of ant colony and particle swarm optimization for short term wind energy forecasting, J. Wind Eng. Ind. Aerodyn. 123 (2013) 163–170. doi:10.1016/j.jweia.2013.10.004.
[17] T. Sen, H.D. Mathur, A new approach to solve Economic Dispatch problem using a Hybrid ACO–ABC–HS optimization algorithm, Int. J. Electr. Power Energy Syst. 78 (2016) 735–744. doi:10.1016/j.ijepes.2015.11.121.
[18] M.K. Patel, M.R. Kabat, C.R. Tripathy, A hybrid ACO/PSO based algorithm for QoS multicast routing problem, Ain Shams Eng. J. 5 (2014) 113–120. doi:10.1016/j.asej.2013.07.005.
[19] H. Hassani, Singular Spectrum Analysis: Methodology and Comparison, J. Data Sci. 5 (2007) 239–257. doi:10.3189/172756506781828863.
[20] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, Time Series Analysis: Forecasting & Control, 1994. doi:10.1016/j.ijforecast.2004.02.001.
[21] D. Asteriou, S.G. Hall, ARIMA Models and the Box–Jenkins Methodology, Appl. Econom. (2011) 265–286. doi:10.1002/(SICI)1099-131X(199705)16:3<147::AID-FOR652>3.0.CO;2-X.
[22] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature. 323 (1986) 533–536. doi:10.1038/323533a0.
[23] D.F. Specht, A General Regression Neural Network, IEEE Trans. Neural Networks. 2 (1991) 568–576. doi:10.1109/72.97934.
[24] J. Wang, W. Zhang, Y. Li, J. Wang, Z. Dang, Forecasting wind speed using empirical mode decomposition and Elman neural network, Appl. Soft Comput. 23 (2014) 452–459. doi:10.1016/j.asoc.2014.06.027.
[25] J.M. Bates, C.W.J. Granger, The Combination of Forecasts, J. Oper. Res. Soc. 20 (1969) 451–468. doi:10.1057/jors.1969.103.
[26] V. Maniezzo, A. Colorni, The ant system applied to the quadratic assignment problem, IEEE Trans. Knowl. Data Eng. 11 (1999) 769–778. doi:10.1109/69.806935.
[27] M. Dorigo, M. Birattari, T. Stutzle, Ant colony optimization, IEEE Comput. Intell. Mag. 1 (2006) 28–39. doi:10.1109/MCI.2006.329691.
[28] P. Jiang, F. Liu, Y. Song, A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting, Energy. 119 (2017) 694–709. doi:10.1016/j.energy.2016.11.034.
[29] P. Du, J. Wang, Z. Guo, W. Yang, Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting, Energy Convers. Manag. 150 (2017) 90–107. doi:10.1016/j.enconman.2017.07.065.
[30] W. Yang, J. Wang, R. Wang, Research and application of a novel hybrid model based on data selection and artificial intelligence algorithm for short term load forecasting, Entropy. 19 (2017). doi:10.3390/e19020052.
[31] F.X. Diebold, R.S. Mariano, Comparing Predictive Accuracy, J. Bus. Econ. Stat. 13 (1995) 253–263. doi:10.1080/07350015.1995.10524599.