Research and application of a combined model based on variable weight for short term wind speed forecasting

Hongmin Li[a], Jianzhou Wang[a*], Haiyan Lu[b], Zhenhai Guo[c]

[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*
[c] *State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 10029, China*

[*]*Corresponding author. Address: School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China*
Tel.:+86 13130476286
E-mail address: wjz@lzu.edu.cn

Abstract

Wind speed forecasting plays a prominent part in the operation of wind power plants and power systems. However, it is often difficult to obtain satisfactory prediction results because wind speed data comprise random nonlinear series. Current some statistical models are not proficient in predicting nonlinear time series, whereas artificial intelligence models often fall into local optima. For these reasons, a novel combined forecasting model, which combines hybrid models based on decomposed methods and optimization algorithms, is successfully developed with variable weighting combination theory for multi-step wind speed forecasting. In this model, three different hybrid models are proposed and to further improve the forecasting performance, a modified support vector regression is used to integrate all the results obtained by each hybrid model and obtain the final forecasting results. To verify the forecasting effectiveness of the proposed forecasting model, 10-min wind speed series from Penglai, China, are used as case studies. The experimental results indicate that the developed combined model not only outperforms other benchmark models but also can be satisfactorily used for planning for smart grids.

Key words: Combined model; Variable weight; Short-term wind speed forecasting; Forecasting accuracy

1. Introduction

Owing to the rapid pace of modern industrial development, the utilization rate of resources is increasing with each passing day, and the scarcity of resources has become an urgent problem. Wind power, a clean and renewable energy source, has been regarded as one of the alternatives to conventional fuel power generation. This led to a collaborative effort to achieve 20% of U.S. electricity supply from wind power by 2030 [1]. Generally, wind speed forecasting is at the core of wind power generation systems, and it plays an important role in their control and operational decision-making. The exact prediction of wind speed is of great use to enhance the utilization rate of wind energy and stabilize the power supply. On the contrary, inaccurate forecasting will result in bad decision-making, which may cause considerable economic losses in wind power generation systems.

Wind power forecasting (WPF) may produce decision risk to power system operation for its forecasting deviations [2]. Traditional methods of wind speed forecasting concentrate on the characteristics of historical data and the effect of numerical weather models on wind speed to perform the forecasting. Fortunately,

statistical models combined with artificial intelligence algorithm applied to the forecasting field have yielded good results. There are many methods for short-term wind speed forecasting, such as statistical models, artificial intelligence models, physical models, hybrid models, and combined models.

Common statistical models for predicting wind speed following the classical Box–Jenkins methodology include the auto-regressive moving average (ARMA (p, q)) and auto-regressive integrated moving average (ARIMA (p, d, q)) [3] models. The ARIMA model has a strong forecasting ability and it exhibits great precision in wind speed forecasting [4-5]. ARIMA models can help in understanding the dynamics of the data in a given application [6]. ARIMA can model linear patterns in time series well; however, it is not applicable to modeling nonlinear patterns [7]. Regression models are widely used in the field of time series prediction, which are suitable for certain time series with obvious trends. The disadvantage of traditional regression models is that they have fewer variable parameters and are difficult to adapt to the prediction of time series.

Artificial intelligence prediction models are mainly focused on artificial neural networks (ANNs), including the back propagation neural network (BPNN), Elman neural network (ENN), and radial basis function neural network (RBF). In recent years, the ANN approach has been widely utilized in the fields of economic parameter forecasting [8], biology forecasting [9], and wind speed forecasting [10], among others. Great achievements have been made in the field of wind speed prediction. ENN has proven useful for forecasting discrete time series, because of its potential capacity to model nonlinear dynamic systems owing to feedback connections and learning time-varying patterns [11-13]. BPNN has a long history in prediction, and it has made outstanding contributions to forecasting, especially in the state of uncovering nonlinearity between the inputs and outputs, even with a lack of sufficient information about the relationship between them [14]. BPNN is popular for forecasting complex nonlinear systems and it can actualize any complex nonlinear mapping function, which was mathematically proven [15]. However, BPNN easily falls into local minima, and often exhibits over-fitting [16].

Physical models need to describe information in detail according to the onsite conditions of the wind farm, and then the numerical weather prediction (NWP) system is applied to forecasting wind speed [7]. Physical models are easy to simulate at low cost. However, compared with statistical models, physical models have a high demand of data and are often utilized in long-term wind speed forecasting with diverse weather variables.

Hybrid models can overcome the shortcomings of single models, and have therefore become increasingly popular. Hybrid models overcome the drawbacks and integrate the advantages of single models by integrating two or more single models. In this way, the overall forecasting accuracy of hybrid models is improved. Wind speed time series data is highly nonlinear and unstable. Empirical mode decomposition (EMD) has been widely used in the analysis of nonlinear signals [17, 18]. Wang et al. pointed out that EMD has many advantages compared with wavelet transformation and Fourier transformation, such as good multi-resolution and extensive applicability [19]. The EMD can decompose an original time series into a small and finite number of oscillatory modes called intrinsic mode functions (IMF), and then the IMFs can be used in the model for prediction. Thus, a hybrid model that contains EMD can improve the accuracy enormously by eliminating the unstable nonlinear parts of the original data [20-21]. Wang et al. [11] suggested that a hybrid model with the ENN, such as EMD–ENN, could improve the prediction accuracy. Liu et al. [22] presented a novel hybrid

model, with fast ensemble empirical mode decomposition and wavelet packet decomposition, and ENN, which has desirable performance in the multi-step ahead wind speed forecasting. Jiang et al. [23] considered that people often ignored similar fluctuations between adjacent wind turbines in the process of the wind speed forecasting, and proposed a hybrid model combining $\upsilon-$support vector machine (v-SVM) and cuckoo search (CS). Liu et al. [24] combined three individual forecasting models (BPNN, RBFNN, and LSSVM) using the adaptive neuro-fuzzy inference system and obtained a great improvement in accuracy with regard to three individual models. Wang et al. [25] built a hybrid model with improved EMD and the GA-BP neural network for short-term wind speed forecasting.

In recent years, the research on the combination forecasting method has entered another peak. Many scholars have pointed out that the combination forecasting method has higher precision than the single model forecasting. Liang et al. [26] pointed out that a short-term wind power combined forecasting model based on error forecast correction can obtain better performance. By comparison with the existing traditional combined models, Xiao et al. [27] put forward a combined model that can always obtain satisfactory forecasting results. A weighting method or a meta-combination method can be used to combine the forecasts of individual models in the final step of the model. The most commonly used combination methods are the weighted median and the weighted average [28]. Zhang et al. [29] proposed that the best weights for the different models can be obtained by an optimization algorithm. The single forecasting model based on the classical combined method has a fixed weight, which cannot adapt to changes of the sample; thus, the adaptive weight can be adapted to different samples to obtain the matching weights. Support vector regression (SVR) is a method to perform a noise-robust and nonlinear regression based on the structural error minimization principle [30]. It can determine the regression model through the training set, and then obtain the prediction results from the test set. It has some parameters that have a profound impact on prediction accuracy, including a penalization term [30]. Chakri et al. [31] discussed the exploration capabilities of bat algorithm and improved it by introducing directional echolocation to standard bat algorithm.

As mentioned above, single models have many drawbacks with poor forecasting accuracy and stability, while hybrid models can overcome the defects of single models and attain higher precision. Therefore, based on the limitations or strengths discussed above, and considering the combined model as the mainstream in forecasting models, a novel combined forecasting model is developed. Generally, combination forecasting models can be divided into fixed-weight combination forecasting models and variable weight combination forecasting models [32]. In this paper, a novel combined forecasting model is proposed on the basis of the variable weight combination theory. More specifically, three different hybrid models are proposed and to further improve the forecasting performance, the modified SVR is used to integrate all the results obtained from each hybrid model and obtain the final forecasting results. Through the experiment, it is found that the variable weight combination forecasting model can yield more robust and higher-accuracy forecasting.

The contributions of the developed model are summarized as follows:

(1) Wind speed series usually have the characteristics of randomness, fluctuation and nonlinear, which often leads to difficulties in forecasting. Data preprocessing is used to improve accuracy by disintegrating the time series into the sum of many time series and extracting the main features of the original time series. Therefore, the removal of noise in wind speed series yields smoother and more predictable series.

(2) This paper proposed three hybrid models that are effective approaches to

noticeably increase the accuracy by decomposing the noisy components in the original series. Moreover, hybrid models combined with the bat algorithm greatly improved the efficiency of neural networks in forecasting. The hybrid models clearly outperform the traditional models in short-term wind speed forecasting.

(3) In hybrid models, two kinds of commonly used neural networks (BPNN, ENN) and a statistical model (ARIMA) are used to forecast the wind speed. In the forecasting process, the statistical model focuses on the linear problem, whereas the neural network excels at nonlinear series forecasting. In addition to wind speed series being irregular and nonlinear, sometimes they also have linear features. Therefore, the integration of three hybrid forecasting models can not only solve the nonlinear problems but also the linear problems.

(4)The novel combined forecasting model overcomes the disadvantages of traditional models that ignored the importance of variable weight. The SVR model is used to integrate all hybrid models' forecasting results as a final forecasting for the original wind speed.

The remainder of this article is organized as follows: Section 2 recalls the preliminaries and describes the method. Section 3 introduces the new proposed model in detail. The experiments and results are presented in Section 4 and Section 5, respectively. Section 6 discusses the results of the experiment and Section 7 presents the conclusions.

2. Methodology

In this section, we will recall several basic theories about the proposed model, and finally introduce the method of proposed combined model.

2.1 Empirical Mode Decomposition (EMD)

EMD is an effective algorithm, usually used to cope with nonlinear and non-stationary time series. The basic method of EMD is to decompose the original time series into a small and limited number of oscillatory modes by local characteristic timescale filtering. The oscillatory modes (IMFs) should satisfy the following principles [33, 34]: (1) In the whole dataset, the number of extremes and zeros must either be equal or differ by at most one; (2) at any point, the mean between the upper and lower envelopes must be zero, which are defined by the local maxima and minima. Let $s(t)(t=1,2,\ldots,l)$ be an original time series. The detailed steps of EMD are as follows:

Step 1: Determine all local maxima and minima of time series.

Step 2: Apply a cubic spline line to connect all local extrema to generate the upper envelop $e_{up}(t)$ and the lower envelop $e_{low}(t)$.

Step 3: Calculate the mean envelop from the upper and lower envelopes

$$m(t)=[e_{up}(t)+e_{low}(t)]/2 \tag{1}$$

Step 4: Calculate the difference between the original time series and the mean envelop then get the detailed components

$$h(t)=s(t)-m(t) \tag{2}$$

Step 5: Check whether $h(t)$ satisfies the IMF's characteristics. If so, $h(t)$ is treated as the ith IMF and $s(t)$ is replaced by the residuals $r(t)=s(t)-h(t)$. If not, $s(t)$ is replaced by $h(t)$.

Step 6: Repeat Steps 1–5, and then terminate the procedure when the size of the standard deviation between two successive sifting results is smaller than the pre-defined

threshold.

Through the above process, a set of IMFs can be picked out from the original time series in the order from high-frequency to low-frequency series. Thus, the original time series is decomposed to n IMFs and one residual as

$$s(t)=\sum_{i=1}^{n} c_i(t)+r_n(t) \tag{3}$$

where n is the number of IMFs, $r_n(t)$ are the final residuals representing a trend in $s(t)(t=1,2,\ldots,l)$, and $c_i(t)(t=1,2,\ldots,n)$ denotes the IMFs, which are periodic and nearly mutually orthogonal. When expressing local properties of an original signal, each IMF is independent and specific. The whole process can be seen in Fig.1 part a**.**

2.2 Bat Optimization Algorithm (BA)

The idea of the algorithm is to develop different bat heuristic algorithms using some of the ultrasonic features of tiny bats [35]. BA is more effective than many other optimization algorithms, because it controls the bat's space and range of movement by using frequency tuning and it parameters are adjustable [36]. The principles of BA are as follows:

(1) All the bats use the difference of ultrasonic echo feelings to determine the distance between food and obstacles.

(2) Bats random flight with velocity v_i at position x_i with fixed frequency $f_{\min}$ (or wavelength λ), and search for prey with different wavelength λ (or frequency f) and volume A_0. They automatically adjust their pulse wavelength (or frequency) according to the proximity of prey.

(3) Volume varies from a large positive value A_0 to a minimum $A_{\min}$.

BA is an iterative optimization technique that is initialized as a set of random solutions; then the optimal solution is searched by iteration and the local solution is generated by random flight around the optimal solution to enhance the local search. Thus, the definition of the new solution x_i^t at time t and the speed v_i^t of the update formula is:

$$f_i = f_{\min}+\left(f_{\max}-f_{\min}\right)\beta \tag{4}$$

$$v_i^t = v_i^{t-1}+\left(x_i^{t-1}-x_{gbest}^t\right) f_i \tag{5}$$

$$x_i^t = x_i^{t-1}+v_i^t \tag{6}$$

where β is a random variable that is uniformly distributed over a range of [0, 1], x_{gbest}^t represents the current global optimal position, and f_i is the frequency of the *ith* bat, which controls the range and speed of movement of the bats.

For local search, once a solution is selected as the current optimal solution, then each bat produces a local new solution according to the random walk rule:

$$x_{new}=x_{old}+\varepsilon A^t \tag{7}$$

where $\varepsilon\in[0,1]$ is a random number, and A^t is the average volume of all bats in the same time period.

This study adopts the BA to optimize the BPNN and ENN for improved prediction accuracy, defining the fitness function as

$$O = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i^{forecast} - y_i^{actual}}{y_i^{actual}}\right| \quad (8)$$

Here, y_i^{actual} represents the actual wind speed, $y_i^{forecast}$ represents the forecast wind speed, and n represents the bat population size. The specific process the BA is shown in Fig.1 part b and Fig.2 part a.

2.3 Back Propagation Neural Network (BPNN)

The BPNN is a kind of multilayer feedforward neural network that has the main characteristic of signal forward propagation and the error back propagation. In the existing studies, single hidden layer BPNNs are widely applied in one-step ahead forecasting. The appropriate number of input and hidden nodes can be determined by experiments or training based on the minimum mean square error of the test data [37]. Moreover, by using an optimization algorithm to determine the optimal weights and thresholds of networks, the forecasting capacity and accuracy can be significantly improved. In this study, the BPNN optimized by BA is applied to forecasting the wind speed.

2.4 Elman Neural Network (ENN)

ENN generally can be divided into four layers: the input layer, hidden layer, linked layer, and output layer. ENN has the characteristic that the output of the hidden layer leads to the input of the hidden layer through the delay and storage of the linked layer. This makes it sensitive to the history data by adding an internal feedback network, thereby enhancing the capacity of network dynamic information processing in order to achieve the purpose of dynamic modeling. The nonlinear state space expression of ENN is:

$$y(k) = g\left(w^3 x(k)\right) \quad (9)$$

$$x(k) = f\left(w^1 x_c(k) + w^2\left(u(k-1)\right)\right) \quad (10)$$

$$x_c(k) = x(k-1) \quad (11)$$

where y is an m-dimensional output node vector; x is an n-dimensional intermediate node unit vector; u is an r-dimensional input vector; x_c is an n-dimensional feedback state vector; w^1, w^2, and w^3 are the connection weights between the layers; $g(*)$ is a transfer function of the output neuron; and $f(*)$ is a transfer function of the intermediate layer neurons. Considering that ENN has a strong ability to deal with the nonlinear data, in this study, we use ENN to forecast the wind speed and obtain good performance.

2.5 ARIMA

ARIMA is a common statistical forecasting method, this study uses the ARIMA model to forecast wind speed. The model is known as ARIMA (p, d, q), where p is the order of the autoregressive part, q is the order of the moving average part, d is the degree of first differencing involved. ARIMA follows the Box–Jenkins methodology for the identification, estimation, diagnostic checking, and forecasting [38]. Because there is usually little historical data available, it is generally preferred to base the method on intra-sample fitting procedures. Therefore, Akaike's information criteria (AIC) was used to determine the lag order of ARIMA, which can penalize the possibility that the model compensates for data that may over-fit, which can also take into account the state of the fit and gives the order in the best fit state [39-40]:

$$AIC = -2\log(L) + 2k \quad (12)$$

where L denotes the likelihood of the model and k represents the total number of parameters and initial states that have been estimated.

The linear expression that defines the previous symbols is as follows [8]:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j} + \varepsilon_t \tag{13}$$

where ϕ_i is the ith autoregressive parameter, θ_j is the jth moving average parameter, and ε_t is the error term at time t.

2.6 Hybrid Models EMD-BA-BPNN, EMD-BA-ENN, and EMD-ARIMA

At present, there are two methods to forecast time series: artificial intelligence methods and statistical methods. In this study, two kinds of commonly used neural networks (BPNN and ENN) and a statistical model (ARIMA) are employed to forecast the wind speed. In the forecasting process, the statistical model focuses on the linear problem, whereas the neural network deals with nonlinear series forecasting. Because wind speed data has the characteristics of high degree of instability and nonlinearity, but occasionally exhibit certain linear features, the two kinds of models are not only selected to solve the nonlinear problem but also the linear series forecasting problem. Then, the EMD method is employed to eliminate noise from the wind speed series before the wind speed is predicted. For the ARIMA model, the parameters p and q have a great influence on the forecasting accuracy. Thus, in order to improve the effect of forecasting, the AIC criterion is applied to determine the order of the model. BPNN easily falls into local optima and exhibits excessive convergence, and its convergence speed is slow. On the contrary, ENN converges quickly, but with low accuracy. Therefore, to improve the forecasting accuracy, this study uses the BA to optimize the weights and thresholds of the two neural networks, and then to improve the forecasting accuracy. It can be seen from the results that the forecasting accuracy of the optimized model is significantly improved compared with that of the model without optimization.

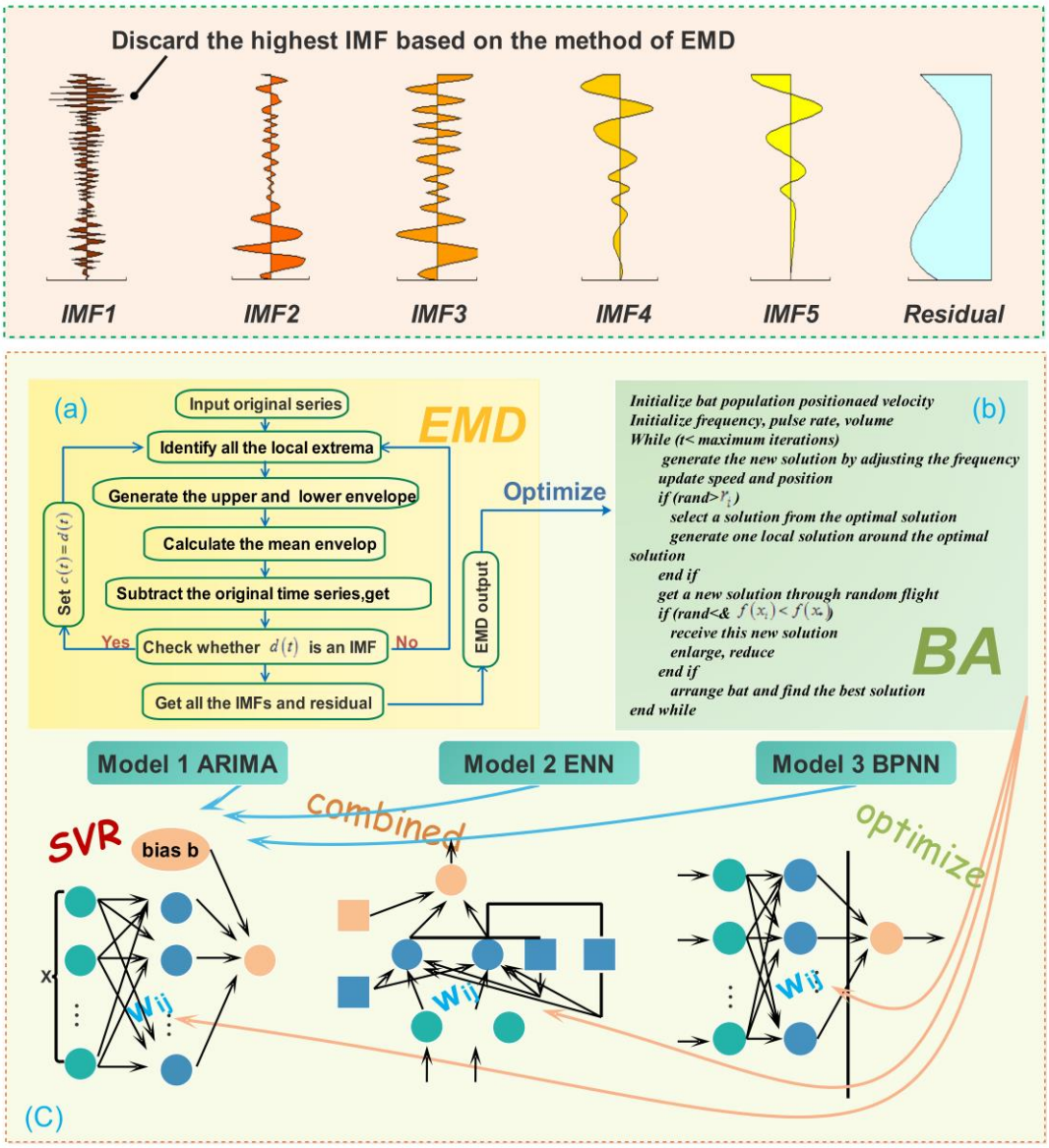

Fig.1. The structure of the paper

3. Combined Model

Combined models which integrate the results from several models are often utilized in forecasting field. In the combination method, the weight coefficient can obtained by simply averaging all available forecasts of a given variable and attributing equal weights to the individual forecasting. However, it is more appealing that minimize some cost function or objective criterion when determining optimal combination weights. [41]

3.1 Support Vector Machine Regression (SVR)

SVR, an extension of support vector machine (SVM), was proposed by Drucker et al. [42]. In recent years, SVR has shown outstanding forecasting performance [43-44]. The main idea of SVM is to establish a classification hyperplane as the decision surface, such that the separation edge between positive and negative cases is maximized. The theoretical basis of SVM is statistical theory; more precisely, SVM is the approximate realization of structural risk minimization.

It has the advantage of versatility, and can construct a function among a wide range of functions. It is theoretically perfect and computationally simple. Furthermore, it is

often regarded as one of the best ways to solve practical problems. The following steps show this method's specific form:

Step 1: Suppose that the training set is as follows:

$$T=\{(x_1,y_1),\ldots,(x_l,y_l)\}\in(X\times Y)^l \tag{14}$$

where $x_i\in X=R^n$, $y_i\in Y=\{1,-1\}(i=1,2,\ldots,l)$, and x_i is an eigenvector.

Step 2: Select the appropriate kernel function $K(x,x')$ and the appropriate parameters C; construct and solve following optimization problems:

$$\begin{aligned}&\min_{\alpha}\frac{1}{2}\sum_{i=1}^{j}\sum_{j=1}^{t}y_iy_j\alpha_i\alpha_jK(x_i,x_j)-\sum_{j=1}^{l}\alpha_j\\&s.t.\quad\sum_{i=1}^{l}y_i\alpha_i=0,\quad 0\le\alpha_i\le C,i=1,\ldots,l\end{aligned} \tag{15}$$

Then, obtain the optimal solution $\boldsymbol{\alpha}^*=(\alpha_1^*,\ldots,\alpha_l^*)^T$.

Step 3: Select a positive component of $\boldsymbol{\alpha}^*$, $0<\boldsymbol{\alpha}_j^*<C$, and calculate the threshold according to the following:

$$b^*=y_j-\sum_{i=1}^{l}y_i\alpha_i^*K(x_i-x_j) \tag{16}$$

Step 4: Structure decision function:

$$f(x)=\operatorname{sgn}\left(\sum_{i=1}^{l}\alpha_i^*y_iK(x,x_i)+b^*\right) \tag{17}$$

The general flowchart followed in this study as shown in Fig.1.

3.2 Traditional Combination Forecasting Method (TCM)

The traditional combination forecasting method (TCM) indicates that when there are n forecasting methods for dealing with a certain forecasting problem, several methods' forecasting results can be added up in order to simply and properly allocate weight coefficients. [45]

Assume that $f_{it}(i=1,2,\ldots,n)$ are the forecasting values of the n forecasting methods at time t, y_t is the actual value, and w_i is the weight of the ith forecasting method. Then the final forecasting value is:

$$\hat{y}_t=\sum_{i=1}^{n}\hat{w}_i f_{it},t=1,2,\ldots,m \tag{18}$$

$$\sum_{i=1}^{n}w_i=1 \tag{19}$$

3.3 Novel Combined Forecasting Model (NCFM)

In order to improve the quality of forecasting, three hybrid models, EMD-BA-BPNN, EMD-BA-ENN and EMD-ARIMA, were selected as the base predictors in this study. In the forecasting process, the EMD-ARIMA model focuses on the linear problem, whereas EMD-BA-BPNN and EMD-BA-ENN deal with the nonlinear series forecasting. Because wind speed data exhibits a high degree of instability and nonlinearity, but sometimes shows certain linear features, the three selected models not only solve the nonlinear problem, but also the linear series forecasting problem. However, optimally choosing the combination weights is the key to combining forecasting models. SVR is an application in the field of regression forecasting. In this

study, SVR is employed to integrate all the forecasted components into an ensemble of results for the final forecasting. The regression coefficient is considered as the weight of the combination. Because the regression coefficient changes with the data, it can be considered that the combination weight is variable. Therefore, a novel combined forecasting model (NCFM) that aggregates the results from three hybrid models based on the variable weight theory was proposed.

Because SVR can find the perfect fitting variables from the training set, it can be used to realize the adaptive change of the weight of every forecasting method. The parameters of SVR have a significant impact on the results and hard to determine. Accordingly, the parameters of SVR were optimized by the BA, and the accuracy is markedly improved. From the experimental results, the proposed combined model (NCFM) performs better than the single model; furthermore, it evidently surpasses the TCM. Fig.2 shows the flowchart for the weighting-based combined methods.

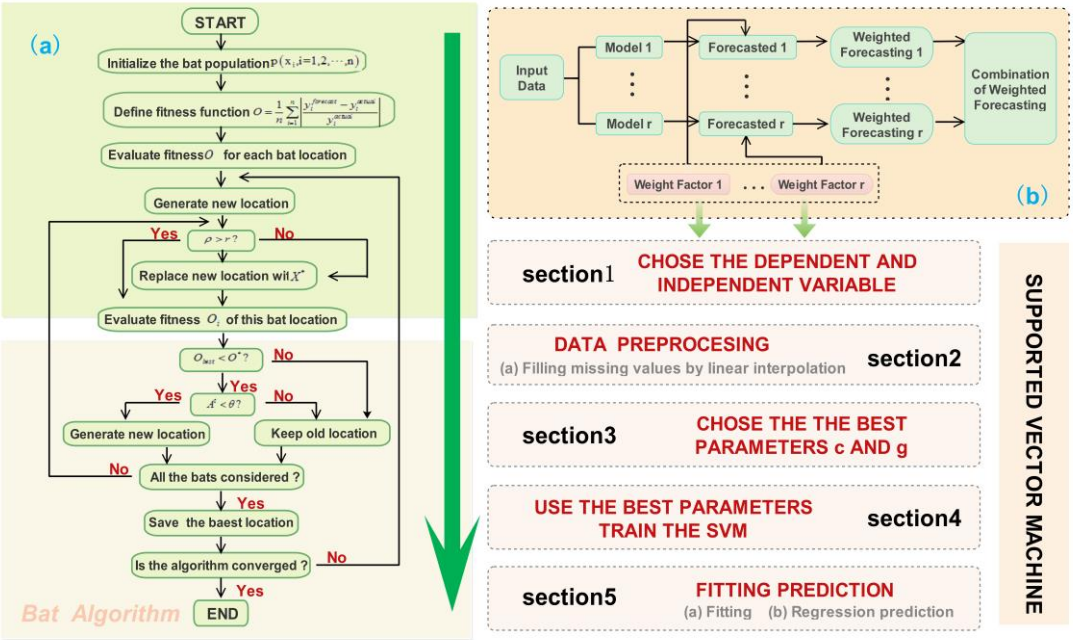

Fig.2. The flowchart for the weighting based combined approaches

4. Numerical Experimentation

Penglai is a seaside city of Shandong province in the north of Jiaodong Peninsula in China. Although the area is not large, because of its unique geographical advantages, it has a wealth of wind resources. Therefore, this study included an experiment based on the wind speed series of Penglai. Accordingly, the datasets A, B, and C are chosen from three adjacent sites of a wind farm in Penglai, and then three experiments were developed. For each experiment, the data is divided into four seasons, to facilitate the analysis of differences in the results of different seasons. This study also examined one-step forecasting and multistep forecasting, and each experiment included one-step ahead forecasting and multistep ahead forecasting.

The experiments performed in our study were implemented on Matlab2016a running on the Windows 8.1 Professional operating system. The specific hardware parameters of hardware were: Intel (R) Core i5-4590 3.30 GHz CPU and 8 GB RAM.

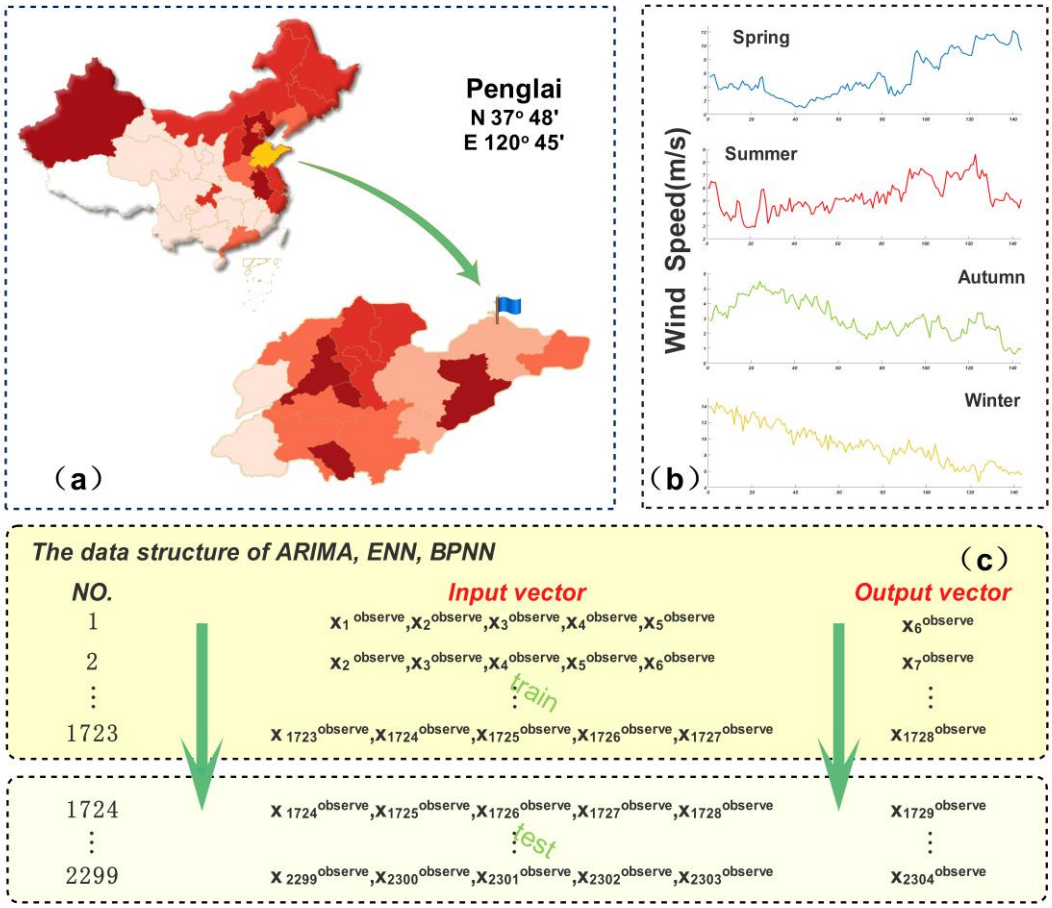

Fig.3. The location of Penglai and data sructure about models. (a) the location of Penglai in Shandong province.(b) the original wind speed from four quarters.(c) the data structure of experiment

4.1 Description of Data Sets

Considering the different wind speed fluctuation of different units, the wind speed is greatly affected by the season. The 10-min wind speed data sample is selected from three observation sites (dataset A, dataset B, dataset C) of Penglai from the four seasons. The observations of each dataset are split into a training set and testing set. The size ratio of the training set to testing set was fixed as 3:1. The input–output structure of the forecasting involved using the five previous days of observations to forecast the next one day wind speed and to replace the latest day. Fig.3 shows the forecasting structure of the datasets of each unit. The detailed arrangements of the experimental dataset are presented in Table 1.

4.2 Evaluation Metrics

Evaluation metrics are used to directly reveal the forecasting accuracy of the combined forecasting model with variable weights. These include the mean absolute percent error (MAPE), mean absolute error (MAE), and mean square error (MSE), of which smaller values indicate higher forecasting accuracy [46]. The metrics are defined as follows:

$$MAPE = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{y_n - \tilde{y}_n}{y_n}\right| \times 100\% \tag{20}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_n - \tilde{y}_n\right| \tag{21}$$

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left(y_n - \tilde{y}_n\right)^2 \tag{22}$$

where y_n is the actual value at time t, and $\tilde{y}_n$ is the predicted value at the corresponding time.

Table 1

The arrangement of the experimental data set in spring

Data Set	Numbers	Statistical Indicator			
		Mean(m/s)	Max(m/s)	Min(m/s)	Std(m/s)
Data Set A					
All samples	2304	7.0809	15.5	0.8	3.0282
Training	1728	7.3711	15.2	1.2	3.0261
Testing	576	6.2102	15.5	0.8	2.8659
Data Set B					
All samples	2304	7.6586	18.7	0.7	3.4504
Training	1728	8.1397	18.7	1.0	3.4649
Testing	576	6.2155	14.9	0.7	2.9737
Data Set C					
All samples	2304	7.0079	18.5	0.6	3.3185
Training	1728	7.4996	18.5	1.1	3.3933
Testing	576	5.5328	14.8	0.6	2.5725

MAPE can be used to effectively measure the prediction performance, and its specific criterion [46] is shown in Table 2. These three metrics measure the average effect of the forecast, and they are very sensitive to the changes of the dataset. When there has a small change in dataset, the metrics change significantly. Similar to reducing the number of training sets, adding a forecasting step, will increase the metrics.

Table 2

Criterion of MAPE

MAPE（%）	**forecasting power**
<10	Excellent
10-20	Good
20-50	Reasonable
>50	Incorrect

4.3 Experiment and Forecasting Model Parameter Selection

The statistical model ARIMA and artificial intelligence neural networks ENN and BPNN play a significant role in this study, as do the weights and thresholds of the network optimized by the BA. Through many experiments, the parameters of the model are determined as follows:

(1) For the BA, the bat population size usually varies from 10 to 25. From the experimental datasets, when the bat population size defined as 10, the optimization result is more stable and the precision is higher. The other parameters of the BA are presented in Table 3.

(2) For ARIMA, the order of autoregressive as well as moving average have enormous implications for the establishment of the model and the forecasting results. The AIC criterion measures the fitting effect of the observed value and takes into account the number of parameters in the best fit state. In terms of the AIC, the best order

of the ARIMA model for forecasting is that with the smallest value of the AIC. Consequently, the lag order p and q determined by the AIC is presented in Table 4.

Table 3
Model Parameters

Model	Experimental Parameters	Default Value
BA	BA population size	10
	Loudness(dB)	0.25
	Pulse rate (%)	50%
	Frequency maximum(kHz)	2
	Frequency minimum(kHz)	0
Elman	the number of input nodes	5
	the number of output nodes	1
	the number of hidden nodes	9--15
	The maximum number of trainings	1000
	Training requirements precision	0.00001
BPNN	the number of input nodes	5
	the number of output nodes	1
	the number of hidden nodes	9--15
	the learning velocity	0.1
	The maximum number of trainings	100
	Training requirements precision	0.00001

(3) For ENN, build the network using the newelm function. The dimensions of the input, hidden, and output layers are presented in Table 3.

(4) For BPNN, build the network using the newff function. The dimensions of the input, hidden, and the output layers are presented in Table 3.

Table 4
The order of ARIMA

Quarters		ARIMA		EMD-ARIMA	
		p	q	p	q
First Quarter	Site1	9	20	12	20
	Site2	6	10	11	16
	Site3	8	15	13	14
Second Quarter	Site1	9	3	14	12
	Site2	4	3	14	20
	Site3	12	12	15	20
Third Quarter	Site1	20	20	18	9
	Site2	19	12	7	6
	Site3	12	10	14	7
Fourth Quarter	Site1	12	12	12	11
	Site2	9	12	9	11
	Site3	13	14	19	17

4.4 Experimental Results for Datasets

For the simulation, the proposed new model is trained based on the wind speed values from the three datasets. The wind speeds of one day were forecasted by single-step and multistep ahead forecasting. The method of multistep ahead forecasting involves updating the input data by discarding the old data for each loop to perform the prediction. The multistep ahead forecasting forecasts the next wind speed value in the form of iterations by using the previous forecasting values rather than the actual value [47]. Through the forecasting results, the effectiveness of the proposed combined model

was verified. The multistep ahead forecasting is described as follows: define the time index h as the forecast origin and the positive integer l as the forecast horizon. Suppose we are at time index h and intended to forecast $\hat{r}_{h+l}$, where $l \geq 1$. Let $\hat{r}_h(l)$ be the forecast of r_{h+l}, then we defined $\hat{r}_h(l)$ as the l-step ahead forecast of r_t at the forecast origin h. When $l = 1$, we defined $\hat{r}_h(1)$ as the one-step ahead forecast of r_t at the forecast origin h [48-49]. Tables 5-6 show the accuracy of different models for different datasets. The results of one-step ahead, two-step ahead, and three-step ahead are shown in Fig.4.

Table5

Comparison of errors of different models for single-step vs multi-step in first two quarters

Dataset	Model	MAPE(%)			MAE(m/s)			MSE(m/s)^2		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Spring										
A	EMD-ARIMA	6.6131	8.1257	9.0714	0.2970	0.3642	0.4083	0.1552	0.2427	0.3145
	EMD-BA-ENN	12.0970	15.5208	18.7744	0.4081	0.5307	0.6841	0.2589	0.4904	0.9369
	EMD-BA-BPNN	6.4298	9.5240	12.4630	0.2748	0.3954	0.5258	0.1236	0.2875	0.6159
	NCFM	5.7459	7.0996	8.2900	0.2565	0.3268	0.3800	0.1088	0.2033	0.2606
B	EMD-ARIMA	6.4762	8.1863	9.2329	0.2944	0.3495	0.3945	0.1651	0.2442	0.3082
	EMD-BA-ENN	13.1992	17.6568	20.6656	0.3956	0.5533	0.6437	0.3278	0.6834	0.9322
	EMD-BA-BPNN	5.9013	9.0931	11.1932	0.2631	0.3967	0.4523	0.1293	0.3468	0.4823
	NCFM	4.6075	7.0229	7.8446	0.2158	0.3332	0.3599	0.0828	0.2353	0.2624
C	EMD-ARIMA	6.6908	7.8840	9.0184	0.2946	0.3332	0.3774	0.1565	0.2228	0.2879
	EMD-BA-ENN	14.7032	20.4802	24.9502	0.3859	0.5338	0.6939	0.2577	0.5572	0.8616
	EMD-BA-BPNN	5.2003	8.5966	11.5258	0.2305	0.3454	0.4572	0.1034	0.2366	0.4013
	NCFM	4.8399	7.3619	8.7817	0.2209	0.3262	0.3800	0.0897	0.2113	0.2877
Summer										
A	EMD-ARIMA	5.5492	5.9353	6.7353	0.2752	0.2930	0.3355	0.1314	0.1367	0.1869
	EMD-BA-ENN	6.9021	8.1961	9.9049	0.3488	0.4083	0.5106	0.2057	0.3210	0.4869
	EMD-BA-BPNN	5.0740	7.1604	8.2231	0.2568	0.3623	0.4232	0.0954	0.2514	0.2972
	NCFM	4.5723	5.9222	6.3354	0.2289	0.2914	0.3208	0.0769	0.1366	0.1714
B	EMD-ARIMA	8.2129	8.7162	9.5838	0.3659	0.3925	0.4389	0.2173	0.2475	0.3195
	EMD-BA-ENN	10.4676	12.6239	14.4143	0.4863	0.5794	0.6865	0.3797	0.5472	0.7844
	EMD-BA-BPNN	7.6444	10.1428	12.3186	0.3428	0.4597	0.5474	0.1768	0.3470	0.6246
	NCFM	6.5418	8.4118	9.4938	0.2884	0.3835	0.4344	0.1331	0.2386	0.3237
C	EMD-ARIMA	7.0995	7.9436	8.7701	0.3165	0.3589	0.3944	0.1549	0.2316	0.2826
	EMD-BA-ENN	8.1754	10.8831	12.5890	0.3790	0.5103	0.5835	0.2390	0.4983	0.6286
	EMD-BA-BPNN	7.0035	9.1102	11.4498	0.3251	0.4189	0.5169	0.1700	0.3562	0.4829
	NCFM	5.6535	7.8147	8.6145	0.2550	0.3561	0.4012	0.0991	0.2348	0.2841

Table 6

Comparison of errors of different models for single-step vs multi-step in last two quarters

Dataset	Model	MAPE(%)			MAE(m/s)			MSE(m/s)^2		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Autumn										
A	EMD-ARIMA	8.1799	7.8101	9.0549	0.2081	0.1962	0.2258	0.0626	0.0647	0.0826
	EMD-BA-ENN	16.7774	21.2795	24.8870	0.3115	0.3998	0.4679	0.1632	0.2615	0.3652
	EMD-BA-BPNN	7.2180	10.5148	13.9602	0.1797	0.2530	0.3175	0.0485	0.1105	0.1771
	NCFM	5.9417	7.3445	8.6143	0.1537	0.1902	0.2206	0.0368	0.0632	0.0801
B	EMD-ARIMA	7.0930	8.2285	8.6828	0.1770	0.2037	0.2160	0.0476	0.0668	0.0757
	EMD-BA-ENN	9.9132	11.8559	13.9097	0.2421	0.2876	0.3175	0.0941	0.1379	0.1644
	EMD-BA-BPNN	6.2094	8.3945	10.3575	0.1528	0.2213	0.2713	0.0362	0.0862	0.1310
	NCFM	6.0736	8.1827	8.6534	0.1506	0.2097	0.2244	0.0333	0.0702	0.0800
C	EMD-ARIMA	6.2787	6.7913	7.8072	0.1731	0.1877	0.2148	0.0500	0.0616	0.0788
	EMD-BA-ENN	12.4135	16.2894	19.6921	0.2861	0.3763	0.4549	0.1255	0.2232	0.3338
	EMD-BA-BPNN	5.3725	7.3716	10.3284	0.1460	0.2018	0.2787	0.0319	0.0662	0.1428
	NCFM	4.7648	6.7536	7.6200	0.1325	0.1862	0.2127	0.0268	0.0586	0.0789
Winter										
A	EMD-ARIMA	4.3643	4.4716	4.5814	0.3958	0.3951	0.4125	0.2625	0.2444	0.2720
	EMD-BA-ENN	6.6654	8.3821	8.9026	0.6750	0.8213	0.8625	0.8791	1.2238	1.3528
	EMD-BA-BPNN	4.5059	5.9472	7.4693	0.4086	0.5466	0.7086	0.2847	0.5359	0.8919
	NCFM	3.5439	4.3845	4.4945	0.3216	0.4040	0.4511	0.1489	0.2566	0.3307
B	EMD-ARIMA	4.9444	4.3035	4.6488	0.4375	0.3854	0.4125	0.2819	0.2173	0.2506
	EMD-BA-ENN	7.2625	10.2637	14.8425	0.7707	1.1604	1.6526	1.4030	4.6385	9.8762
	EMD-BA-BPNN	4.0534	6.1036	7.5744	0.3800	0.5882	0.7484	0.2431	0.6857	1.1326
	NCFM	3.4926	4.2143	4.6145	0.3164	0.4041	0.4584	0.1436	0.2532	0.3310
C	EMD-ARIMA	4.5921	4.0730	4.3032	0.3920	0.3419	0.3608	0.2286	0.1926	0.2097
	EMD-BA-ENN	6.5974	8.3258	9.4818	0.6518	0.8269	0.9137	0.9968	1.6353	1.8391
	EMD-BA-BPNN	4.5442	6.2406	8.2066	0.3972	0.5620	0.7708	0.2689	0.6330	1.4048
	NCFM	3.9828	4.0147	4.2968	0.3277	0.3648	0.3938	0.1688	0.2125	0.2492

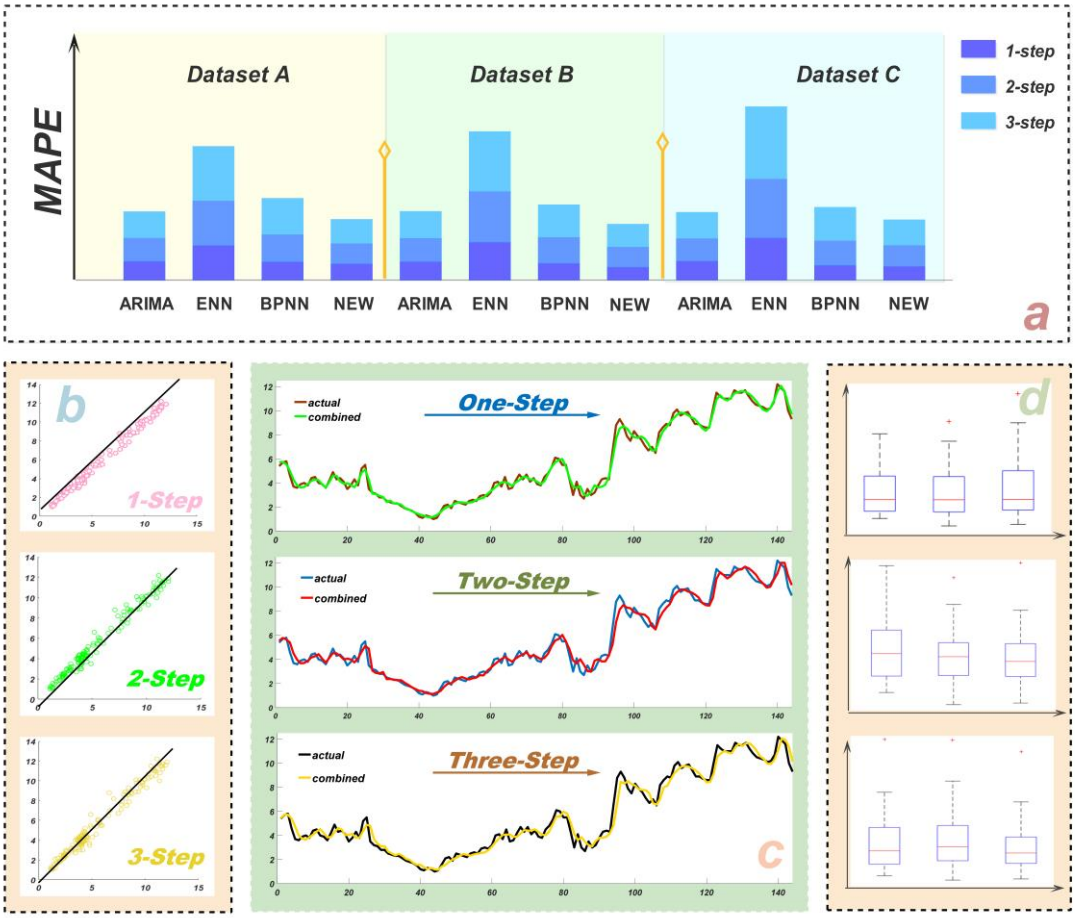

Fig.4. The results for proposed model for one-step, two-step and three-step

(1) First, we compare the hybrid and combined models in one-step forecasting. The hybrid models achieve the MAPE value of 5.2003%, 5.0740%, 5.3725% and 4.5059%, in four quarters respectively. The combined models perform better, with MAPE values of 4.6075%, 4.5723%, 4.7648%, and 3.4926%, respectively. Therefore, the forecasting results of the combined model are better than the hybrid model. From Tables 5-6, the accuracy of the combined model is far better than the hybrid model.

(2) Second, compared with multistep forecasting, the one-step forecasting performs better, and the MAPE values of the proposed model are 4.6075%, 7.0229%, and 7.8446% in one-step, two-step, and three-step forecasting, respectively, in spring. Multistep forecasting is based on single-step ahead forecasting, whereby the number of outputs is changed, and each forecasting input is added to the previous forecast. Therefore, with the increase in the number of forecasting steps, the forecasting accuracy gradually decreases.

(3) Finally, compared with conventional models, the proposed model performs better. According to Fig.4, the new model has a high fitting degree with the true value. More importantly, the new model has higher forecasting accuracy.

Remark: The combined forecasting model considers more information, and thus, it performs better. However, the multistep forecasting procedure is complex and it introduces high error. In the process of multistep forecasting, there is less historical information and the predicted values of each step will enter into the forecasting process

as input. As a result, each stack of the process will produce errors, causing an accumulation of errors decreased accuracy, and increased uncertainty.

5. Analysis of the Experimental Results

In this part, we will analyze the experimental results in detail, and demonstrate the validity of the new developed model, which is divided into two parts. First, from three angles to analyze one-step forecasting, one-step forecasting has the advantages of simple procedure, high forecasting accuracy shown in Tables 5-6 and Fig.4. Next, on the basis of one-step forecasting, the multi-step forecasting is analyzed.

5.1 Analysis for One-Step Forecasting

In this section, the experiments arranged to evaluate the forecasting results will be divided into three parts to demonstrate the validity of the proposed model. The first part is to verify the effectiveness of hybrid models compared with single models. The second part is conducted to analyze the wind speed forecasting of different datasets. The third part is compare the influence of seasonal factors on wind speed forecasting.

5.1.1 Analysis Single Model vs Hybrid Model

Take dataset A from the first quarter as an example. The accuracy data of single models and hybrid models are presented in Table 7. The actual wind speed fluctuated violently can be seen in Fig.5. The reconstructed wind speed series used EMD by removing the noise is shown in Fig.5. By comparing the actual series with the reconstructed series, the absolute residuals are calculated in Fig.5. We draw the following conclusions from the analysis.

(1) From Fig.5, the extreme instability of wind speed series is very obvious and the reconstruction of wind speed series is helpful to overcome it, by comparing the forecasting results of denoising series with the actual series. From Table 7, the accuracy is all greatly improved when using the denoising wind speed series, the EMD-ARIMA decreases the MAPE value by 4.0471%, 3.1815%, and 4.9287%, respectively. Similarly, for the EMD-ENN model, the accuracy has the improvement of 2.3857%, 4.6630%, and 4.0122%, respectively. The EMD-BPNN model shows the biggest improvement of 8.2018%, 4.6389%, and 6.9979%. The MAPE of the novel proposed model NCFM decreases by 4.2003%, 4.8081%, and 5.1002%. By comparing the results, it can be determined that EMD has good validity.

(2) From Fig.6, the hybrid models with the BA perform better than the single model in the wind speed forecasting. This study used the BA to optimize the weights and thresholds of the network, and the results reveal that the BA is effective. From Table 7, the BA-ENN shows MAPE values decreased by 1.3360%, 1.6301%, and 1.1425%, respectively. Moreover, BA-BPNN shows the improvement in accuracy of 3.57%, 0.4681%, and 1.1379%. The value of the evaluation metrics all decreased significantly. It can be concluded that the BA contributes much to the forecasting process and hybrid models are effective.

Remark : Consequently, it is necessary to remove the noise from the wind speed series before forecasting, and the EMD is one of the effective methods for this purpose. In conclusion, using the EMD to denoise the wind speed in the process of forecasting the wind series is valuable. The BA also positively impacts the forecasting process. Properly changing the weights and thresholds can greatly improve the network

forecasting ability. As a result, the hybrid model with the proper compound mode can exceed the single model in forecasting.

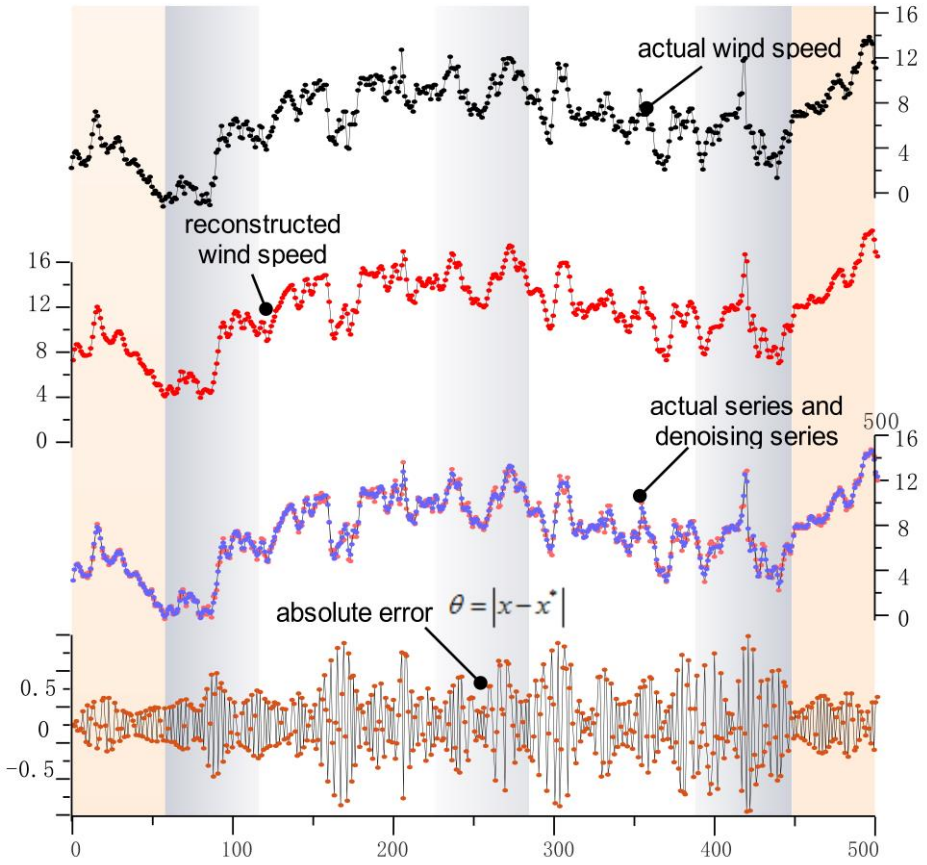

Fig.5. The denoising results of datasetA from the first quarter

Table 7

The accuracy of the model with denoising and without denoising

Model	Dataset A			Dataset B			Dataset C		
	MAPE (%)	MAE (m/s)	MSE $(m/s)^2$	MAPE (%)	MAE (m/s)	MSE $(m/s)^2$	MAPE (%)	MAE (m/s)	MSE $(m/s)^2$
ARIMA	10.6615	0.4818	0.4426	9.6649	0.4395	0.3816	11.6221	0.5117	0.4928
EMD-ARIMA	6.6144	0.2970	0.1552	6.4834	0.2944	0.1651	6.6934	0.2946	0.1565
ENN	14.8504	0.5205	0.4858	18.1247	0.5547	0.5210	19.0835	0.5917	0.6102
BA-ENN	13.5144	0.4953	0.4248	16.4946	0.5178	0.4914	17.9410	0.5375	0.5123
EMD-ENN	12.4647	0.4033	0.2559	13.4617	0.3819	0.2851	15.0713	0.3985	0.2819
BP	14.8937	0.5367	0.4825	11.1124	0.4468	0.3804	12.8814	0.4889	0.4366
BA-BP	11.3237	0.4679	0.4113	10.6443	0.4536	0.3893	11.7435	0.4626	0.3991
EMD-BP	6.6919	0.2730	0.1660	6.4735	0.2794	0.1440	5.8835	0.2373	0.1083
CFM	9.9531	0.4598	0.4057	9.4216	0.4354	0.3865	9.9438	0.4439	0.3932
NCFM	5.7528	0.2565	0.1088	4.6135	0.2158	0.0828	4.8436	0.2209	0.0897

5.1.2 Analysis of the Results of Different Datasets

In the second part, the wind speed series of the three nonadjacent wind turbine units are selected to forecast the wind speed. Take the data from the first quarter as an

example. The results from all the different models are clearly visible in Fig.4 part a, and Tables 5-6 show the accuracy of different turbine units from four quarters. Then, we can draw the following conclusions:

(1) From the view of the 10-min point wind speed series forecasted for a whole day, we can see that the wind speed series from the three wind turbine units have the same fluctuation trend in general, but the gap is obvious at a single time point. The reason for this is that the wind speed of the same place is roughly the same, but specific to different wind turbine units, the size and direction of the wind speed changes, creating the wind speed difference between different wind turbine units.

(2) From Table 5, the MAPE values of the four quarters from the three turbine units are different. For example, the MAPE of the NCFM from the three datasets are 5.7459%, 4.6075%, and 4.8399%. From the forecasting accuracy, we can see that dataset A shows the best accuracy.

Remark : The direction and magnitude of wind speed are uncertain, and the wind speed values of different adjacent wind turbine units are also different. Moreover, the wind speed fluctuates distinctly in different periods of a day. Therefore, the forecasted accuracy varies diversely from sites and the periods.

5.1.3 Analysis of Season Features

This part compares the forecasted results and the forecasting accuracy of different wind turbines in different seasons. For example, Tables 5-6 present the accuracy of forecasting models. The following conclusions can be drawn:

(1) The size and the fluctuation trend of wind speed are closely related to the seasons. From the original series of four quarters shown in Fig.3, the fluctuations in the second quarter are some of the more intense; the gap between the actual value and the forecast value for each model are slightly larger. The fourth quarter trend is obvious; the wind speed fluctuations in each period is more stable, so the predicted value is close to the actual value, and the forecast effect is better.

(2) From Tables 5-6, it can be seen that the accuracy of different models from the four quarters is diverse. In the four quarters, the NCFM performs best all the time. It can still be concluded that the hybrid model is more accurate than the single model. The MAPE values of the NCFM from dataset A are 5.7459%, 4.5723%, 5.9417%, and 3.5439%. MAPE of the fourth quarter is the smallest, because the wind speed of the fourth quarter shows smaller fluctuations.

Remark: The accuracy of wind speed forecasting is greatly influenced by the fluctuation of wind speed. However, the forecasted results indicate that the novel developed method is more accurate.

5.2 Analysis of Multi-Step Forecasting

Multi-step forecasting plays a very important part in many forecasting experiments. Accordingly, the proposed model that integrates three hybrid models is applied to multistep forecasting in this study. Then the experimental results can be used to testify the effectiveness of the combined forecast model. The forecasting results of two-step and three-step from different models are shown in Tables 5-6 and Fig.4. The results indicate that the developed model also performs better than other benchmark models. Further comparison results are illustrated below.

(1) Similar to the analysis of one-step forecasting, the accuracy of the hybrid models is higher than the single model in multi-step forecasting. Similarly, the combined model is still the most accurate for multi-step forecasting, but its accuracy is less than that in the one-step forecasting.

(2) From a comparison of the results of one-step forecasting and multistep forecasting, the one-step forecasting has obvious advantages in accuracy. Take the results of spring as an example. The MAPE of NCFM is 5.7459% in one-step forecasting of dataset A, and in two-step and three-step forecasting, it is 7.0996% and 8.29%, respectively. The MAE values are 0.25, 0.32, and 0.38, respectively. The MSE values are 0.10, 0.20, and 0.26 respectively.

These findings can be summarized as follows:

(1) When comparing the results of different seasons, whether in single-step or multistep ahead forecasting, the proposed model performs best.

(2) When comparing results of different datasets, we can draw the same conclusion that the developed combined model performs better than any other models, whether in single-step or multistep ahead forecasting.

Remark: From the above analysis, we can conclude that the developed model always performs better than any other models by comparing the three benchmark indexes MAPE, MAE, and MSE, whether in the single-step or multistep forecasting. Therefore, the newly developed model is effective.

Table 8

The real value and forecasting results of models on March 26th from 0:00 to 21:00

Time	Real (m/s)	ARIMA		ENN		BPNN		Traditional Combined		NCFM	
		Forecast(m/s)	FE(%)	Forecast(m/s)	FE(%)	Forecast(m/s)	FE(%)	Forecast(m/s)	FE(%)	Forecast(m/s)	FE(%)
Dataset A											
0:00	5.4000	6.1808	14.4588	5.8590	8.4996	6.1312	13.5410	5.7372	6.2444	5.7232	**5.9849**
3:00	4.1000	4.2702	4.1505	4.4629	8.8512	4.3698	6.5815	3.9637	**3.3238**	3.9348	4.0295
6:00	1.8000	2.1305	18.3585	2.3794	32.1902	2.0505	13.9141	1.8745	4.1393	1.8151	**0.8381**
9:00	2.6000	2.4661	5.1516	2.7478	5.6831	2.6913	**3.5098**	2.4635	5.2499	2.4116	7.2469
12:00	3.9000	4.3664	11.9584	4.4865	15.0397	4.4718	14.6626	4.1137	5.4790	4.0875	**4.8074**
15:00	4.3000	3.7270	13.3257	3.6840	14.3248	3.9889	**7.2347**	3.7174	13.5498	3.6849	14.3040
18:00	8.9000	8.3094	6.6360	8.5736	3.6676	8.4466	5.0948	8.6044	3.3212	8.6091	**3.2684**
21:00	11.0000	10.7944	1.8692	10.6733	2.9701	10.9739	**0.2371**	10.9493	0.4608	11.0344	0.3132
Dataset B											
0:00	5.3000	5.2008	1.8720	5.1094	3.5966	5.2229	1.4552	5.3166	**0.3134**	5.4295	2.4427
3:00	4.0000	3.5610	10.9743	3.6466	8.8354	3.7355	6.6123	3.8349	4.1265	3.9059	**2.3526**
6:00	1.9000	2.1910	15.3148	2.7017	42.1961	2.2812	20.0649	2.2014	15.8631	1.9306	**1.6118**
9:00	3.1000	2.7705	10.6295	3.0164	**2.6958**	2.9921	3.4799	2.9254	5.6315	2.8374	8.4704
12:00	3.5000	3.2953	5.8480	3.5123	**0.3524**	3.0773	12.0770	3.5331	0.9468	3.4342	1.8808
15:00	2.8000	3.1396	12.1290	3.2490	16.0350	3.3691	20.3250	3.0649	9.4605	2.9906	**6.8072**
18:00	8.9000	8.8202	0.8964	8.9873	0.9814	8.9835	0.9384	8.7277	1.9365	8.9422	**0.4744**
21:00	11.0000	11.4935	4.4868	11.5933	5.3937	11.4685	4.2590	10.7984	1.8331	10.9745	**0.2319**
Dataset C											
0:00	4.1000	4.1459	1.1188	3.9952	2.5569	4.1477	1.1629	4.0724	**0.6722**	4.2128	2.7512
3:00	3.3000	3.4371	4.1534	3.3929	2.8145	3.4678	5.0851	3.3815	**2.4705**	3.3950	2.8785
6:00	2.3000	2.3014	**0.0606**	2.7398	19.1205	2.4223	5.3189	2.4960	8.5202	2.3166	0.7206
9:00	2.7000	2.5850	4.2593	2.9369	8.7747	2.6075	3.4253	2.7121	**0.4464**	2.4980	7.4810
12:00	3.6000	3.8351	6.5319	3.8259	6.2751	3.7322	**3.6716**	3.8236	6.2122	3.8241	6.2255
15:00	3.2000	2.7578	13.8176	2.9933	6.4595	2.4372	23.8368	3.1103	2.8030	3.1535	**1.4526**
18:00	7.2000	6.9905	2.9100	6.7498	6.2524	6.7189	6.6826	7.2974	1.3528	7.2391	**0.5426**
21:00	8.9000	9.7871	9.9677	9.4114	5.7465	9.4860	6.5840	9.1831	**3.1807**	9.2807	4.2772

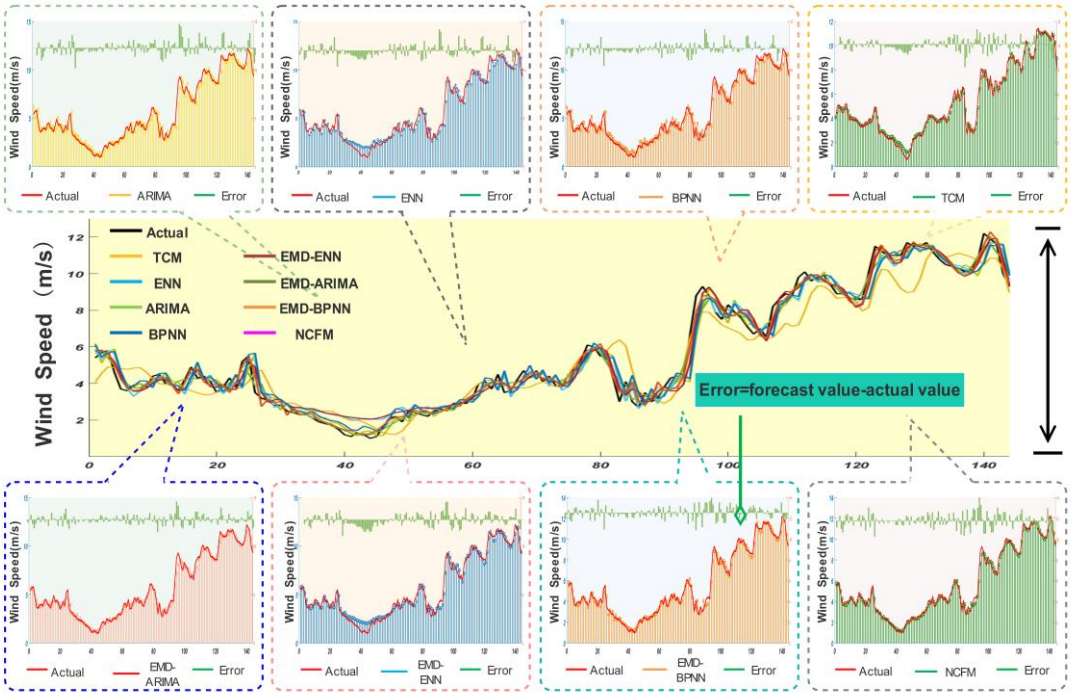

Fig.6. The forecasting results about the combined model

6. Discussion

This section provides a profound discussion of the experimental results, which includes forecasting models, each component in combined models, and the influence factors of metaheuristics.

6.1 Hybrid Model and Combined Model

Hybrid models and combined models are mainstream in the forecasting field. Hybrid models have many forms, combining algorithms with traditional models, and are the extension of single models. In the experiment, the developed hybrid model outperforms the corresponding single model, and the decreases in MAPE are 4.0471%, 2.7534%, and 8.4639% in dataset A, which verifies the validity of the hybrid models.

Combined models aim at taking full advantage of more methods to increase accuracy as much as possible. The key of a combined model is the estimation of weight coefficients. Traditional methods focus on searching for the best weight between the models; however, the weight is fixed. Therefore, this study developed the variable weight combination forecasting model considering that wind speed series changes with time. The success of the experiment proves that variable weight combination forecasting model outperforms the hybrid models in this study, decreasing the MAPE values by 0.8672%, 6.3511%, and 0.6831%.

6.2 Variable Weight Combination Forecasting Model vs Constant Weight Combination Forecasting Model

The forecasting model NCFM proposed in this paper is superior to the other models. Table 8 presents the time point data of the wind speed forecasting results and the forecasting precision of each model for 3-h intervals during one day. Then, Fig.6 shows the forecasting results of one day ahead for different models. We can conclude

the following: (1) The black bold in Table 8 represents the minimum value of each model forecasting error at the same time point; we can see that the error of the variable weight combination forecasting model (NCFM) is shown to be the smallest many times, and it performs best among the models. (2) The fitting degree between the forecasted value and the true value of each model is clearly visible and the forecasted error of the models are clearly shown in Fig.6. Even though BPNN is the most prominent in the single model forecasting, all the hybrid models are better than the corresponding single model. However, of all hybrid models, the NCFM is also the validated model with high precision.

The variable weight combination forecast method is obviously better than the fixed weight combination forecast method: (1) From Table 8, the predicted series of the two combined forecasting models are very close to the actual series; however, the variable weight combination forecasting model NCFM has a better forecasting effect. (2) By comparing the forecasting results of the description in Fig.6, it can be seen that the accuracy of the two combined models is higher than that of the hybrid models, but the NCFM has changed the equal weight characteristics of the traditional combined model, which improves the performance.

Remark: From the above analysis, we can see that the NCFM forecasting accuracy is higher, which proves that the variable weight combination forecast method is better than the fixed weight combination forecast method. The variable weight combination forecasting model can better adapt to changes of the sample, and match the weight of the sample points in the corresponding model. Therefore, the NCFM is more applicable to forecast the wind speed.

6.3 Steps of Forecasting

To testify the performance of the developed forecasting model, multi-step ahead wind speed forecasting is also conducted in this study. Table 9 compares the multi-step forecasting accuracy in spring. From the accuracy of dataset A, the one-step forecasting improves by 1.3537% and 2.5441% compared with two-step and three-step forecasting, respectively. For dataset B, the improvements with two-step and three-step forecasting are 2.4154% and 3.2371%, respectively. For dataset C, the accuracy of one-step forecasting increases 2.5220% and 3.9418% for two-step and three-step forecasting, respectively. As the comparison reveals, the proposed combined model in multi-step forecasting is effective; moreover, it is more effective in one-step forecasting.

6.4 Effectiveness of Data Preprocessing Approach

The wind speed data is irregular, which always includes high fluctuation and noise. Therefore, it is essential to eliminate the noise in the series. To validate the importance of denoising, we compare the results of ARIMA, ENN, BA-ENN, BP, BA-BPNN, and NCFM before and after denoising. The MAPE decreases by 4.05%, 2.39%, 1.41%, 8.2%, 4.89%, and 4.2% for the above models, respectively, in dataset A. From the improvement of accuracy, we can see that all MAPE decreased significantly, verifying that the accuracy improves significantly and the BPNN increases the most. Furthermore, the NCFM has the best forecasting performance, and its MAPE improves 4.2%. As Fig.7 reveals, the three metrics all decreased significantly and it demonstrates that the EMD not only enhance the forecasting accuracy but also can effectively reduce the

MAE and MSE. The results demonstrate that when data preprocessing methods were applied to denoise the original data, the forecasting accuracy improved enormously in this study.

Table 9

Comparison of multi-step forecasting accuracy

Metric	1-Step	2-Step	Improvement	3-Step	Improvement
			Dataset A		
MAPE (%)	5.7459	7.0996	1.3537	8.2900	2.5441
MAE(m/s)	0.2565	0.3268	0.0703	0.3800	0.1235
MSE(m/s)^2	0.1088	0.2033	0.0945	0.2606	0.1518
			Dataset B		
MAPE (%)	4.6075	7.0229	2.4154	7.8446	3.2371
MAE(m/s)	0.2158	0.3332	0.1174	0.3599	0.1441
MSE(m/s)^2	0.0828	0.2353	0.1525	0.2624	0.1796
			Dataset C		
MAPE (%)	4.8399	7.3619	2.5220	8.7817	3.9418
MAE(m/s)	0.2209	0.3262	0.1053	0.3800	0.1591
MSE(m/s)^2	0.0897	0.2113	0.1216	0.2877	0.1980

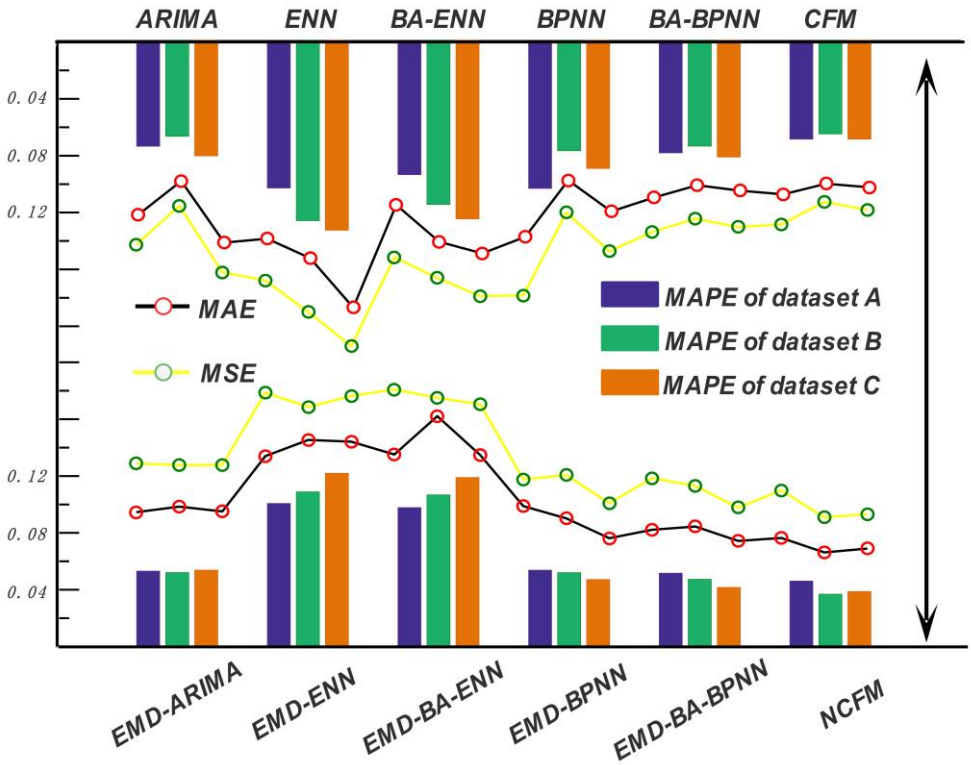

Fig.7. The metrics between denoising and without denoising

6.5 Significance of the Proposed Combination Forecast Model

The ability to predict the evaluation model can not only be based on the error and the value of MAPE, MAE, MSE. But statistical tests can also be employed to verify the forecasting ability well. Diebold and Marino [50] proposed a comparing test, which is mainly based on the comparison of two predictive models. Therefore, from the

statistical perspective, the bias–variance statistics framework and the Diebold–Mariano (DM) test are used to evaluate the significance of the developed combined model.

Table 10

Bias-variance and Diebold-Mariano test of experiments among different models for the average value of four quarters and three sites

Model	Bias²-Variance						D_t		
	Bia²(m/s)²			*Var(m/s)²*					
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
ARIMA	0.4426	0.4067	0.3964	0.4451	0.4089	0.3988	4.9588*	4.6202*	3.5508*
ENN	0.4248	0.4136	0.4138	0.4275	0.4139	0.4121	5.5181*	5.5471*	5.6011*
BPNN	0.4113	0.4215	0.4158	0.4127	0.4098	0.4086	4.9514*	4.5863*	4.6734*
EMD-ARIMA	0.1552	0.2427	0.3145	0.1562	0.2439	0.3160	2.6922*	2.4848**	2.9062*
EMD-BA-ENN	0.1721	0.4904	0.9369	0.2583	0.4915	0.9428	6.0321*	5.1339*	4.0468*
EMD-BA-BPNN	0.1236	0.2875	0.6159	0.1244	0.2895	0.6202	1.6646**	2.4441**	3.5239*
NCFM	0.1088	0.2033	0.2606	0.1093	0.2045	0.2620	-	-	-

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

The bias and variance of the bias–variance statistics framework are shown in Table 10. The Bia^2 and *var* of the three hybrid models are much lower than the three corresponding single models for wind speed forecasting. Correspondingly, the hybrid model has higher forecasting accuracy. The Bia^2 and var of the proposed model is still lower than the hybrid model, and the developed model is much better than the hybrid model in predicting ability.

Table 10 shows the DM statistics values and reveals that (a) the DM statistics values of the ARIMA, ENN, BPNN, EMD-ARIMA, and EMD-ENN models are greater than the critical value at a 1% significance level; (b) the DM statistics values of EMD-BA-BPNN model is greater than the critical value at 15% significance level;(c) the proposed model is far better than the other models when comparing DM statistics values at a 1% significance level.

6.6 Metaheuristics

In this study, metaheuristics were used to find the optimal weight and threshold values of networks to obtain high accuracy. In this section, we will focus on discussing the factors that would have an impact on the forecasting results.

6.6.1 Comparison of Different Population Size.

A comparative study of different population sizes is conducted in this section. The comparison results reveal that the performance of BA gets worse when increasing the population size of bats beyond 10 and keeping other parameters fixed. Moreover, decreasing the population size from 10, also degrades the performance. Therefore, we can conclude that the population size plays a vital role in the optimization algorithm. From Table 11, when population size is 10, the search ability of the algorithm is best.

6.6.2 Comparison of Different Train-to-Verify Ratio.

The train-to-verify ratios denote the degree of the usage of recent series and we can determine how it influences the forecasting results. We configured several train-to-verify ratios to research the effect that ratios have on the results. The ratios were configured to 1:1, 2:1, 3:1, 4:1, and 5:1 for the wind speed data of three datasets. Large ratios mean that there is more recent series put into training. However, small ratios

indicate that there are fewer recent series for training. In our experiment, we found that increasing the ratios can obtain better accuracy. The reason is that using more recent data can improve the efficiency of training. However, this does not mean that the ratios can infinitely expand in practical application, because there is a lack of reliability when there are too few samples to verify. Therefore, we suggest selecting the higher ratios when the number of samples is taken into account.

Table 11

Selection of population size of bat algorithm (MAPE)

Population Size		5	10	15	20	25
Spring	Dataset A	6.43%	**6.25%**	6.55%	6.61%	6.66%
	Dataset B	6.34%	**5.69%**	6.29%	5.91%	5.95%
	Dataset C	5.33%	**4.87%**	5.39%	5.29%	5.31%
Summer	Dataset A	5.14%	**5.03%**	5.10%	5.21%	5.30%
	Dataset B	7.70%	**7.63%**	7.67%	7.71%	7.83%
	Dataset C	7.09%	**7.05%**	7.10%	7.23%	7.29%
Autumn	Dataset A	7.06%	**6.74%**	7.02%	7.19%	7.25%
	Dataset B	6.22%	**6.14%**	6.17%	6.24%	6.19%
	Dataset C	5.51%	**5.42%**	5.47%	5.51%	5.59%
Winter	Dataset A	4.74%	**4.67%**	4.87%	4.79%	4.86%
	Dataset B	3.90%	**3.82%**	4.08%	4.27%	4.25%
	Dataset C	4.79%	**4.57%**	4.97%	4.89%	4.79%

7. Conclusion

Wind power generation is now developing rapidly in the world. Many scholars have carried out thorough research into wind speed forecasting. However, owing to the instability of wind speed, the models at this stage cannot still yield satisfactory results. This study proposed a new model to forecast short-term wind speed, which has several features:

The first is to decompose the original data and reconstruct the time series. On the basis of the denoising model and the characteristics of the decomposed series, a proper model is established to fit and forecast the short-term wind speed series. As a result, EMD makes an outstanding contribution to the forecasting of unstable time series; moreover, the fitting ability and forecasting capacity have well improved.

The second idea is that there are three models, which are from statistical models and artificial intelligence models, used to forecast the wind speed considering that wind speed data have both nonlinear and linear features. The BA is used to optimize the parameters of the models in order to improve the forecasting capacity. Accordingly, three hybrid models (EMD-ARIMA, EMD-BA-ENN, and EMD-BA-BPNN) are used in the forecasting process. The results reveal that the hybrid models show enormous improvements in the forecasting accuracy, stability, and trend.

The last feature is that the developed model integrated the three hybrid models based on a variable weight method. Although the hybrid model improves the predictive ability of the single predictive model, it still has some defects. Consequently, the combination forecast method is used to further improve the forecasting ability of the hybrid model. In view of the fact that the SVR model has a good ability for fitting and regression forecasting, the predictive results of the hybrid model are fitted and forecasted by SVR, and the final results are obtained.

In our experiments, the novel combined model NCFM that integrates three hybrid models could cope with not only the nonlinear series forecasting problem, but also certain linear series forecasting problems. Compared with the hybrid models EMD-ARIMA, EMD-BA-ENN, and EMD-BA-BPNN, the experimental results reveal that the average MAPE of the proposed combined model were significantly reduced by 23.21%, 61.56%, and 13.16% respectively. Thus, the forecasting efficiency of NCFM was verified. Therefore, the developed forecasting model NCFM, which has the highest accuracy, is a potential model for use in the future. The combined model can also be applied in many other fields, such as power load forecasting, stock price forecasting, and traffic flow forecasting.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No.71671029 and Grant No. 41475013).

Reference

[1]U.S. Department of Energy (DOE. "20% Wind Energy By 2030: Increasing Wind Energy's Contribution to US Electricity Supply." (2008).

[2] Yan, Jie, et al. "Reviews on uncertainty analysis of wind power forecasting." Renewable & Sustainable Energy Reviews 52(2015):1322-1330.

[3] Kavasseri, Rajesh G., and K. Seetharaman. "Day-ahead wind speed forecasting using f -ARIMA models." Renewable Energy 34.5(2009):1388-1393.

[4] Jung, Jaesung, and R. P. Broadwater. "Current status and future advances for wind speed and power forecasting." Renewable & Sustainable Energy Reviews 31.2(2014):762-777.

[5] Cadenas, Erasmo, and W. Rivera. "Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model." Renewable Energy 35.12(2010):2732-2738.

[6] Babu, C. Narendra, and B. E. Reddy. "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data." Applied Soft Computing 23.10(2014):27-38.

[7] Mengjiao Qin, Zhihang Li, and Zhenhong Du. "Red tide time series forecasting by combining ARIMA and deep belief network." Knowledge-Based Systems.

[8] Kordanuli, Bojana, et al. "Appraisal of artificial neural network for forecasting of economic parameters." Physica A Statistical Mechanics & Its Applications 465(2016):515-519.

[9] Reyes-Alvarado, Luis C., et al. "Forecasting the effect of feast and famine conditions on biological sulphate reduction in an anaerobic inverse fluidized bed reactor using artificial neural networks." Process Biochemistry (2017).

[10] Liu, Hui, et al. "New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks." Renewable Energy 83(2015):1066-1075.

[11] Wang, Jujie, et al. "Forecasting wind speed using empirical mode decomposition and Elman neural network." Applied Soft Computing 23.5(2014):452–459.

[12] Li, Penghua, et al. "Application of a hybrid quantized Elman neural network in short-term load forecasting." International Journal of Electrical Power & Energy Systems 55.2(2014):749-759.
[13] Hongmei Liu, Jichang Zhang, and Chen Lu. "Performance degradation forecasting for a hydraulic servo system based on Elman network observer and GMM–SVR." Applied Mathematical Modelling 39(2015):5882–5895.
[14] Xu, Zhun, et al. "A Monte-Carlo-based network method for source positioning in bioluminescence tomography." Journal of Biomedical Imaging 2007.2(2007):48989.
[15] Ren, Chao, et al. "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting." Knowledge-Based Systems 56.3(2013):226-239.
[16] Wang, Lin, Y. Zeng, and T. Chen. "Back propagation neural network with adaptive differential evolution algorithm for time series forecasting." Expert Systems with Applications 42.2(2015):855-863.
[17] Huang, Norden E, Z. Shen, and S. R. Long. "A new view of nonlinear water waves: the Hilbert spectrum." Annual Review of Fluid Mechanics 31.1(2003):417-457.
[18] Yu, Lean, et al. "Linear and nonlinear Granger causality investigation between carbon market and crude oil market: A multi-scale approach." Energy Economics 51(2015):300-311.
[19] Wang, Shouxiang, et al. "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method." Renewable Energy 94(2016):629-636.
[20]Yi, Siqi, K. Guo, and Z. Chen. "Forecasting China's Service Outsourcing Development with an EMD-VAR-SVR Ensemble Method ☆." Procedia Computer Science 91(2016):392-401.
[21]Wang, Jie, and J. Wang. Forecasting stochastic neural network based on financial empirical mode decomposition. Elsevier Science Ltd. 2017.
[22] Liu, Hui, et al. "Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks." Applied Energy 157(2015):183-194.
[23] Jiang, Ping, Y. Wang, and J. Wang. "Short-term wind speed forecasting using a hybrid model." Energy (2016).
[24] Liu, Jinqiang, X. Wang, and Y. Lu. "A Novel Hybrid Methodology for Short-term Wind Power Forecasting Based on Adaptive Neuro-fuzzy Inference System." Renewable Energy (2017).
[25] Wang, Shouxiang, et al. "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method." Renewable Energy 94(2016):629-636.
[26] Liang, Zhengtang, et al. "Short-term wind power combined forecasting based on error forecast correction." Energy Conversion & Management 119(2016):215-226.
[27] Xiao, Ling, et al. "Combined forecasting models for wind energy forecasting: A case study in China." Renewable & Sustainable Energy Reviews 44.17(2015):271-288.
[28] Barrow, Devon K., and S. F. Crone. "A comparison of AdaBoost algorithms for time series forecast combination." International Journal of Forecasting 32.4(2016):1103-1119.

852 [29] Zhang, Wenyu, et al. "A combined model based on CEEMDAN and modified
853 flower pollination algorithm for wind speed forecasting." Energy Conversion &
854 Management 136(2017):439-451.
855 [30] Santamaria-Bonfil, Guillermo, A. Reyes-Ballesteros, and C. Gershenson. "Wind
856 Speed Forecasting For Wind Farms: A Method Based on Support Vector Regression."
857 Renewable Energy 85(2015):790–809.
858 [31] Chakri, Asma, et al. "New directional bat algorithm for continuous optimization
859 problems." Expert Systems with Applications 69(2016):159-175.
860 [32] Chen, H. The Validity of the Theory and Its Application of Combination Forecast
861 Methods; Science Press: Beijing, China, 2008. (In Chinese)
862 [33] Zhang, Chi, et al. "Short-term wind speed forecasting using empirical mode
863 decomposition and feature selection." Renewable Energy 96(2016):727-737.
864 [34] Lei, Yu, H. Cai, and D. Zhao. "Improvement of the forecasting accuracy of polar
865 motion using empirical mode decomposition." Geodesy & Geodynamics (2017).
866 [35] Yang, Xin She. "A New Metaheuristic Bat-Inspired Algorithm." Computer
867 Knowledge & Technology 284(2010):65-74.
868 [36] Yang, Xin She. Nature-Inspired Optimization Algorithms. Elsevier Science
869 Publishers B. V. 2014.
870 [37] Wang, Lin, Y. Zeng, and T. Chen. "Back propagation neural network with adaptive
871 differential evolution algorithm for time series forecasting." Expert Systems with
872 Applications 42.2(2015):855-863.
873 [38] Shukur, Osamah Basheer, and M. H. Lee. "Daily wind speed forecasting through
874 hybrid KF-ANN model based on ARIMA." Renewable Energy 76(2015):637-647.
875 [39]Sen, Parag, M. Roy, and P. Pal. "Application of ARIMA for forecasting energy
876 consumption and GHG emission: A case study of an Indian pig iron manufacturing
877 organization." Energy 116.Part 1(2016):1031-1038.
878 [40] Santos, Nicolau, and R. Rui. Performance of state space and ARIMA models for
879 consumer retail sales forecasting. Pergamon Press, Inc. 2015.
880 [41] Conflitti, Cristina, C. D. Mol, and D. Giannone. "Optimal combination of survey
881 forecasts." International Journal of Forecasting 31.4(2012):1096-1103.
882 [42] Drucker, Harris, et al. "Support Vector Regression Machines. " Advances in Neural
883 Information Processing Systems 28.7(2008):779-784.
884 [43] Shi, Jing, J. Guo, and S. Zheng. "Evaluation of hybrid forecasting approaches for
885 wind speed and power generation time series." Renewable & Sustainable Energy
886 Reviews 16.5(2012):3471-3480.
887 [44] Xu, Yunzhen, W. Yang, and J. Wang. "Air quality early-warning system for cities
888 in China." Atmospheric Environment 148(2016).
889 [45]Yang, Yi, et al. "Modelling a combined method based on ANFIS and neural network
890 improved by DE algorithm: A case study for short-term electricity demand forecasting."
891 Applied Soft Computing 49(2016):663-675.
892 [46]Liu, Li, et al. "A Rolling Grey Model Optimized by Particle Swarm Optimization
893 in Economic Forecasting." Computational Intelligence 32.3(2016):391-419.
894 [47] Cadenas, Erasmo, and W. Rivera. "Short term wind speed forecasting in La Venta,
895 Oaxaca, México, using artificial neural networks." Renewable Energy 34.1(2009):274-

278.
[48] 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.
[49] Tsay, Ruey S. Analysis of Financial Time Series. Analysis of financial time series: Wiley, 2002:5880-5885.
[50] Diebold, Francis X, and R. S. Mariano. "Comparing Predictive Accuracy." Journal of Business & Economic Statistics 20.1(2002):134-144.