



A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm

Yanbin Cui¹ · Chenchen Huang² · Yanping Cui³

Received: 17 September 2019 / Accepted: 12 December 2019 / Published online: 28 December 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Wind power, a clean and renewable resource, is regarded as one of the most promising and economical resources during the transformation from fossil fuels to new energy resources. Thus, the accuracy of wind speed forecasting work is very important to integrate the wind resource into electrical power system on a large scale. To improve the short-term wind speed forecasting accuracy, a novel compound model is introduced in this paper. For the proposed model, the fast ensemble empirical mode decomposition method was employed to do the data preprocessing. After the data preprocessing, phase space reconstruction was used for choosing each sub-series' input and output vectors for the forecasting model dynamically. Then, the bat algorithm was applied to optimize the connection weights and thresholds of the traditional back propagation neural network. The forecasting results can be obtained through the aggregation of sequential prediction. The performance evaluation of this proposed model indicates that it can capture the nonlinear characteristics of the wind speed signal efficiently. The proposed model shows better performance when being compared with the parallel models.

Keywords Wind speed forecasting · Fast ensemble empirical mode decomposition · Phase space reconstruction · Bat algorithm · Back propagation neural network

Introduction

With the rapid economy development in twenty-first century, the energy consumption keeps increasing over the years. Nowadays, the energy resources have been treated as one of the fundamental factors which impact the economy development of all countries of the world. In the *IEA world energy outlook 2017*, the global energy demands will increase by 30% now until the end of 2040. With the development of renewable energy, the structure of energy consumption will change dramatically, and the renewable energy may become

the backbone of the energy structure. The technical stacks of wind power, one of the renewable energy resources, have made a great progress to countries all over the world thanks to the widespread distribution, mature developing techniques, and non-pollution characters of the wind power. By the end of 2017, the global total rated capacity of wind power generations has reached 539,291 MW. And the annual growth is as high as 10.8% which is 52,552 MW rated capacity for the year 2017. As the largest developing country in the world, China has also made its own efforts to develop and utilize the wind energy resources. In 2010, the newly rated capacity of wind power of China has reached at 44.7 GW, making China the largest rated wind generation capacity of the world. However, there are also some problems that need to be addressed during the wind power development. The wind speed, which produces random, intermittent, uncontrollable output, may affect the operation safety and stability of the power grid negatively. To solve this problem, an accurate short-term wind speed forecasting is needed. The high precision of the short-term wind speed forecasting not only can help the system operator to do the dispatching work accordingly in real time but also can reduce the spinning capacity of the system which benefits in economic performance.

Responsible editor: Marcus Schulz

✉ Chenchen Huang
973118497@qq.com

¹ Department of Mechanical Engineering, North China Electric Power University, Baoding 071000, China

² Department of Economic Management, North China Electric Power University, Baoding 071000, China

³ University of Victoria, Victoria, Canada

In general, the wind speed forecasting can be classified into three categories in terms of time scale: long-term forecasting, medium-term forecasting, and short-term forecasting. The long-term wind speed forecasting is mainly used to help the wind power enterprises to develop annual generation strategy or assess a new wind energy project before the construction. The medium-term forecasting is used as a reference for the wind farms in developing unit maintenance and repair plans. The short-term wind speed forecasting which is referring to the forecast for the next 3 days (72 h) which plays an important role during power dispatching cooperation as well as ensuring the operation safety and stability of the power system.

From another perspective, the wind speed forecasting can be classified into three categories in terms of the forecasting approaches: physics methods, statistic-theory-based data-driven/artificial intelligence approaches. The numerical weather prediction (NWP) is one of the most commonly used physics method of wind speed forecasting. Hoolohan et al. (2018) presented a hybrid numerical weather prediction model (NWP) and a Gaussian process regression (GPR) model for near ground wind speed prediction up to 72 h ahead using data partitioned on atmospheric stability level to improve performance of proposed model. Considering that the hybrid methods can outperform statistical methods, particularly with longer forecasting timeframe, the author investigated how the atmospheric stability level can affect the NWP forecasting accuracy and if the forecasting accuracy can be improved with data partitioning. The result showed that the power forecasting can be enhanced with the wind predicting method proposed in this paper. The advantage of physics methods is that it does not require for much historical data, and the geological and meteorological factors of the wind farm area can be well described. However, since it is difficult to do comprehensive meteorological factors collection and analysis in a very short time, the physics predicting techniques fit the medium or long-term wind speed forecasting better rather than the short term.

Different from the physics methods, the statistical models which put more focus on analyzing the wind speed in time series, has more applications when it comes to short-term wind speed forecasting. Auto-regression model (AR), moving average model (MA), auto-regressive moving average model (ARMA), and auto-regressive integrated moving average model (ARIMA) are four of the most commonly used models in industry. Lydia et al. (2016) indicated that it is needed to develop the models which can do the hourly wind speed forecasting with higher reliability and greater accuracy. Thus, the authors built a hybrid model based on linear and non-linear auto-regressive moving average models with and without external variables to predict wind speed from 10-min intervals up to 1 h. A genetic algorithm-support vector regression model (GA-SVR) model was proposed by Santamaria-Bonfil et al. (2016) to predict the wind speed, where the genetic algorithm

is employed to tune the parameters of support vector regression which is used for predicting the wind speed. Based on the auto regressive moving average (ARMA) method, Erdem and Shi (2011) proposed four approaches to perform the wind speed and wind direction forecasting. Besides the AR-based models, the support vector regression (SVR), and Gaussian process regression (GPR) are also widely used statistical methods. A hybrid model was proposed by Zhang et al. (2016) which combines the auto regressive model and Gaussian process regression for probabilistic wind speed forecasting. The results of this work indicate that the proposed method can not only improve the forecasting accuracy at single point but also produce good forecasting intervals when it is compared with other techniques.

The time series models require less data with simple and fast operation, but there are still some existing defects for such methods. First, the models are usually based on a certain statistical assumption, while in practical, the wind speed data does not follow the assumption very well. Second, there is no clear standard during the process of model identification and parameter estimation, which makes the performance affected by great human subjectivity.

With the development of computer technology, the artificial intelligence neural networks are employed by researchers to solve the existing problems. In recent years, the artificial-intelligent-based methods have increasing number of deployments in wind power predicting works. Zhang et al. (2018) presented a wind speed forecasting method using radial basis function (RBF) neural network based on principal component analysis (PCA) and independent component analysis (ICA). Compared with the traditional neural network models, the model in this work reduced the effect of wind speed sequence volatility and enhanced the short-term wind speed forecasting accuracy. Lazarevska (2016) introduced an approach to forecast the wind speed based on extreme learning machine (ELM). According to the analysis result from this paper, the ELM is a simple method with good approximation performance, and fast computation. Back propagation (BP) neural network is another method that has been widely used in wind speed forecasting due to its strong function reproducibility. Guo et al. (2011) proposed a hybrid wind speed forecasting method based on BP neural network. This method is able to predict the average daily wind speed for the next year, and can obtain relatively low average absolute error in its case studies. The wind speed forecasting model proposed by Sun and Wang (2018) combines fast ensemble empirical mode decomposition (FEEMD), sample entropy, phase space reconstruction, and BP neural network. The experimental results of this work show better performance than the regular model in short-term wind speed forecasting. Liu et al. (2017) proposed a novel hybrid methodology which combines three individual forecasting models, the adaptive neuro-fuzzy inference system, the BP neural network, and the radial basis function (RBF)

neural network. It has been proved in practice that BP neural network can accurately approximate any non-linear function which result in better forecasting results in wind speed predicting. Therefore, this paper chooses back propagation neural network as the main forecasting method.

However, simply using the neural network model for wind speed forecasting cannot meet the requirements for the accuracy of wind speed forecasting. Thus, optimization algorithms are often employed to improve the accuracy of the predicting model. Jiang et al. (2018) used Cuckoo search algorithm (CSA) and harmony search algorithm (HSA) to optimize the parameters of BP neural network which improve the predicting accuracy in short-term wind speed forecasting. Su et al. (2014) introduced a wind speed forecasting model based on ARIMA. Particle swarm optimization (PSO) was used to optimize the model parameters. The optimization algorithm improves the fitting accuracy and avoids over-fitting. Zhou et al. (2018) optimized the model parameters by implementing the backtracking search algorithm (BSA). Performance results show that the optimization algorithm improves the performance of the short-term wind speed forecasting model. In recent years, a meta-heuristic optimization algorithm called bat algorithm (BA) has been proved to have better performance come to genetic algorithm (GA), particle swarm optimization (PSO), and so on (Mirjalili et al., 2014). Yang and Gandomi (2012) have proved that BA is better than the existing methods in optimization. The conjugate gradient (CG) method has been proposed by Xiao et al. (2017) to improve the BA performance by optimizing the initial weight of the neural network. Wu and Peng (2016) also utilized BA to optimize the parameters of least squares support vector machine (LSSVM) model which improves the performance of the model. Tavakkoli et al. (2015) is another work which combined BA, GA, and PSO with SVR in case studies. The experimental results show that the SVR model optimized by BA performed better than the model optimized by other methods. Since the BA shows better performance in optimizing the parameters of the neural network, the BA will be employed in this work to improve the forecasting accuracy.

With the further study about the wind power forecasting, scholars found that the chaotic nature and inherent complexity of wind speed in time series can affect the stability of forecasting model. Thus, to construct a stable forecasting model, it is necessary to make analysis of the original data features. The wavelet transform (WT), variational mode decomposition (VMD), singular spectrum analysis (SSA), empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and fast ensemble empirical mode decomposition (FEEMD) had all been used in previous signal analyzing research works for wind speed in time series.

In this work, the authors decomposed the wind speed signal into two components by WT: an approximation signal to maintain the major fluctuations and a detail signal to remove the

stochastic volatility (Liu et al., 2014). Ali et al. (2018) presented a method for short-term wind speed forecasting by using a hybrid approach based on VMD in conjunction with both linear and nonlinear predicting models. After being decomposed into multiple intrinsic narrow band components, the data was predicted using ARIMA and artificial neural network. Another short-term wind speed and wind power forecasting model proposed by Liu et al. (2018) is singular spectrum analysis (SSA) and locality-sensitive hashing (LSH) based, where SSA is for decomposing and LSH for selecting similar segments of the mean trend segments in order to enhance the accuracy and efficiency of forecasting. Yu et al. (2017) came up with EMD to decompose the original sequence into several Intrinsic Mode Functions (IMFS) and a residue value generated due to the non-stationary and nonlinear characteristics of wind speed sequences. To overcome the shortcoming of EMD mixed mode, Wu and Huang (2009) improved the EMD by adding white noise into original series and named it as EEMD. Santhosh et al. (2018) preprocessed the original wind speed data through EEMD, and the results revealed that the EEMD data preprocessing can enhance the forecasting accuracy. Compared with EMD, EEMD has a better performance in the decomposition of non-stationary signals. In 2014, Wang (2014) introduced a method to optimize the real-time computational performance of EEMD which is known as FEEMD. Jiang and Ma (2016) used the FEEMD as noise canceling method to help forecasting the key indicators in electric power system, including short-term wind speed, electrical load, and electricity price. Liu et al. (2015) combined FEEMD, mind evolutionary algorithm (MEA), GA, and multilayer perceptron neural network (MLP) together for the wind speed multi-step forecasting. The experimental results indicate that FEEMD could address the wind speed fluctuation efficiently and improve the performance of the MLP neural networks significantly. Thus, FEEMD is chosen in this work to preprocess the original wind speed data. EMD as well as WT are also conducted for a comparison purpose.

As mentioned above, the wind speed forecasting accuracy can be improved by applying signal decomposition algorithms. However, other characteristics of wind speed like chaotic should also be considered. The chaos is one of the universal natural phenomena which seem random but has its internal rules. To consider the chaotic characteristic of wind speed, phase space reconstruction (PSR) is employed into wind speed forecasting. Han and Liu (2015) studied the online forecasting of short-term wind speed and power generation at wind farm based on PSR. After reconstructing the sample space, the short-term wind speed is predicted by BP neural network. Gao et al. (2013) combined the chaos PSR and numerical weather prediction (NWP) for a short-term wind speed forecasting in which the historical wind speed data reconstructed by PSR and the external factors from NWP data are used as input, and the general regression neural network

(GRNN) was developed for forecasting. Compared with the chaos GRNN model, NWP GRNN model and Persistence Model, the method presented in this work shows an improvement in forecasting accuracy.

To consider the volatility and chaotic characteristics of wind speed, this work combined the fast ensemble empirical mode decomposition, phase space reconstruction, and neural network for short term wind speed forecasting. In the developed FEEMD-PSR-BA-BPNN model, the FEEMD technology is employed to decompose the raw wind speed time sequence into a series of more stable subsequences, and the phase space reconstruction (PSR) is applied to determine the input and output matrix of neural network. The back propagation neural network (BPNN) with two hidden layers is utilized for forecasting, and the bat algorithm (BA) is introduced for the optimization of connection weights and thresholds of BPNN to enhance the forecasting stability of model. There are few relevant literatures on wind speed forecasting which proposed the hybrid model as what in this paper. This paper will fill this gap. The rest of the paper is organized as follows.

“Materials and methods” section gives a brief description about the fast ensemble empirical mode decomposition, phase space reconstruction and the back propagation neural network optimized using bat algorithm; a holonomic short-term wind speed forecasting framework is introduced in “The framework of the hybrid model” section; “Data preprocessing” section is designed for data preprocessing and result forecasting when discussion are provided in “Short-term wind speed forecasting” section; final conclusion is shown in “Conclusions” section.

Materials and methods

This part gives an introduction of all methods and algorithms which will be utilized in the proposed hybrid model.

Fast ensemble empirical mode decomposition

The fast ensemble empirical mode decomposition (FEEMD) was derived from empirical mode decomposition (EMD). FEEMD is a time domain signal decomposing method which is proposed by the scientist N. E. Huang (1998). The calculating steps of EMD are shown as follows:

- 1) Determine the maximum value $x_{\max}(t)$ and minimum value $x_{\min}(t)$ of the original time series $x(t)$, and the upper envelope $U_i(t)$ and lower envelope $L_i(t)$ of $x(t)$ are constructed through interpolation method.
- 2) Calculate the mean curve $m_i(t)$:

$$m_i(t) = \frac{U_i(t) + L_i(t)}{2} \quad (1)$$

- 3) Subtract $m(t)$ from $x(t)$ to get $h(t)$

$$h_i(t) = x_i(t) - m_i(t) \quad (2)$$

- 4) There are two conditions that needs to be met to determine if $h(t)$ is an intrinsic mode function (IMF): (1) the difference between the number of extreme points and the number of zero points is less than or equal to one; (2) the mean sequence $m_i(t)$ approaches 0. If all abovementioned conditions are met, an intrinsic mode function IMF_i is obtained, otherwise, return to the first step.
- 5) Calculate the remaining data:

$$r_i(t) = x_i(t) - h_i(t) \quad (3)$$

Repeat the above steps to obtain number of n IMF components and one residual component R :

$$x(t) = \sum_{i=1}^n IMF_i(t) + R \quad (4)$$

To improve the EMD, the fast ensemble empirical mode decomposition added different Gaussian white noise series implementation into the original signal. This step overcomes the mode-mixing problem existing in EMD algorithm and speeds up the computing process. Two important parameters need to be determined in the process of FEEMD algorithm, the amplitude k of white noise and the number of iterations M . The specific steps of the FEEMD algorithm are as follows:

- 1) The white noise $n_m(t)$ is randomly added to the original sequence $x(t)$, and the new sequence $x_m(t)$ can be obtained
- $$x_m(t) = x(t) + n_m(t) \quad (5)$$
- 2) According to the decomposition steps of EMD, $x_m(t)$ is decomposed into several IMFs and residual sequences R , which are expressed as $z_{i,m}(t)$ and $r_{i,m}(t)$, respectively.
 - 3) Repeat the first step and the second step with different frequency of white noise until the number of iterations reach the number of M .
 - 4) Compute the mean value to get the final IMF components and residual component R :

$$IMF_i(t) = \frac{\sum_{m=1}^M z_{i,m}(t)}{M} \quad (6)$$

$$R_i(t) = \frac{\sum_{m=1}^M r_{i,m}(t)}{M} \quad (7)$$

Considering the characteristic of wind speed data, this paper sets the amplitude k of white noise as 0.05–0.5 times and the iteration time M as 100, and please refer to Liu et al. (2015) for details on the decomposition procedure.

Phase space reconstruction

The chaos theory describes the irregularity and randomness of chaos phenomena which exist in the nature and human society universally. To address the nonlinearity and chaotic characteristic of the wind speed in time series, this paper reconstructs the phase space of each subsequence by the phase space reconstruction which originates from the Chaos theory to eliminate the disorder. And in this way, the input and corresponding output vectors of the extreme learning machine can be determined.

In general, for the time series $X = \{x_1, x_2 \dots x_N\}$, the reconstructed phase space is presented as follows:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{N-m+1} \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \dots & x_{1+(m-1)\tau} \\ x_2 & \ddots & & x_{2+(m-1)\tau} \\ \vdots & & & \vdots \\ x_{N-m+1} & x_{N-m+1+\tau} & \dots & x_N \end{bmatrix} \quad (8)$$

Use the above matrix as input, the output of the model is expressed as:

$$\begin{bmatrix} x_{2+(m-1)\tau} \\ x_{3+(m-1)\tau} \\ \vdots \\ x_N \\ x_{N+1} \end{bmatrix} \quad (9)$$

It can be seen from Eqs. (8) and (9) that there are two parameters τ and m in the matrix. The two parameters are called as delay time and minimum embedding dimension, respectively, which play crucial roles for the PSR. After certain time delay τ , the original time series can be treated as an independent coordinate. But, a small value of τ leads to the two components contained in phase space too close in value to be distinguished which fails to provide two independent coordinate components. On contrary, a big value of τ can lead to two independent components and the trajectory projection of the chaotic attractor in two directions are irrelevant from each other. The purpose of embedding dimension is to produce a topology equivalent thing between the primitive attractor and chaotic attractor. Thus, this paper will give a brief introduction about the techniques which are employed to determine the value of τ and m .

The mutual information is presented by Fraser and Swinney (1986) to determine the nonlinear correlation of the system. For the two independent discrete time series $X = \{x_1, x_2, \dots, x_m\}$ and $Y = \{y_1, y_2, \dots, y_n\}$, their information entropy can be obtained based on the information theory:

$$H(X) = - \sum_{i=1}^m P(x_i) \log P(x_i) \quad (10)$$

$$H(Y) = - \sum_{j=1}^n P(y_j) \log P(y_j) \quad (11)$$

where $P(x_i)$ and $P(y_j)$ are the probability of x_i and y_j , respectively.

On the value of given X, the information that we know about system Y is denoted as the mutual information between X and Y.

$$I(X, Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y) \quad (12)$$

$$H(X, Y) = - \sum_{i=1}^m \sum_{j=1}^n P_{x,y}(x_i, y_j) \log P_{x,y}(x_i, y_j) \quad (13)$$

where the $H(X, Y)$ represents the combined information entropy of X and Y.

The value of $I(X, Y)$ represents the correlation between Y and a fixed value of X. The smaller the value, the greater the non-correlation between each other will be. Hence, the first minimum value of the function $I(X(i), X(i + \tau))$ is considered as the optimal delay time, which can be regarded as expressing the irrelevance between the time series $X(i)$ and the $X(i + \tau)$ with τ delay time. Please refer to Ma (2013) for the details on this method.

For the embedding dimension selection, on the one hand, it needs to ensure that all the kinds of chaotic invariants can be computed accurately, while in another hand, it needs to minimize the impact of noise and computation load. This paper applies the Cao method, an improvement derived from false nearest neighbors (FNNs), to determine the parameter m . Two non-adjacent points in higher dimensional space could be close to each other when they are projected to one-dimensional space attributes to the disorder of chaotic time series. With the increase of embedding dimension, the trajectory of chaotic motion will be gradually open as well as the rejection of the false nearest neighbors. However, Cao has proposed an enhanced false nearest neighbors method due to the FNNs which is sensitive to the signal noises. By taking the advantage of distinguishing between stochastic and deterministic signals as well as the usability when applied to small dataset, this paper utilizes the Cao method to confirm the value of m . The detail information about this method is presented in Cao (1997).

Back propagation neural network

The back propagation neural network is a multilayer feed-forward network based on an error back-propagation algorithm. The BP neural network consists of input layer, hidden layer, and output layer. There are two outstanding features about it: forward information transmission and error back propagation. The input signal is transmitted and processed from the input layer to the hidden layer and then to the output layer. The neural status of each layer will only affect the next layer. If there is no expected output result from output layer, the network will switch to the back propagation. The weight and threshold will be adjusted accordingly to the forecasting

error which will make the output moving closer to the desired output.

The hidden layer can be classified into single layer or multilayer according to the number of layers. Compared with single layer, the multiple hidden layers have better generalization capability as well as higher accuracy. In this work, two hidden layers are applied in BP neural network for short-term wind speed forecasting. The topology structure is showed in Fig. 1. where the X_1, X_2, \dots, X_n indicate the n input values while the Y_1, Y_2, \dots, Y_m are the m output results.

The bat algorithm

Based on the echolocation of microbats, Yang (2010) proposed a new meta-heuristic method which is called bat algorithm. This algorithm combines the advantages of both the genetic algorithm and particle swarm optimization with the superiority of parallelism, quick convergence, well distributed, and less parameter adjustment. In d dimensions of search space during the global search, the BA i has the position of x_i^t , and velocity v_i^t at the time of t , whose position and velocity is updated as Eqs. (14) and (15), respectively:

$$x_i^{t+1} = x_i^t + v_i^{t+1}; \quad (14)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x) \times F_i \quad (15)$$

where x^t is the current global optimal solution; and F_i is the sonic wave frequency which can be obtained through Eq. (16):

$$F_i = F_{\min} + (F_{\max} - F_{\min})\beta \quad (16)$$

where β is a random number between 0 and 1; F_{\max} and F_{\min} are the maximum and minimum sonic wave frequencies

of the bat i . In the process of flying, each initial bat is assigned to one random frequency between F_{\min} and F_{\max} .

In local search, once a solution is selected in the current global optimal solution, each bat would produce a new alternative solution in the mode of random walk according to Eq. (17):

$$x_n(i) = x_0 + \mu A^t \quad (17)$$

where x_0 is a solution that is chosen in current optimal disaggregation randomly; A^t is the average volume of the current bat population; and μ is a D dimensional vector within $-1, 1$.

The balance of bats is controlled by the impulse volume $A(i)$ and impulse emission rate $R(i)$. Once the bat locks the prey, the volume $A(i)$ will be reduced and the emission rate $R(i)$ will be increased at the same time. The update of $A(i)$ and $R(i)$ are expressed as Eqs. (18) and (19), respectively:

$$A^{t+1}(i) = \gamma A^t(i) \quad (18)$$

$$R^{t+1} = R^0(i) \times (1 - e^{-\theta t}) \quad (19)$$

where γ and θ are both constants and that γ is within 0, 1 and $\theta > 0$. In this paper, the two parameters as $\gamma = \theta = 0.9$.

The framework of the hybrid model

In this section, an overview about the proposed model will be shown. The flowchart of the proposed method is presented in Fig. 2. In this flowchart, the original data is decomposed into several IMF components and one residual component R by FEEMD, which eliminates the volatility and randomness of the original wind speed raw data in part 1. By processing each IMF component and residual component R using the phase

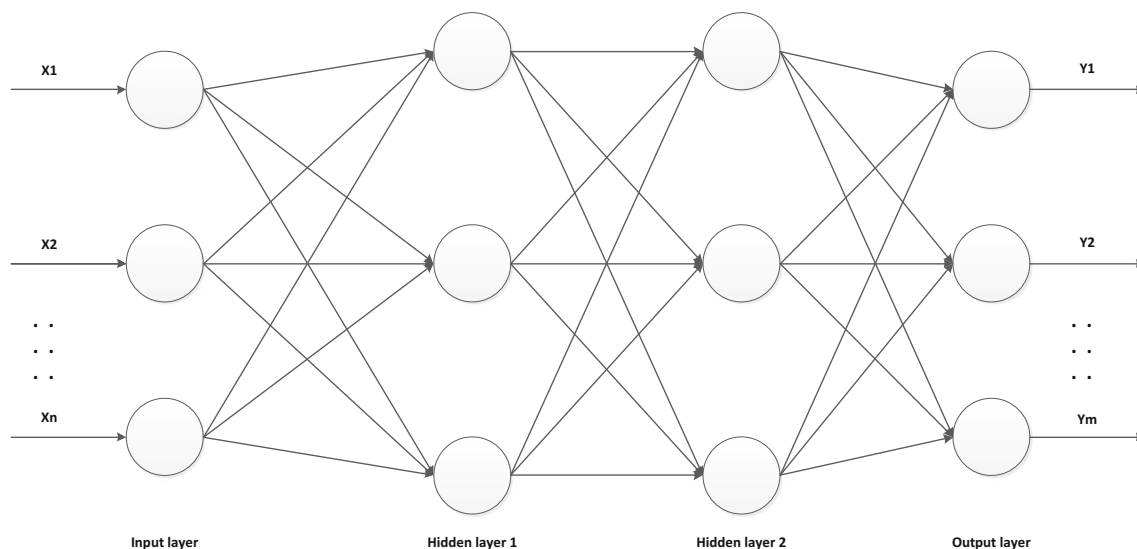


Fig. 1 The topology structure of BP neural network

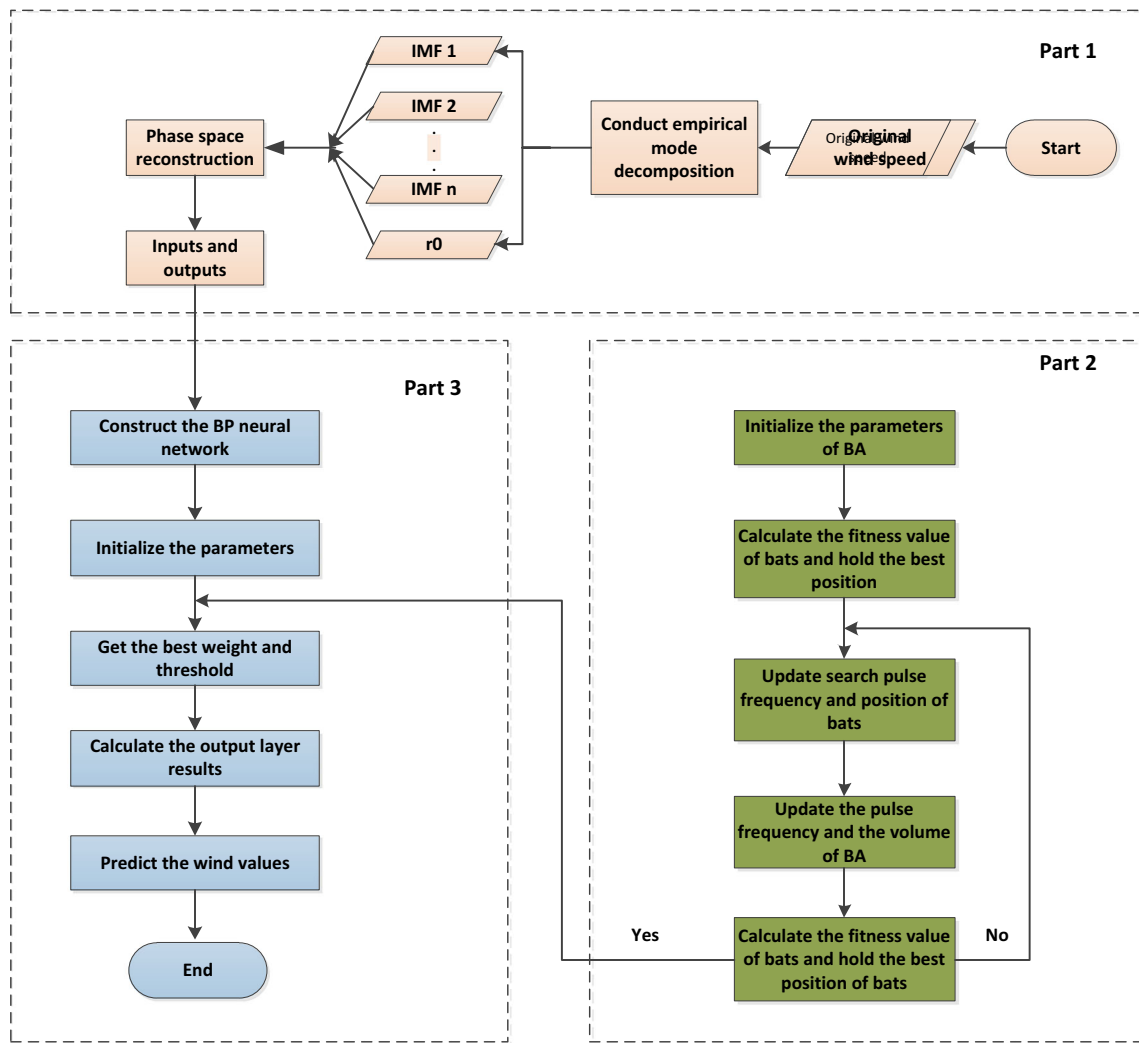


Fig. 2 The overview flowchart of the BA-BPNN based model

space reconstruction method, the input and output vectors of the neural network can be determined. In the next step, the BP neural network which is optimized by BA is used to forecast each component. In the last step, the forecasting results of each

component will be accumulated to get the final forecasting value. In this flowchart, part 2 demonstrates the workflow of the BA while the forecasting procedure of the BP neural network with optimized parameters is illustrated in part 3.

Fig. 3 The original wind speed in time series

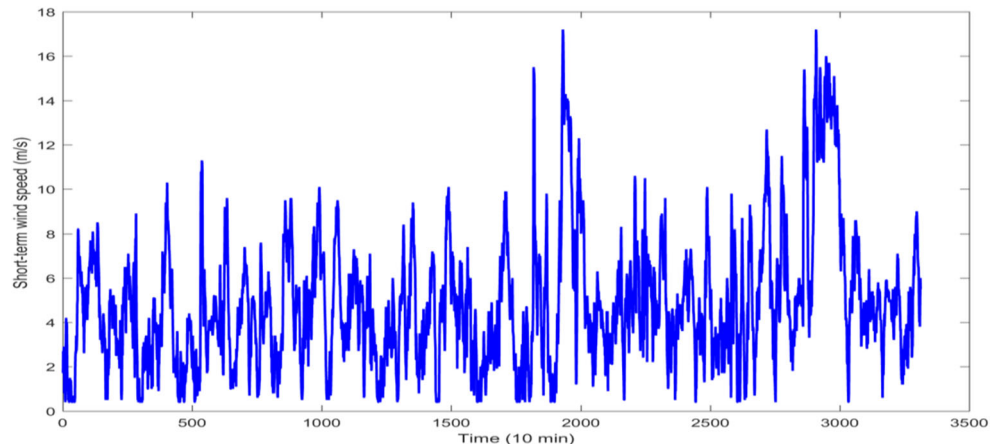
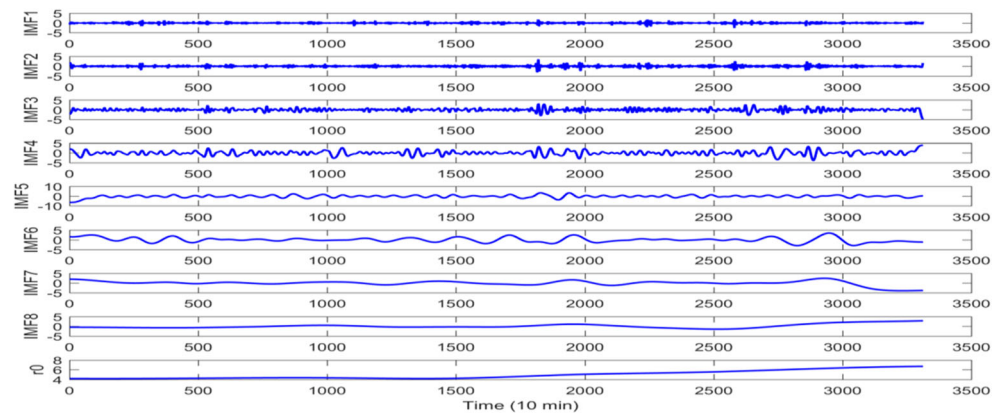


Fig. 4 The decomposition result of FEEMD



Data preprocessing

To evaluate the performance and practicability of the proposed hybrid model, the data preprocessing will be described in detail in this section. The data spans from March 1 to March 23, 2017. The time interval of wind speed data is 10 min, and the total number of data is 3312. Among them, the data of the first 20 days (2880 in total) is employed as training samples to build forecasting model. The remaining 3 days of data (432 in total) is used as test set to evaluate the performance of the forecasting model. The training set includes about 85% of the total data. The measuring height of wind speed is 100 m.

Data decomposition

The wind speed raw data in time series is shown in Fig. 3, which shows 3312 data points in total over 23 days.

It can be seen from Fig. 3 that there is no obvious wind speed regulation information for severe fluctuation. Therefore, this paper utilizes FEEMD to decompose the original data, so as to reduce the non-stationary characteristics of the original wind speed in time series, and then digs out the available information for forecasting. Considering the feature of dataset

the paper chose, the decomposition result consists of 8 intrinsic mode functions and 1 residue, which is presented in Fig. 4.

To further evaluate the effectiveness of the proposed model, empirical mode decomposition (EMD) and wavelet transform (WT) are used as comparison. The results of EMD and WT decomposition are shown in Fig. 5 and Fig. 6, respectively. In Fig. 5, the EMD method decomposes the raw wind data into 8 IMFs and one residual sequence. Figure 6 presents three sections that are all obtained from the WT process which are the original wind speed raw data in time series, the detail D1 and the approximation A1, respectively. In theory, the approximation A1 gives a smoother performance since the high frequency components of the wind speed are included in detail D1 results. Hence, the D1 result is abandoned and the A1 is chosen for the short-term wind speed forecasting result.

Phase space reconstruction

The phase space reconstruction plays a crucial role in determining the input and the output vectors of forecasting model. As previously mentioned, the phase space will be reconstructed on the basis of the two parameters, time delay τ and the embedding dimension m . Table 1 shows the values of the time delay τ and the embedding dimension m of each subsequence

Fig. 5 The decomposition result of EMD

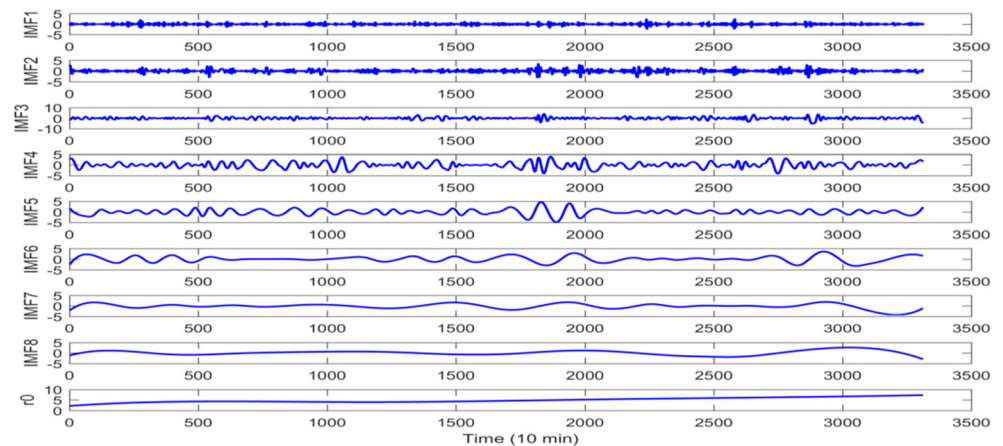
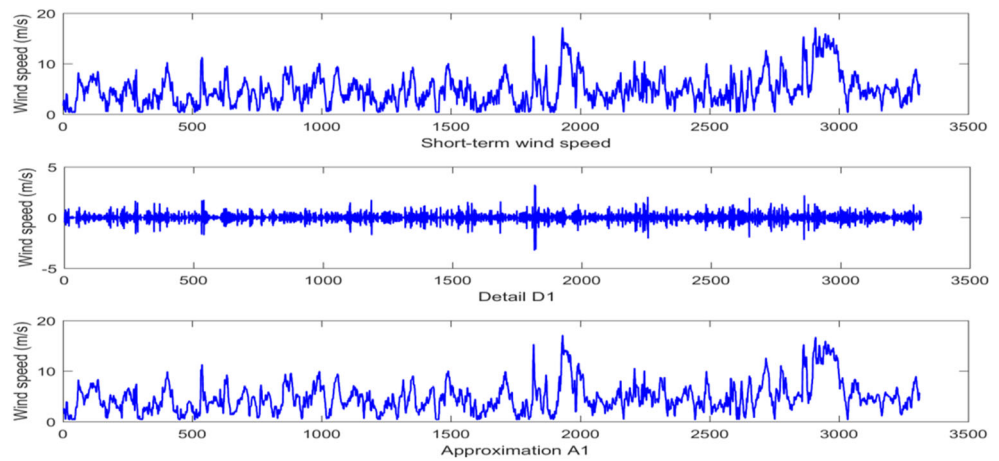


Fig. 6 The decomposition result of WT



obtained by FEEMD. For example, τ and m of the Imf_1 sequence are 3 and 14, respectively. Then the input and output matrices of the Imf_1 sequence after phase space reconstruction are shown in Eqs. (20) and (21), respectively. In addition, Table 1 also demonstrates the partitioning information of training set and test set for each subsequence.

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{3273} \end{bmatrix} = \begin{bmatrix} x_1 & x_4 & \cdots & x_{40} \\ x_2 & x_5 & \cdots & x_{41} \\ \vdots & \vdots & \cdots & \vdots \\ x_{3273} & x_{3276} & \cdots & x_{3312} \end{bmatrix} \quad (20)$$

$$\begin{bmatrix} x_{41} \\ x_{42} \\ \vdots \\ x_{3313} \end{bmatrix} \quad (21)$$

Short-term wind speed forecasting

In this section, a detailed description will be given for the short-term wind speed forecasting. In this paper, the BP neural network model optimized by BA is utilized to forecast the wind speed in the next 3 days, and the predicting interval is 10 min. The performance environment is Windows 7 system and the procedures are all implemented in MATLAB R2015b. Based on the evaluation criteria, an assessment about the proposed model will be given. Besides, a few comparison models

are also implemented to illustrate the superiority and the efficiency of the proposed hybrid model.

Model performance evaluation

In order to evaluate the forecasting performance, multiple statistical indices are chosen between the observed wind speed and the predicted wind speed in time series to evaluate the forecasting capability of the developed model comprehensively. The chosen indices include the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE). These three criteria are well-known for measuring of the deviation between the actual and forecasting values. The smaller the indices denotes the better the forecasting performance. The equations are as follows:

$$MAE = \frac{1}{p} |y_i - y_i^*| \quad (22)$$

$$MAPE = \frac{1}{p} \sum_{i=1}^p \left| \frac{y_i - y_i^*}{y_i} \right| \times 100\% \quad (23)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^p (y_i - y_i^*)^2}{p}} \quad (24)$$

where p is the number of the test samples, y_i ($i = 1, 2, \dots, p$) denotes the i -th actual data, and y_i^* is the corresponding forecasting value.

In addition, to compare the forecasting performance of the two different models, three enhanced percentage indicators

Table 1 The parameters of the forecasting subsequences obtained from FEEMD

	Imf_1	Imf_2	Imf_3	Imf_4	Imf_5	Imf_6	Imf_7	Imf_8	R_0
τ	3	5	8	4	10	13	17	15	20
m	14	13	13	7	11	6	8	6	5
Train	$X_1 \sim X_{2840}$	$X_1 \sim X_{2818}$	$X_1 \sim X_{2883}$	$X_1 \sim X_{2855}$	$X_1 \sim X_{2778}$	$X_1 \sim X_{2813}$	$X_1 \sim X_{2759}$	$X_1 \sim X_{2803}$	$X_1 \sim X_{2799}$
Test	$X_{2841} \sim X_{3272}$	$X_{2819} \sim X_{3250}$	$X_{2884} \sim X_{3315}$	$X_{2856} \sim X_{3287}$	$X_{2779} \sim X_{3210}$	$X_{2814} \sim X_{3245}$	$X_{2760} \sim X_{3191}$	$X_{2804} \sim X_{3235}$	$X_{2800} \sim X_{3231}$

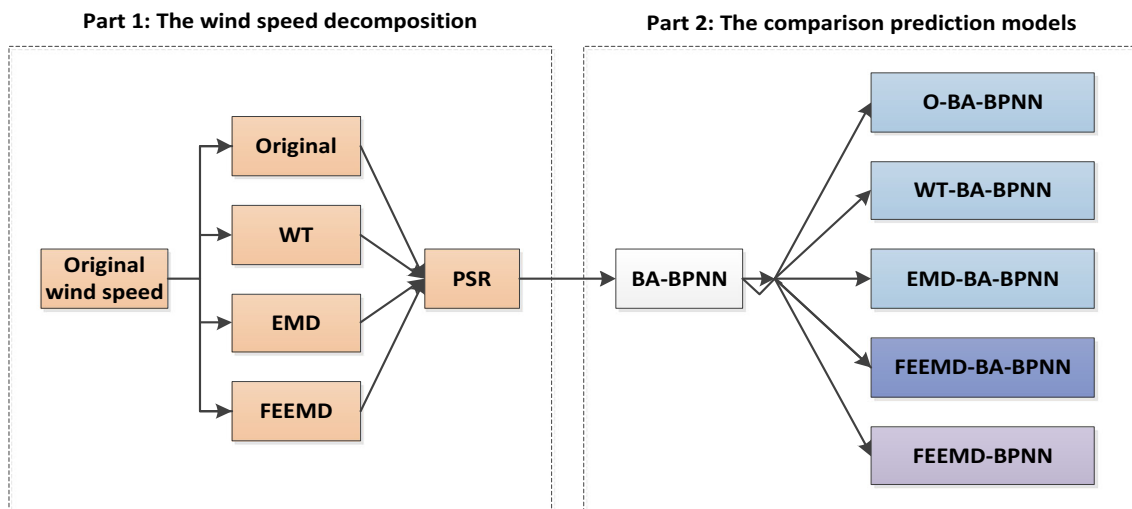


Fig. 7 The structure of the comparison models

P_{MAE} , P_{MAPE} , and P_{RMSE} are employed for describing the model superiority degree. The specific expressions are defined as follows:

$$P_{MAE} = \frac{MAE_1 - MAE_2}{MAE_1} \times 100\% \quad (25)$$

$$P_{MAPE} = \frac{MAPE_1 - MAPE_2}{MAPE_1} \times 100\% \quad (26)$$

$$P_{RMSE} = \frac{RMSE_1 - RMSE_2}{RMSE_1} \times 100\%. \quad (27)$$

Among them, subscript 1 denotes the statistical index value of the reference model, and subscript 2 denotes the statistical index value of the comparative model.

The structure of comparison models

Considering the characteristic of the proposed model, the paper used five different forecasting models compared and analyzed which are illustrated in Fig. 7. Part 1 demonstrates the

decomposition methods of WT, EMD, and FEEMD. This part shows the advantage of FEEMD method compared with the other two. The comparison models are showed in Part 2, where the first four models are the BA-based but with different data preprocessing methods. And the last two forecasting results are computed using different algorithms but with the same data preprocessing procedure.

Short-term wind speed forecasting

The fitting curves between the actual wind speed and the different forecasting from different models are given in time series in Fig. 8. To evaluate the performance of the data decomposition algorithm proposed in this paper, a study is conducted by comparing the same forecasting models with different decomposition methods and the original wind speed data. To illustrate the effectiveness of the FEEMD-BA-BPNN model, three evaluation indicators as well as the three pairwise comparative indices are calculated to evaluate the models. Table 2 shows the results of MAE, MAPE, and RMSE of all

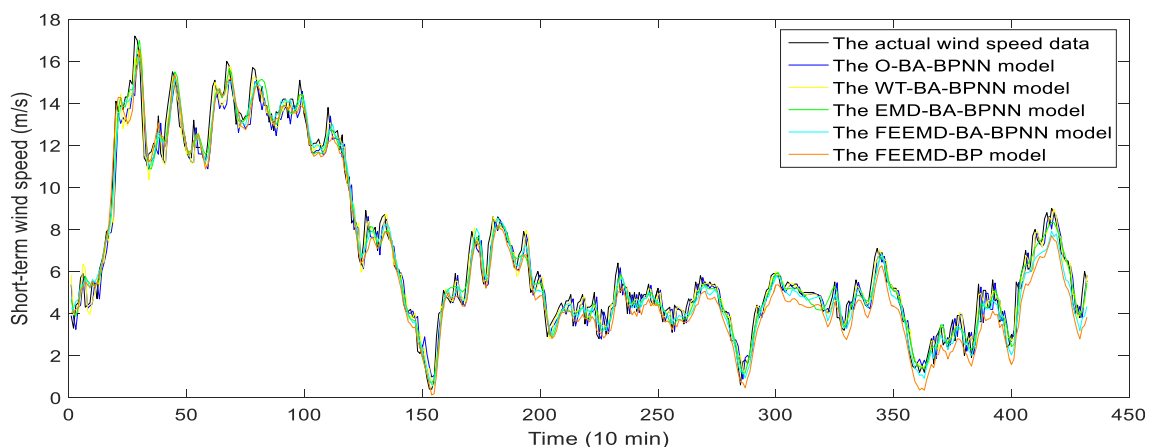


Fig. 8 The forecasting results of 7 models

Table 2 The evaluation indicators of parallel models

Index	O-BA-BPNN	WT-BA-BPNN	EMD-BA-BPNN	FEEMD-BA-BPNN	FEEMD-BPNN
MAE (m/s)	0.60008	0.48756	0.47854	0.46667	0.65475
MAPE (%)	14.88915	10.34489	9.99891	9.08783	11.88407
RMSE (m/s)	0.82345	0.65210	0.62237	0.64332	0.90005

the models. Meanwhile, the other three enhanced percentage indicators P_{MAE} , P_{MAPE} , and P_{RMSE} between two diverse forecasting models are given in Table 3.

The following conclusions can be drawn from Fig. 8 and Table 2:

- a) The proposed model FEEMD-BA-BPNN presents better fitting degree to the actual wind speed data than other methods shown in the picture. From the MAE index, the performance ranking of the comparison model from the best to the worst is FEEMD-BA-BPNN, EMD-BA-BPNN, WT-BA-BPNN, O-BA-BPNN, and FEEMD-BP. According to the MAPE index, the ranking is FEEMD-BA-BPNN, EMD-BA-BPNN, WT-BA-BPNN, FEEMD-BP, and O-BA-BPNN. As for the RMSE index, the model of FEEMD-BA-BPNN is slightly better than WT-BA-BPNN, but a little worse than EMD-BA-BPNN.
- b) The model of FEEMD-BA-BPNN has a better performance than FEEMD-BPNN, which demonstrates the significance of the optimization algorithm. The MAE, MAPE, and RMSE results of FEEMD-BA-BPNN model are 0.46667 m/s, 9.08783%, and 0.64332 m/s while the MAE, MAPE, and RMSE results of FEEMD-BPNN model results are 0.65475 m/s, 11.88407%, and 0.90005 m/s, respectively. The difference between the

three indices of two models demonstrates the improvement made by the BA.

- c) The model without any noise-canceling process delivers the worst performance. The single BA-BPNN without decomposed method shows an apparent shortcoming when compared with the FEEMD-BA-BPNN, EMD-BA-BPNN, and the WT-BA-BPNN.

Table 3 summarizes the pairwise comparison results as a percentage from the aspect of forecasting models and the decomposition methods, respectively. The value in table presents the ratio defined as row over column, where the positive value indicates that the model in row has a better performance than the model in column, and vice versa. For instance, the values of P_{MAE} in Table 4 tell that the FEEMD-BA-BPNN shows better performance in comparison with the other decomposition-based model at a percentage of 28.588, 4.476, and 2.544%, respectively. The values of P_{MAPE} suggest that the FEEMD-BA-BPNN shows better performance in comparison with the other decomposition-based model at a percentage of 63.836, 13.832, and 10.025%, respectively. The values of P_{RMSE} demonstrate that the FEEMD-BA-BPNN shows better performance in comparison with the O-BA-BPNN and the WT-BA-BPNN at a percentage of 28.000 and 1.365%, respectively. In summary, it can be observed that:

Table 3 The improved percentage indicators of proposed models

Index	O-BA-BPNN	WT-BA-BPNN	EMD-BA-BPNN	FEEMD-BA-BPNN
$P_{MAE}(\%)$				
WT-BA-BPNN	23.078			
EMD-BA-BPNN	25.398	1.885		
FEEMD-BA-BPNN	28.588	4.476	2.544	
FEEMD-BPNN	-8.350	-25.535	-26.913	-28.725
$P_{MAPE}(\%)$				
WT-BA-BPNN	43.928			
EMD-BA-BPNN	48.908	3.460		
FEEMD-BA-BPNN	63.836	13.832	10.025	
FEEMD-BPNN	25.287	-12.952	-15.863	-23.529
$P_{RMSE}(\%)$				
WT-BA-BPNN	26.277			
EMD-BA-BPNN	32.309	4.777		
FEEMD-BA-BPNN	28.000	1.365	-3.257	
FEEMD-BPNN	-8.511	-27.548	-30.852	-28.524

(a) the gap between O-BA-BPNN and the models with WT, EMD, or FEEMD reveal the usefulness of the signal decomposition methods. (b) The decomposition scale level of WT is usually determined by experience, which is considered as one of the reasons for poor performance when compared with EMD. (c) The errors decreasing between EMD-BA-BPNN and FEEMD-BA-BPNN show the usefulness of the fast ensemble empirical mode decomposition. By adding random white noise to FEEMD, the mode mixing phenomenon which may occur in EMD is mitigated.

Conclusions

An accurate wind speed forecasting method is not only important for the wind power large-scale integration but also relates to the security and stability of the state grid. So, in this work, a novel compound short-term wind speed forecasting model is proposed. The proposed model employs fast ensemble empirical mode decomposition, phase space reconstruction, and back propagation neural network optimized by bat algorithm. The FEEMD is used for filtering the noises out from the raw wind speed data. Considering the chaotic characteristic of wind speed in time series, phase space reconstruction is applied to determine the input vector and corresponding output for each series. And the back propagation neural network with bat algorithm optimized parameters is conducted to do the short-term forecasting for the wind speed.

To validate the effectiveness and usefulness of the proposed model, some comparison forecasting models are also implemented in this work. The comparison models include O-BA-BPNN, WT-BA-BPNN, EMD-BA-BPNN, FEEMD-BA-BPNN, and FEEMD-BPNN. From the comparison results, it can be concluded as: (a) signal decomposition methods are useful in short-term wind speed forecasting works. (b) The proposed method which combined the fast ensemble empirical mode decomposition and phase space reconstruction together delivered solid performance for short-term wind speed forecasting. FEEMD reduces the volatility and randomness of the signal significantly while the PSR address the chaotic property of wind speed in time series which makes the proposed method an innovative one. (c) The back propagation neural network with bat algorithm optimized parameters presented good performance in this work because of its strong generalization and nonlinear mapping capability.

References

- Ali M, Khan A, Rehman NU (2018) Hybrid multiscale wind speed forecasting based on variational mode decomposition. *Int Trans Electr Energy Syst*:28(1)
- Cao L (1997) Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D* 110:43–50
- Erdem E, Shi J (2011) ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl Energy* 88:1405–1414
- Fraser AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. *Phys Rev A* 33(2):1134–1140
- Gao S, Dong L, Liao XZ, Gao Y (2013) Very-short-term prediction of wind speed based on chaos phase space reconstruction and NWP. In: 2013 32nd Chinese Control Conference (CCC)
- Guo Z et al (2011) A case study on a hybrid wind speed forecasting method using BP neural network. *Knowledge Based Syst* 24(7):1048–1056
- Han YJ, Liu J (2015) The online forecasting research of short-term wind speed and power generation at wind farm based on phase space reconstruction. In: 2015 Seventh International Conference On Measuring Technology And Mechatronics Automation (Icmtma 2015)
- Hoolohan V, Tomlin AS, Cockerill T (2018) Improved near surface wind speed predictions using Gaussian process regression combined with numerical weather predictions and observed meteorological data. *Renew Energy* 126:1043–1054
- Jiang P, Ma XJ (2016) A hybrid forecasting approach applied in the electrical power system based on data preprocessing, optimization and artificial intelligence algorithms. *Appl Math Model* 40:10631–10649
- Jiang P, Li R, Zhang K (2018) Two combined forecasting models based on singular spectrum analysis and intelligent optimized algorithm for short-term wind speed. *Neural Comput Applic* 30(1):1–19
- Lazarevska E. Wind speed prediction with extreme learning machine. In: 2016 IEEE 8th international conference on intelligent systems (is).
- Liu D, Niu DX, Wang H, Fan LL (2014) Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renew Energy* 62:592–597
- Liu H, Tian HQ, Liang XF, Li YF (2015) New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, mind evolutionary algorithm and artificial neural networks. *Renew Energy* 83:1066–1075
- Liu JQ, Wang XR, Lu Y (2017) A novel hybrid methodology for short-term wind power forecasting based on adaptive neuro-fuzzy inference system. *Renew Energy* 103:620–629
- Liu L, Ji TY, Li MS, Chen ZM, Wu QH (2018) Short-term local prediction of wind speed and wind power based on singular spectrum analysis and locality-sensitive hashing. *J Mod Power Syst Clean Energy* 6:317–329
- Lydia M, Kumar SS, Selvakumar AI, Kumar GEP (2016) Linear and non-linear autoregressive models for short-term wind speed forecasting. *Energy Convers Manag* 112:115–124
- Ma DL (2013) Determination of time delay τ in phase space reconstruction. *Silicon valley* 19: 51+48
- Mirjalili S, Mirjalili SM, Yang X (2014) Binary bat algorithm. *Neural Comput Applic* 25(3–4):663–681
- Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond Ser Math Phys Eng, Sci*:903–995
- Santamaria-Bonfil G, Reyes-Ballesteros A, Gershenson C (2016) Wind speed forecasting for wind farms: a method based on support vector regression. *Renew Energy* 85:790–809
- Santhosh M, Venkaiah C, Kumar DMV (2018) Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction. *Energy Convers Manag* 168:482–493
- Su Z et al (2014) A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting. *Energy Convers Manag* 85:443–452
- Sun W, Wang Y (2018) Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space

- reconstruction, sample entropy and improved back-propagation neural network. *Energy Convers Manag* 157:1–12
- Tavakkoli A, Rezaeenour J, Hadavandi E (2015) A novel forecasting model based on support vector regression and bat meta-heuristic (Bat-SVR): case study in printed circuit board industry. *Int J Inform Technol Decis Mak* 14(1):195–215
- Wang YH, Yeh CH, Young HW, Hu K, Lo MT (2014) On the computational complexity of the empirical mode decomposition algorithm. *Physica A* 400:159–167
- Wu Z, Huang NE (2009) Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Adv Adapt Data Anal* 1(1): 1–41
- Wu Q, Peng C (2016) Wind power generation forecasting using least squares support vector machine combined with ensemble empirical mode decomposition, principal component analysis and a bat algorithm. *Energies* 9(4)
- Xiao L, Qian F, Shao W (2017) Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. *Energy Convers Manag* 143:410–430
- Yang X, Gandomi AH (2012) Bat algorithm: a novel approach for global engineering optimization. *Eng Comput* 29(5–6):464–483
- Yang XS (2010) A new metaheuristic bat-inspired algorithm. *Comput Knowl Technol* 284:65–74
- Yu M, Zhou WM, Wang B, Jin J. The short-term forecasting of wind speed based on emd and arma. In: proceedings of the 2017 12th IEEE conference on industrial electronics and applications (iciea)
- Zhang C, Wei HK, Zhao X, Liu TH, Zhang KJ (2016) A Gaussian process regression based hybrid approach for short-term wind speed prediction. *Energy Convers Manag* 126:1084–1092
- Zhang YG, Zhang CH, Zhao Y, Gao S (2018) Wind speed prediction with RBF neural network based on PCA and ICA. *J Electr Eng* 69:148–155
- Zhou J et al (2018) A novel decomposition-optimization model for short-term wind speed forecasting. *Energies* 11(7)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.