IOP Conference Series: Earth and Environmental Science

A Hybrid Approach for Short-term Wind Speed Prediction in Huan County of China

IOP Conf. Series: Earth and Environmental Science **295** (2019) 012030 doi:10.1088/1755-1315/295/2/012030

A Hybrid Approach for Short-term Wind Speed Prediction in Huan County of China

Tonglin Fu[1,a] , Mingxia Yang[1], Pengyue Ju[1], Kun Liu[1], Huani Zhao[1]

[1]School of Mathematics & Statistics, Longdong University, Qingyang 745000, China

[a] Corresponding author: futonglin2008@163.com

Abstract. Wind energy, which is intermittent due to the irregular and non-stationary characteristics of wind speed, can have a significant impact on power grid security. It is important to improve the accuracy of wind speed forecasting models for the wind generation. However, due to the nonlinear and intrinsic complexity of weather parameters, it is difficult to predict wind speed accurately by using different patterns in different locate. In this paper, a new hybrid wind speed forecasting model is constructed based on a back-propagation neural network(BPNN) and the idea of eliminating noise effects by using ensemble empirical mode decomposition(EEMD) method and eliminating seasonal effects from actual wind speed dataset using seasonal exponential adjustment(SEA). The hybrid EEMD-SEA-BPNN models are proposed to forecast the wind speed effectively in Huan County of Loess Plateau in China; numerical results demonstrate that the hybrid EEMD-SEA-BPNN model has better forecasting performance.

1. Introduction

Wind energy, which is generated by atmospheric motions, is one of the predominant alternative sources of new renewable energy resource with its pollution-free and reproducibility. According to The Global Wind Power Casting Industry 2016 Market Research Report [1], the global wind power industry installed another 54.6GW and the global cumulative installed capacity of wind power reached 486.7GW with an increase of 12.6%.

It is important foundation and premise of wind power and wind farm generation prediction to improve the accuracy of wind speed forecasting model for wind park management. However, due to the disadvantages of intermittent and fluctuating, as well as low energy density, it is difficult to predict wind speed accurately by using different patterns in different locate. There are many statistical models and machine learning models have been developed to predict the wind speed. Statistical models primarily use a time series approach and have been successfully applied for forecasting based on the assumption that a linear correlation structure exists among time series values. [2-5]. Therefore, non-linear patterns cannot be captured using these statistical models. To overcome these shortcomings, neural network models such as BPNN, support vector machines (SVM) and radial basis function (RBF) neural network have been used to improve the accuracy of wind speed predictions [6-13].

In general, wind speed forecasting models can be classified as short-term predictions and long-term predictions according to the different time scale. For short-term predictions, it is important in minimizing scheduling errors due to the wind speed characteristics of the nonlinear and intrinsic complexity which will have an impact on the power grid reliability and add additional costs in market-based ancillary service [14]. Riahy et al. [15] constructed a linear short-term wind speed forecasting model based on the observation that filtering out less effective frequency components from a wind speed spectrum can increase the correlation between real and predicted wind speed. Sancho et al.[16] presented a method of exploiting the diversity in input data using banks of neural networks model for short term wind speed forecasting, which have better forecasting performance compared with single neural networks. For long-term wind speed predictions, precise forecasting is important for performance prediction, site selection and selection of the optimal wind machine size for a particular site et al. [17]. Jujie Wang et al. [18] presented a hybrid SAM-ESM-RBFN model based on the seasonal adjustment method (SAM), exponential smoothing method (ESM), and radial basis function neural network (RBFN) to predict the mean hourly wind speed of two meteorological stations in the Hexi corridor of China and obtained better prediction performance. Zhenhai Guo et al. [19] constructed a new hybrid SEA-ARMA models and SEA-GARCH to forecaste the long-term wind speed in the Hexi corridor of China, the numerical results show that the developed models have higher accuracy than the single time series model such as ARMA and GARCH.

In this paper, the technique of time series analysis combined with an artificial neural network is adopted to deal with the chaotic and intrinsic complexity of the wind speed time series. The focus of this study falls into the short-term predictions and the predictions of the daily average wind speed one year ahead. In order to evaluate the daily average wind speed in Huan County of China, the seasonal exponential adjustment (SEA) method is used to eliminate seasonal effects and the ensemble empirical mode decomposition (EEMD) method is used to eliminate noise effects from actual wind speed datasets. The hybrid EEMD-SEA-BPNN models are presented to forecast the wind speed in Huan County of Loess plateau in China; numerical results show that the hybrid EEMD-SEA-BPNN models have better forecasting performance compared with forecasting models such as BPNN, SEA-BPNN and EEMD-BPNN.

2. Theoretical model and calculation method

2.1 Ensemble empirical mode decomposition (EEMD)

In reality, the wind speed are influenced by a variety of factors and the noisies are mixtured with the wind speeds, forecasting the wind speed time series directly often has large error. In order to overcome this disadvantages, the data pre-process technique EEMD is used to counterbalance this weakness .

EEMD, which is extended from EMD to overcome the drawback of mode mixing, is an effective self-adaptive data processing method to decompose the non-linear and non-stationary time series[25]. The key of EMD is to decompose complex signals into finite intrinsic mode function (IMF) which satisfying the following two conditions:1) the number of maxima and minima and number of zero crossings must be equal, or different at the most by one in each whole function, and 2) the functions must be symmetric with respect to local zero mean. The decomposed IMF components contain local characteristic signals of different time scales of original signals. The EMD decomposition process is as follows:

1) Find out all the upper and lower envelope of the original time series $x(t)$ through the cubic spline function.

2) The first IMF is calculated as

$$h_1(t) = x(t) - m_1(t) \qquad (1)$$

where $m_1(t)$ denote the mean value of upper and lower envelope. It is difficult to obtain the IMF by one decomposition because wind speed time series consists of linear and nonlinear patterns and there are still some asymmetric waves exist in $h_1(t)$, therefore, repeat the above process k times for $h_1(t)$ until $h_1(t)$ satisfy the conditions of IMF, that is

$$h_k(t) = h_{k-1}(t) - m_k(t) . \qquad (2)$$

The first component of IMF denoted as $f_1(t) = h_k(t)$.

3) The residual signal is denoted as $r_1(t)$, where

$$r_1(t) = x(t) - f_1(t). \qquad (3)$$

4) The second component of high frequency $f_2(t)$ can be obtained by regarding $r_1(t)$ as the original time series and repeat the above process. Repeat above steps n times, there are n IMFs are obtained, that is

$$\begin{cases} r_2(t) = r_1(t) - f_2(t), \\ r_3(t) = r_2(t) - f_3(t), \\ \quad \vdots \\ r_n(t) = r_{n-1}(t) - f_n(t). \end{cases} \qquad (4)$$

The decomposition process stopped when the termination condition of the residual component is a monotonic function and it cannot be decomposed as IMF.

The original wind speed time series $x(t)$ can be repressed as

$$x(t) = \sum_{i=1}^{n} f_i(t) + r_n(t), \qquad (5)$$

where $\sum_{i=1}^{n} f_i(t)$ is the average variation trend of original series and the residual component $r_n(t)$ regarding as noises. Because each IMF of the EMD has different frequency component, which leads to the frequency mixing phenomenon occurs and unachieved effective result. The EEMD, which is extended from EMD to overcome the drawback of frequency mixing, is widely used to decompose non-linear and non-stationary signal sequences [33,36]. It defines the true IMFs components as the mean of an ensemble of trials and each trial consists of the decomposition results of the signal plus a white noise of finite amplitude. EEMD decomposition principle is: when added white noise is uniformly distributed throughout the time-frequency space, the time-frequency space is divided into different scales by the filter group components. When the signal with uniform distribution of white noise background, the different scales of signal area will be automatically mapped to the appropriate scale of associated with the white background noise, each individual test may produce very noisy as a result, this is because each additional noise components including the white noise signal and the additional. Since noise is different in each individual test, when using enough to test all the average, the noise will be eliminated. The average of the last will be considered the real results, as more and more testing, additional noise is eliminated, the only persistent part is the signal itself.

In this paper, the following steps show the algorithm of EEMD:

Step 1: Add a white noise $w(t)$ to the wind speed time series $x(t)$ and the new series is obtained as

$$X(t) = x(t) + w(t). \qquad (6)$$

The effect of the added white noise can be controlled by $\varepsilon_{ne} = \varepsilon/\sqrt{NE}$, where ε_{ne} is the final standard deviation of error defined as the difference between the input signal and the corresponding IMFs, NE is the number of ensemble members and ε is the amplitude of the added noise,

Step 2: Decompose the time series $X(t)$ into IMFs using the EMD algorithm;

Step 3: Repeat Step 1 and Step2 with different white noise series each time.

Step 4: The final result can be obtained by computing the means of corresponding IMFs.

2.2 Seasonal exponential adjustment (SEA)

$,P_t$ In this study, T, m, l are given as integers and $\mathrm{T=m\cdot l}$. In addition, x_t denotes the wind speed time series at time $\mathrm{t} \in \{1,2,\cdots,T\}$, the seasonal and trend components are denoted as I_t and P_t respectively. The multiplicative wind speed x_t at time t can be expressed as

$$\mathrm{x}_t = \mathrm{I}_\mathrm{t} \times P_t \quad (7)$$

Thus, the seasonal index I_t can be computed as:

$$\mathrm{I}_\mathrm{t} = \mathrm{x}_t / P_t \quad (8)$$

Next, dividing the time series $\mathrm{x}_t, \mathrm{t} \in \{1, 2, \cdots, T\}$ into l groups, and each group represents one cycle, and m time series are used in each cycle because $\mathrm{T} = \mathrm{m} \cdot 1$. In addition, $\mathrm{x}_t, \mathrm{t} \in \{1, 2, \cdots, T\}$ is denoted as

$\mathrm{x}_{11}, \mathrm{x}_{12}, \cdots, \mathrm{x}_{1s}, \cdots, \mathrm{x}_{1m}, \mathrm{x}_{21}, \mathrm{x}_{22}, \cdots, \mathrm{x}_{2s}, \cdots, \mathrm{x}_{2m}, \cdots, \mathrm{x}_{k1}, \mathrm{x}_{k2}, \cdots, \mathrm{x}_{ks}, \cdots, \mathrm{x}_{km}, \cdots, \mathrm{x}_{l1}, \mathrm{x}_{l2}, \cdots, \mathrm{x}_{ls}, \cdots, \mathrm{x}_{lm}$,
$k = 1, 2, \cdots, l, s = 1, 2, \cdots, m.$

where x_{ks} represents the *s-th* datum of the k-*th* cycle. The unknown trend component can be approximated by computing the average of each cycle [20]. The average of the k-*th* cycle is obtained by the following equation :

$$\overline{\mathrm{x}} = (\mathrm{x}_{k1} + \mathrm{x}_{k2} + \cdots + \mathrm{x}_{km}) / m, k = 1, 2, \cdots, l. \quad (9)$$

Let I_{ks} denote the normalization data of items x_{ks}, then

$$I_{ks} = x_{ks} / \overline{x}_k,\ k = 1, 2, \quad , l, s = 1, 2, \quad , m, \quad (10)$$

and I_s is denoted as

$$I_s = (I_{1s} + I_{2s} + \cdots + I_{ls}) / l, s = 1, 2, \cdots, m. \quad (11)$$

The definition of I_s indicates the normalization process and can be expressed as

$$\sum_{s=1}^{l} I_s = \frac{1}{l} \sum_{k=1}^{l} \sum_{s=1}^{m} I_{ks} = \frac{1}{l} \sum_{k=1}^{l} \left(\sum_{s=1}^{m} \mathrm{x}_{ks} / \overline{x}_k \right) = \frac{1}{l} \sum_{k=1}^{l} l = l. \quad (12)$$

From above process, the time series without infect of seasonal component is computed as

$$x'_{ks} = x_{ks} / I_s, k = 1, 2, \cdots, l, s = 1, 2, \cdots, m. \quad (13)$$

The new wind speed time series without the seasonal component can be obtained if the data items $\mathrm{x}'_{11}, \mathrm{x}'_{12}, \cdots, \mathrm{x}'_{1l}; \cdots$, and $\mathrm{x}'_{m1}, \mathrm{x}'_{m2}, \cdots, \mathrm{x}'_{ml}$ are re-recorded back to $\mathrm{x}'_1, \mathrm{x}'_2, \cdots, \mathrm{x}'_T$.

For the additive decomposition model,the following equations are used to replace Eqs. (10), (11), (12), and (13) similar to the multiplicative decomposition model:

$$\mathrm{x}_t = \mathrm{I}_\mathrm{t} + P_t \quad (14)$$

$$\mathrm{I}_\mathrm{t} = \mathrm{x}_t - P_t \quad (15)$$

$$I_{ks} = x_{ks} - \overline{x}_k, k = 1, 2, \cdots, l, s = 1, 2, \cdots, m. \quad (16)$$

$$x'_{ks} = x_{ks} - I_s, k = 1, 2, \cdots, l, s = 1, 2, \cdots, m. \quad (17)$$

In this paper, the cycle length of the wind speed time series $\mathrm{m} = 30(31)$.

2.3 Back-Propagation neural network(BPNN)

The structure of a BPNN is composed by the input layer, the hidden layer and the output layer[18]and the goal of training process is minimizing the global mean sum squared error E between the output z_l of real network and the desired output t_l, which denoted as

$$E = \frac{1}{2} \sum_l (t_l - z_l)^2 \quad (18)$$

where

$$z_l = f(\sum_j v_{lj} \varphi(\sum_j w_{ji} x_i - \theta_j) - \theta_l), \quad (19)$$

w_{ji} denote the weight of the connection from input i to neuron j and v_{lj} denote the weight of the connection from neuron j to output l, x_i is input time series, θ_l, θ_j denote the threshold and $f(\bullet), \varphi(\bullet)$ denote activation function. The inner product of the input vector and weight vector by a nonlinear transfer function is calculated to to get a scalar result for each node in the network.

To achieve better forecasting performance, the node number of the hidden layer is determined by using the Hecht-Nelson algorithm [20]: if the node number of the input layer is n, then the number of the hidden layer is 2n+1. In this paper, The structure of BPNN is composed with one output neuron, $2n+1$ hidden neurons and n input neurons. To ensure the accuracy of predicted results, the input data is normalized in advance, the compute formula is

$$\mathrm{x}' = \{x_i'\} = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}}, i = 1, 2, \quad , n. \quad (20)$$

In this paper, a three-layer BPNN is selected to forecast the daily average wind speed.

2.4 The hybrid EEMD-SEA-BPNN model

In reality, the wind speed time series are non-linar, irregular and highly-noisy due to the nonlinear and intrinsic complexity of weather parameters, forecasting wind speed with the original time series directly will lead to lower accuracy , this disadvantage can be overcomed by using the data preprocess technical EEMD.

According to Zhang [21],if the time series are characterized by increasing seasonal variations, the multiplicative decomposition method is appropriate to eliminte the seasonal and trend components. However, if the seasonal variation is relatively consistent with the trend, the additive decomposition model should be used. In reality, it is very difficult to determine whether the addition or multiplication operations model is more suitable [22]. Thus, both multiplicative and additive decomposition models [23] are used to decompose the seasonal and trend components in wind speed series.

BPNN was proposed by Rumelhart and McCelland in 1986. It is a multilayered feedforward network trained according to the error inverse propagation algorithm and is one of the most widely used neural network models at present. BPNN is one of the effectively models in uncovering nonlinearity between the input layer and the output layer without sufficient information, and BPNN is widely used in back analysis due to the fact that it can extract useful information from training processes without prior assumptions regarding the form of functions related to input and output layers. It can approximate an arbitrary nonlinear function with satisfactory performance.

The EEMD-SEA-BPNN algorithm composed as following steps :

Firstly, the wind speed time series are decomposed into several layers by using the data preprocess EEMD ;

Secondly, the SEA method (multiplicative and additive decomposition) are used to eliminate seasonal effects from actual wind speed data ;

Thirdly, BPNN model is used to forecast the signal wtihout seasonal effects;

Finally, the end forecasting results are obtained by combining the seasonal effects.

3 The numerical results and discussion

3.1 Study area and evaluation criteria

Huan County of China, which is the study area, lies in the northeast of Gansu province. It is a mountainous area bordering on Shanxi province, Gansu province and the Ningxia Hui Autonomous Region. There are 20 towns with a total land area of 9236 square kilometers and a population of 347000 people, of which agricultural population accounts for 93.7%. With an altitude of 1200-2089 meters, an average annual rainfall of about 300 mm, its typical climate is cool and droughty. Huan County is classified as one of the key counties of the national poverty alleviation plan because of its special geographical environment and backward economic development. In June 2011, The total investment is 4.015 billion yuan projects are started to construct the million kilowatt wind farmer in Huan County, the total installed capacity is about 100MW. In October 2012, the Tianshui wind farmer has generated more than 1000 kwh.

In this paper, the daily average wind speed time series data set was collected from 1 January 2012 to 31 December 2018 from the Huan County in China. The daily average wind speed in 2018 are forecasted by using the data set in the corresponding months from 2012 to 2017 based on BPNN and the idea of eliminating noise effects by using EEMD from actual wind speed datasets and eliminating seasonal effects from actual wind speed data set by using SEA.

The M-W test and K-S test are two effective non-parameter test method to compare the difference between the distributions of two continuous random variables. Suppose that the daily average wind speed in 2018 of the Huan County are forecasted by using the wind speed time series from 2012 to 2017. First, the two sample M-W test and K-S test are used to determine whether the distributions between the paired samples were significantly different. In the experiments, the two effective non-parameter M-W test and K-S test tool in the SPSS software were used to compare the differences between the distributions of the paired samples (years 2012 and 2013, years 2013 and 2014, years 2014 and 2015, years 2015 and 2016, years 2016 and 2017, and years 2017 and 2018). The results of each paired samples and the probabilities are shown in Table 1 and Table 2. The conclusion can be obtained from Table 1 and Table 2 that the differences among the paired samples are not significant since the p-values are all larger than the significance level of 0.05, which means that the hybrid EEMD-SEA-BPNN model of using the wind speed of January to December from 2012 to 2017 to forecast the wind speed in the corresponding months of 2018 is resonable.

Table 1: The M-W test results of daily average wind speed from 2012 -2017

Year	2012 -2013	2013 -2014	2014 -2015	2015 -2016	2016 -2017
p-values	0.313	0.463	0.494	0.905	0.287
p-values	0.795	0.510	0.390	0.496	0.773

Table 2:The K-S test results of daily average wind speed from 2012 - 2017

Year		2012 -2013	2013 -2014	2014 -2015	2015 -2016	2016 -2017
Most Extreme Differences	Absolute	0.194	0.129	0.129	0.097	0.194
	Positive	0.065	0.032	0.129	0.065	0.194
	Negative	-0.194	-0.129	-0.032	-0.097	-0.032
K-S		0.762	0.508	0.508	0.381	0.762
p-values		0.607	0.959	0.959	0.999	0.607
Most Extreme Differences	Absolute	0.133	0.267	0.167	0.200	0.100
	Positive	0.133	0.267	0.067	0.033	0.100
	Negative	-0.067	-0.100	-0.167	-0.200	-0.033
K-S		0.516	1.033	0.645	0.775	0.387
p-values		0.952	0.236	0.799	0.586	0.998

In this paper,order to evaluate the forecasting performance of the hybrid EEMD-SEA-BPNN model, the mean square error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are adopted and these measures are defined as following:

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(x_i\text{-}\hat{x}_i)^2, \quad (21)$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|x_i\text{-}\hat{x}_i\right|, \quad (22)$$

$$\mathrm{MAPE}=\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left|\frac{x_i\text{-}\hat{x}_i}{x_i}\right|\times 100\%, \quad (23)$$

where x_i and $\hat{x}_i$ represent the *i-th* original and predicted values, respectively.

3.2 Experimentation design and numerical results

The wind speed data set of the Huan County in the Loess Plateau was predicted by the proposed hybrid EEMD-SEA-BPNN model. The model superiority can be reflected by comparing with other popular single forecasting models recommended by recent works. Among those time series forecasting models, the single ARIMA model, BPNN model and other simple hybrid models such as EEMD-BPNN, SEA-BPNN, SEA-ARIMA, EEMD-ARIMA, EEMD-SEA-ARIMA are adopted as the benchmarks.

The simulation process is as following：

Step 1: Perform the M-W test and K-S test for the data set to judge whether the wind speed data sets (training set and test set) have significant difference each other or not.

Step 2: The EEMD technique is employed to eliminate noise of the wind speed time series.

Step 3: Two models, multiplicative and additive decomposition models are used to eliminate seasonal effects from data after noise elimination.

Step 4: The hybrid forecasting models based on the data pre-processing techniques including EEMD-SEA-BPNN, EEMD-BPNN, SEA-BPNN, SEA-ARIMA, EEMD-ARIMA, EEMD-SEA-ARIMA, and two single models BPNN and ARIMA are used to predict the daily average wind speed.

Step 5: The forecasting performance of the above models are compared in different benchmarks such as MSE, MAE and MAPE.

The numerical results of MSE, MAE and MAPE are calculated and shown in Table 3 and Table 4; Fig. 1 and Fig. 2 shows the forecasting results of the single ARIMA model, BPNN model and other hybrid EEMD-BPNN, SEA-BPNN, SEA-ARIMA, EEMD-ARIMA, EEMD-SEA-ARIMA models and the original daily average wind speed.

3.3 Comparison and discussion

As shown in Table 3 and Table 4, Fig. 1 and Fig. 2 , the ARIMA, BPNN, EEMD-BPNN, EEMD-ARIMA, SEA-BPNN, SEA-ARIMA, EEMD-SEA-ARIMA and EEMD-SEA-BPNN models all have good predictive effects, it is clear that the hybrid EEMD-SEA-BPNN model performs much better than the hybrid model EEMD-BPNN, SEA-BPNN, SEA-ARIMA, EEMD-SEA-ARIMA, EEMD-ARIMA and two single models BPNN and ARIMA model, both multiplicative and additive decomposition models are used to eliminate seasonal effects from data. The MSE, MAE and MAPE are all smaller than the MSE, MAE and MAPE of the hybrid model EEMD-BPNN, SEA-BPNN, EEMD-ARIMA, SEA-ARIMA, EEMD-SEA-ARIMA and two single models BPNN and ARIMA model. The MAPE of the hybrid EEMD-SEA-BPNN are 21.24% and 22.34% respectively, more precisely, compared with ARIMA model, the MAPE of proposed model leads to reductions of 5.78% and 10.04%, which means that the the hybrid EEMD-SEA-BPNN model has better forecasting performance.

The conclusion can be obtained from Table 3 and Table 4 that the forecast accuracy are improved by using multiplicative and additive decomposition to eliminate seasonal effects. The data preprocessing technical of EEMD can improve data quality effectively and forecasting performance.

The data pre-processing technology EEMD is an efficient algorithm in improving the forecasting performance. In general, the BPNN model has higher performance than the time series model ARIMA in the daily average wind speed predicting processes.

In conclusion, the novel hybrid wind speed forecasting model EEMD-SEA-BPNN, which based on data pre-processing method is adopted to predict the wind speed effectively in the Huan County of Loess Plateau in China, the simulation results show that the hybrid EEMD-SEA-BPNN model has higher robustness than the other models.

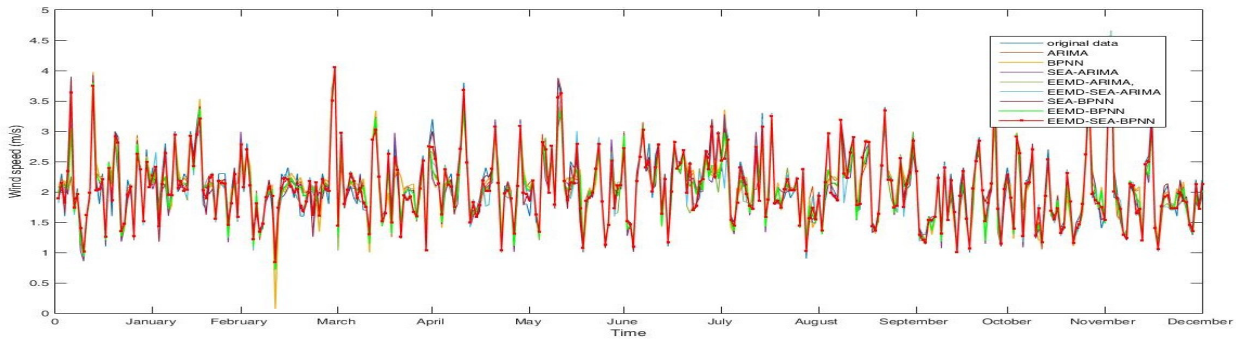

Fig. 1 The forecasting performance of EEMD-SEA-BPNN, EEMD-BPNN, SEA-BPNN, EEMD-SEA-ARIMA, EEMD-ARIMA,
SEA-ARIMA ,BPNN and ARIMA for the multiplicative decomposition model

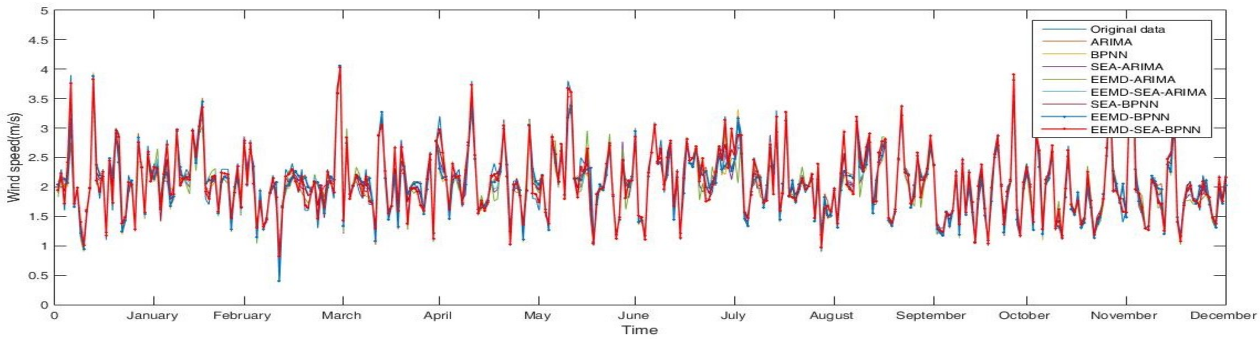

Fig. 2 The forecasting performance of EEMD-SEA-BPNN, EEMD-BPNN, SEA-BPNN, EEMD-SEA-ARIMA, EEMD-ARIMA,
SEA-ARIMA ,BPNN and ARIMA for the additive decomposition model

Table 3. The MSE，MAE and MAPE of EEMD-SEA-BPNN, EEMD-BPNN,SEA-BPNN, EEMD-SEA-ARIMA, EEMD-ARIMA,
SEA-ARIMA ,BPNN and ARIMA for the multiplicative decomposition model

Evaluation criteria	EEMD-SEA-BPNN	EEMD-BPNN	SEA-BPNN	EEMD-SEA-ARIMA	EEMD-ARIMA	SEA-ARIMA	BPNN	ARIMA
MSE(m/s)	0.2314	0.2502	0.2827	0.2678	0.2838	0.2821	0.3856	0.4011
MAE(m/s)	0.3908	0.4052	0.4123	0.4186	0.4310	0.4372	0.45354	0.5289
MAPE (%)	21.24%	22.58%	23.99%	22.95%	25.59%	26.65%	27.02%	28.16%

Table 4. The MSE，MAE and MAPE of EEMD-SEA-BPNN, EEMD-BPNN,SEA-BPNN, EEMD-SEA-ARIMA, EEMD-ARIMA,
SEA-ARIMA ,BPNN and ARIMA for the additive decomposition model

Evaluation criteria	EEMD-SEA-BPNN	EEMD-BPNN	SEA-BPNN	EEMD-SEA-ARIMA	EEMD-ARIMA	SEA-ARIMA	BPNN	ARIMA
MSE(m/s)	0.2414	0.2603	0.2828	0.2758	0.3832	0.2771	0.3956	0.4211
MAE(m/s)	0.4008	0.4083	0.4223	0.4096	0.4670	0.4378	0.5354	0.4378
MAPE (%)	22.34%	22.59%	23.89%	22.94%	26.59%	27.65%	31.02%	32.37%

4. Conclusions

It is important for wind farm management to obtain accurate wind speed forecasting results due to the uncertainty about wind power. The associated benefits can be gained from estimating power output if the bias in wind speed prediction is reduced by 10% in the electricity market [24]. In this paper, the model that combined the EEMD method and SAM with BPNN was constructed and tested with original wind speed data sets for 2018 in the Huan County of China. Numerical results show that the hybrid EEMD-SEA-BPNN model can forecast the daily average wind speed one year ahead with a better accuracy compared with forecasting models such as EEMD-BPNN, SEA-BPNN, EEMD-SEA-ARIMA, EEMD-ARIMA, SEA-ARIMA, BPNN and ARIMA. The numerical results show that the values of MAPE of the hybrid EEMD-SEA-BPNN at most are 19.23% and 22.88% for the two real wind speed datasets.

Acknowledgement

This work was supported by the Western Project of the National Social Science Foundation of China (Grant No.18XTJ003)

References

[1] https://www.researchgate.net/publication/316699909_Global_Wind_Report_2016-Annual_Market_Update.
[2] E. Erdem, J. Shi, Applied Energy, **88**, 1405–1414 (2011)
[3] R. G. Kavasseri, K. Seetharaman, Renewable Energy, **34**, 1388–1393(2009)
[4] N. Bokde, G.Asencio-Cortés, F. Martínez-Álvarez, K. Kulat, R J., **9**, 324–333(2017)
[5] A. Gupta, N. Bokde, K. Kulat, Water Resour. Manag, **32**,1133–1151(2018) .
[6] A. Sfetsos, Renewable Energy, **27**, 163–174 (2002)
[7] K. W. Chau, Journal of Hydrology, 329,363-367 (2006)
[8] C. T. Cheng, C. P. Ou, K. W. Chau, Journal of Hydrology, **268**, 72–86 (2002)
[9] P. Flores, A. Tapia, G. Tapia, Renewable Energy, **30** 523-536(2005)
[10] Z. X. Guo, W. K. Wong, Information Sciences,**193**, 54–71(2012)
[11] C. Hamzace, Information Sciences,**178**,4550-4559 (2008)
[12] K. P. Lin, P. F. Pai, Y. M. Lu, P. T. Chang, Information Sciences, **220**,196–209(2013)
[13] J. J. Wang, J. Z. Wang, Z. G. Zhang, S. P. Guo, Omega-International Journal of Management Science, **40**,758–766(2012)
[14] R. G. Kavasseri, K. Seetharaman, Renewable Energy, **34**,1388–1393(2009)
[15] G. H. Riahy, M. Abedi, Renewable Energy, **33**,35–41(2008)
[16] S. S. Sancho, M. P. B. Angel, G. O. G. Emilio, P. F. Antonio, P. Luis, C. Francisco, Neurocomputing, **72**,1336–1341(2009)
[17] Jianzhou Wang, Yiliao Song, Feng Liu, Ru Hou,Renewable and Sustainable Energy Reviews, **60**,960–981(2016)
[18] L. Xiao, W. Shao, M. Yu, J. Ma, Appl. Energy, **198**,203–222(2017)
[19] Zhenhai Guo, Yao Dong, Jianzhou Wang, Haiyan Lu, Mathematical Problems in Engineering, **10**,684742 (2010)
[20] T. L. Fu, C. Wang, Sustainability,**10**, 3913(2018)
[21] G. P. Zhang, M. Qi, European Journal of Operational Research, **160**,501–514(2005)

[22] Z. H. Guo, J. Wu, H. Y. Lu, J. Z. Wang, Knowledge-Based Systems, **24**,1048–1056(2011)
[23] Tonglin Fu, Chen Wang, Mathematical Problems in Engineering, **2018**, 2506157 (2018)
[24] P. Jiang, X. J. Ma, Applied Mathematical Modelling, **40**,10631–10649 (2016)
[25] H. Jiang, J. Wang, J. Wu, W. Geng, Renew. Sustain. Energy Rev,**69**,1199–1217(2017)