Research Article

A Hybrid Model Based on Ensemble Empirical Mode Decomposition and Fruit Fly Optimization Algorithm for Wind Speed Forecasting

Zongxi Qu,1 Kequan Zhang,1 Jianzhou Wang,2 Wenyu Zhang,1 and Wennan Leng1

1Key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China
2School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

Correspondence should be addressed to Kequan Zhang; zhangkq@lzu.edu.cn

Received 26 February 2016; Revised 10 July 2016; Accepted 4 August 2016

Academic Editor: Ferhat Bingol

Copyright © 2016 Zongxi Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a type of clean and renewable energy, the superiority of wind power has increasingly captured the world’s attention. Reliable and precise wind speed prediction is vital for wind power generation systems. Thus, a more effective and precise prediction model is essentially needed in the field of wind speed forecasting. Most previous forecasting models could adapt to various wind speed series data; however, these models ignored the importance of the data preprocessing and model parameter optimization. In view of its importance, a novel hybrid ensemble learning paradigm is proposed. In this model, the original wind speed data is firstly divided into a finite set of signal components by ensemble empirical mode decomposition, and then each signal is predicted by several artificial intelligence models with optimized parameters by using the fruit fly optimization algorithm and the final prediction values were obtained by reconstructing the refined series. To estimate the forecasting ability of the proposed model, 15 min wind speed data for wind farms in the coastal areas of China was performed to forecast as a case study. The empirical results show that the proposed hybrid model is superior to some existing traditional forecasting models regarding forecast performance.

1. Introduction

The world’s current sources of fossil fuels will eventually be depleted, mainly due to high demand and, in some situations, extravagant consumption [1]. The recently posted Energy Outlook 2035 of British Petroleum predicts that primary energy consumption will increase by 37% between 2013 and 2035, with growth averaging 1.4% per year. Approximately 96% of the expected growth will be in countries that are not members of the Organization for Economic Cooperation and Development (OECD), with energy consumption growing at 2.2% per year [2]. According to some statistics, energy demand worldwide will grow rapidly by one-third from 2010 to 2035, and China and India will become the largest contributors, accounting for 50 percent of the growth during that period. Moreover, China is expected to be the largest oil importer by 2020 [2, 3]. To cope with the growing demand for energy, countries such as China can look to renewable energy sources to provide an opportunity for sustainable development. The significance of renewable sources was recently underpinned by a plethora of advocates and reports, which have mostly focused on wind energy studied by the related institutions and energy commissions of several countries [2, 4–7]. According to reports from the China National Renewable Energy Center (CNREC), wind resources in China are rich and promising prospects, carrying a potential of more than 3.0 TW, mostly in the Three North Areas, with an onshore potential of more than 2.6 TW. Before 2020, land-based wind power will dominate, with offshore wind power in the demonstration status. Furthermore, the annual discharge of carbon dioxide will be reduced to 1.5 billion tons and 3.0 billion tons in 2050 in the conservative and aggressive scenarios, and an estimated 720,000 jobs and 1,440,000 jobs will be created, respectively [4, 5]. Based on these figures, wind energy should be regarded as an appealing energy option because it is both abundant and
environmentally friendly; as such, wind energy will be able to satisfy the growing demand for electricity.

Wind energy has great influence on power grid security, power system operation, and market economics due to its intermittent nature, especially in areas with high wind power penetration. Thus, the analysis and assessment of wind energy are a meaningful but markedly difficult task for research. Because wind power generation hinges on wind speed, obtaining accurate wind speeds is important. To improve the precision of wind speed predictions, numerous methods have been proposed and developed in recent decades. These methods can be divided into three general types: physical models, conventional statistical models, and artificial intelligence models [8–11]. Physical models use weather prediction data, such as temperature, pressure, orography, obstacles, and surface roughness, for the best forecasting accuracy but are poor at short-term wind speed simulation. Conventional statistical models, in contrast, draw on vast historical data based on mathematical models usually involving conventional time series analysis, such as ARMA, ARIMA, or seasonal ARIMA models [12, 13], and achieve more accurate short-term wind speed predictions than physical models. However, conventional statistical models are imperfect. The fluctuating and intermittent characteristics of wind speed sequences require more complicated functions to capture the nonlinear relationships rather than assuming a linear correlation structure [14]. Given the development of statistical models along with the advent of artificial intelligence techniques, artificial intelligence models, including artificial neural networks (ANNs) and other mixed methods, have been proposed and are used in the field of wind speed forecasting [15–20]. For instance, because of the chaotic nature of wind time series, Alanis et al. [15] proposed a higher order neural network (HONN) based on an extended Kalman filter for model training, which provides accurate one-step-ahead predictions. Guo et al. [20] proposed a hybrid wind speed forecasting method employing a backpropagation (BP) neural network and seasonal exponential adjustment to remove seasonal effects from actual wind speed datasets. Wang et al. [21] exploited a radial basis function (RBF) neural network for wind speed prediction, and the effectiveness of this method was proved by a practical case. Zhou et al. [17] proposed a prediction method based on a support vector machine (SVM), for short-term wind speed prediction. De Giorgi et al. [19] adopted the ANNs to forecast wind speeds and compared them to the linear time-series-based model, with the ANNs providing a robust approach for wind prediction. All of these methods have improved the precision of wind speed predictions to some extent.

However, wind speed time series are highly noisy and unstable; therefore, using the primary wind speed series directly to establish prediction models is subject to large errors [22–24]. To build an effective prediction model, the features of original wind speed datasets must be fully analyzed and considered. The ensemble empirical mode decomposition (EEMD) [25] is an advanced, effective technology, which makes up for the deficiency of EMD [26] and has certain advantages over other typical decomposition approaches such as the wavelet decomposition and the Fourier decomposition [27]. With direct, intuitive, empirical, and adaptive data processing, EEMD was especially devised for nonlinear and complicated signal sequences, such as wind speed series. For example, Hu et al. [22] proposed a hybrid method based on the EEMD to disassemble the original wind speed datasets into a series of independent Intrinsic Mode Functions (IMFs) and use SVM to predict the values for IMFs in different frequencies. Jiang et al. [28] also proposed a hybrid model for high-speed rail demand forecasting based on EEMD, in which the original series are decomposed into certain signals with different frequencies and then the grey support vector machine (GSVM) is employed for forecasting. Zhou et al. [29] additionally proposed a hybrid method based on EEMD and the generalized regression neural network (GRNN). In this method, the original data are decomposed into different IMFs with corresponding frequencies and the residue component by EEMD, and then each component is taken as an input to establish GRNN forecasting model.

Each of the aforementioned models only employs a single ANN model to predict all of the signal sequences decomposed by EEMD; nevertheless, different signals have different characteristics, meaning that a simple individual model can no longer adapt to all properties of the data. Moreover, previous literature has not addressed which features are best suited for choosing the most appropriate approach. Thus, in our study, we propose a hybrid model based on a model selector that combines RBF, GRNN, and SVR to address signal data series with different characteristics to further improve forecasting accuracy.

In existing neural network training structures, model parameters are very vital factors affecting prediction precision, and different types of data require different parameters. The genetic algorithm (GA) and particle swarm optimization (PSO) algorithms are the most common approaches to optimize the parameters of neural network structures. Liu et al. [30] used the genetic algorithm to determine the weight coefficients of a combined model for wind speed forecasting. Zhao et al. [31] developed a combined model for energy consumption prediction based on model parameters optimization with the genetic algorithm. Ren et al. [32] applied the particle swarm optimization to set weight coefficients of a forecasting model for 6-hour wind speed forecasting. However, these meta-heuristic algorithms have the drawbacks of being hard to understand and achieving the global optimal solution slowly. The fruit fly optimization algorithm (FOA) [33] was a new optimization and evolutionary computation technique, which has distinct advantages in its simple computational process, fewer parameters to be fine-tuned, and stronger ability to search for global optimal solutions and outperforms other metaheuristics [34, 35]. In our study, we introduce the FOA algorithm to automatically determine the necessary parameters of the RBF, GRNN, and SVR models to achieve better performance.

The rest of the paper is organized as follows. Section 2 briefly introduces related methods while Section 3 describes the proposed hybrid approach in detail. Section 4 describes the dataset used for this study and discusses the forecasting results of proposed model compared with other prediction models. Section 5 concludes the work.
2. Related Methodology

This section briefly introduces EEMD, FOA, and three classical forecasting models: RBF, GRNN, and SVR, which will be used in our research.

2.1. RBF. The radial basis function (RBF) neural network is a type of feedforward network developed by Broomhead and Lowe [36]. This type of neural network is based on a supervised algorithm and has been widely applied to interpolation regression, prediction, and classification [37–39]. It has three layers of architecture, where there are no weights between the input hidden layers, and each hidden unit implements a radial-activated function. The Gaussian activation function is used in each neuron at the hidden layer, which can be expressed as

\[ h_j(x) = \exp \left( -\frac{\sum_{i=1}^{M} (x_i - \mu_j)^2}{2\sigma_j^2} \right), \]  

where \( x_i \) is the \( i \)th input sample, \( \mu_j \) is the mean value of the \( j \)th hidden unit presenting the center vector, \( \sigma_j \) is the covariance of the \( j \)th hidden unit denoting the width of the RBF kernel function, and \( M \) is the number of training samples.

The network output layer is linear so that the \( k \)th output is an affine function that can be expressed as

\[ y_k = \sum_{j=1}^{L} h_j w_{jk} + \beta_k, \]  

where \( w_{jk} \) is the weight between the \( k \)th output and \( j \)th hidden unit, \( \beta_k \) is the biased weight of the \( k \)th output, and \( L \) is the number of hidden nodes.

2.2. GRNN. The general regression neural network (GRNN), first proposed by Specht [40], is a very powerful computational technique used to solve nonlinear approximation problems based on nonlinear regression theory. The advantages of GRNNs include its good feasibility, simple structure, and fast convergence rate. It consists of four layers, and its basic principles are presented in Figure 1.

2.3. Support Vector Regression (SVR). SVR is a version of an SVM for regression and was introduced by Lasala et al. [41]. In the model, a regression function \( y = f(x) \) is applied to a forecast based on an input set. Attempts are made to minimize the generalization error that will impact generalization performance. Figure 2 illustrates the basic rules of SVR, and the more detailed information can be referenced in [42].

2.4. EEMD. The empirical mode decomposition (EMD) method, as an adaptive data analysis technique, has proven to be effective in analyzing nonlinear and nonstationary time series, such as wind speed series. It decomposes complex signals into IMFs that satisfy the following conditions.

1. In the whole data sequence, the number of extrema and the number of zero crossings in the entire sampled dataset must either be equal or differ at most by one.

2. The mean value at any point of the envelope defined by the local maxima and the envelope defined by the local minima is zero. With the hypothesis of decomposition and the definition of the IMF above, the EMD process of a raw data series \( x(t) (t = 1, 2, \ldots, T) \) can be formulated as

\[ x(t) = \sum_{k=1}^{m} \text{imf}_m(t) + r_m(t), \]  

where \( x(t) \) denotes any nonlinear and nonstationary signal, \( \text{imf}_m(t) \) is the \( m \)th IMF of the signal, and \( r_m(t) \) is the residual item, which can be a constant or the signal mean trend.

However, the EMD method is imperfect, and the mode-mixing problem [43] is encountered frequently in practical application. Due to the mentioned drawback of EMD, the advent of the EEMD method was proposed by Wu and Huang [25], and the procedures of EEMD can be presented as follows.

Step (a). Add a white noise series to the original data.

Step (b). Decompose the data with added white noise to IMFs through the EMD algorithm.

Step (c). Repeat the abovementioned two steps, but add white noise series at different scales each time.

Step (d). Calculate the means of each IMF of the decomposition to constitute the final IMFs.

As a result, the white noise series incorporated into the original signal can provide a uniform reference scale to facilitate the EMD process and, consequently, help extract the true IMFs. The relationship between the ensemble number, the error tolerance, and the added noise level can be described
according to the well-established statistical rule proved by Wu and Huang:

$$N_{\varepsilon} = \frac{\varepsilon^2}{\varepsilon_n^2},$$  \hspace{1cm} (4)

where \( \varepsilon \) is the amplitude of the added noise, \( \varepsilon_n \) is the final standard deviation of error, and \( N_{\varepsilon} \) is the value of ensemble members. Generally, it is suggested that an amplitude fixed at 0.2 will result in an exact result. In this study, we set the value of ensemble members to 100 and select the optimal standard deviation of white noise series from 0.1 to 0.2 with a \( k \)-fold cross-validation method.

2.5. Fruit Fly Optimization Algorithm (FOA). The fruit fly optimization algorithm (FOA), imitated by the food-finding behavior of the fruit fly, is a new swarm intelligence algorithm that was put forward by Pan in 2012 [33]. It is an interactive evolutionary computation method for finding global optimization and has been shown to perform better than traditional metaheuristic algorithms. The FOA succeeds in solving optimization challenges and has received significant attention in multiple scientific and academic fields.

The fruit fly, a type of insect, is superior to other species in visual and olfactory sensory abilities. It can make the most of its instinctive advantages to find food, even capable of smelling a food source from 40 km away. The fruit fly’s method of searching for food starts by using the olfactory organ to smell food odors in the air and then flies towards that location. Upon getting closer to the food location, it continues to seek food and the company's flocking location using its keen eyesight, and then it flies to that position too. Figure 3 shows the iterative process of food searching of a fruit fly swarm.

A rudimentary FOA algorithm is outlined as shown in Algorithm 1.

### 3. Combined Model

The combined model first applies the EEMD technique to decompose the original time series into a collection of relatively stationary subseries, and the model selection is used to select the optimal model above artificial neural networks based on FOA optimization for predicting each subseries. The prediction results are then aggregated to obtain the final prediction values of wind speed series.

#### 3.1. Model Selection

Through the process of EEMD, distinct information scales in the original wind speed series can be determined and decomposed into a set of IMFs. Additionally, different IMFs exhibit different frequency characteristics, and the instantaneous frequency of each IMF has its meaning at any point. Moreover, no clear theory exists to determine which characteristic is best suited for choosing the most suitable approach. Thus, we must describe some performance metrics to comprehensively measure the strengths of different models. To evaluate the forecast capacity of the proposed models, three evaluation criteria are applied in model selection. They are the mean absolute error (MAE), root mean-square error (RMSE), and index of agreement (IA), as shown in Table 1.

Here, \( y_n \) and \( \hat{y}_n \) denote the real and predicted values at time \( n \), respectively. \( N \) is the sample size. The IA is a dimensionless indicator that portrays the similarity between the observed and forecasted tendencies. The range of IA is from 0 to 1 and for a “perfect” model the value of IA is close to 1 while the MAE and RMSE are equivalent to 0.

#### Table 1: Four evaluation rules.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Equation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>( \text{MAE} = \frac{1}{N} \sum_{n=1}^{N}</td>
<td>y_n - \hat{y}_n</td>
</tr>
<tr>
<td>RMSE</td>
<td>( \text{RMSE} = \left( \frac{1}{N} \sum_{n=1}^{N} (y_n - \hat{y}_n)^2 \right)^{1/2} )</td>
<td>The root mean-square forecast error</td>
</tr>
<tr>
<td>IA</td>
<td>( \text{IA} = 1 - \frac{\sum_{t=1}^{T} (y_t - \hat{y}<em>t)^2}{\sum</em>{t=1}^{T} (</td>
<td>y_t - \bar{y}</td>
</tr>
</tbody>
</table>
Objective:
Maximizes smell concentration

Output:
The best smell concentration (Smellbest)

Parameters:
- Iteration number (Maxgen)
- Population size (sizepop)
- Location range (LR)
- Random fly direction and distance zone of fruit fly (Smellbest)

Algorithm 1: FOA.

The main processes of the proposed hybrid model are demonstrated in Figure 4. The detailed steps of the hybrid model are as follows.

Step 1 (EEMD process). The raw data series are decomposed into 7 different IMFs and a residue $R$. Because the first IMF with high frequency is evoked by noise, it is removed directly, and the rest are used for forecasting.

Step 2 (model selection and optimization of model parameters). First, select the appropriate parameter from the RBF, GRNN, and SVR models by the FOA. Next, the aforementioned models are then selected by model selection to forecast IMFs and a residual $R$.

Step 3 (ensemble forecast). Combine the forecasting results of each signal component to obtain the final result.

4. Results and Analysis

In this section, the process descriptions of RBF, GRNN, and SVR models optimized by the FOA are presented firstly and then followed by the process descriptions of the model selection. Results conclude with the final forecasting results of the hybrid model compared to other different forecasting models.

4.1. Data Selection. Shandong Province located in eastern China has abundant wind energy resources. In our study, the wind speed series from the wind farm in Weihai was used to examine the performance of the combined model. Figures 5(a) and 5(b) present the statistical measures and visual graphs of four wind speed datasets, which show apparent differences between the four seasons. Thus, the original wind speed data, picked randomly corresponding to the four seasons of the year, are used to test whether the proposed models can be applied on different occasions. The wind speed data were sampled at an interval of 15 min, so there are 96 data records per day. Data from 4 days, providing a total of 384 points of 15 min data, were selected for model training, and the next 48 of the 15 min data values were used to test the effectiveness of the developed hybrid model (as shown in Figure 5(b)).

4.2. The Performance Metric. Forecasting accuracy is an important criterion in the evaluation of forecasting models. In this paper, three metric rules were applied to evaluate the
accuracy of forecasting models, as shown in Table 1. In addition, two benchmark models and bias-variance framework are used to test the hybrid model.

4.2.1. Persistence Model. The persistence model as a simple statistical model, which has simple calculation and provides accurate prediction in a very short time, has been widely used as benchmark model to evaluate the accuracy of more advanced forecasting model. The persistence model can be given by

\[ \hat{p}_{t+k} = p_t, \]  

(5)

where \( \hat{p} \) is the forecasting value, \( t \) is a time index, and \( k \) is the look-ahead time.

4.2.2. Autoregressive Integrated Moving Average (ARIMA). ARIMA model is widely used because it can characterize nonlinear data. A general ARIMA model is known as ARIMA \((p, d, q)\), where \( p \) is the order of the autoregressive part, \( d \) is the number of differences from the original time series data to make it stationary, and \( q \) is the order of the moving average portion. The general equation for ARIMA models is

\[ y_k = \sum_{m=1}^{p} f_m y_{k-m} + \sum_{n=1}^{q} \sigma_n e_{k-m} + \epsilon_k, \]  

(6)

where \( y_k \) is the observed value at time \( k \), \( f_m \) is the \( m \)th autoregressive parameter, \( \sigma_n \) is the \( n \)th moving average parameter, and \( \epsilon_k \) is the error at time \( k \).

4.2.3. Bias-Variance Framework. To estimate the availability of the wind speed forecasting models, bias-variance framework [44] was employed to evaluate accuracy and stability of the proposed hybrid model and single models. Let \( x_i - \bar{x}_i \) be the difference between observed value \( x_i \) and predicted value \( \bar{x}_i \), and the average difference over all points is

\[ \frac{1}{T} \sum_{t=1}^{T} (x_t - \bar{x}_t) = \frac{1}{T} \sum_{t=1}^{T} x_t - \frac{1}{T} \sum_{t=1}^{T} \bar{x}_t, \]  

(7)

where \( t \) is the \( t \)th data for performance evaluation and \( T \) is all the forecasting data used for performance evaluation. The expectation of the total number of forecasting values is \( E(\bar{x}) = (1/T) \sum_{t=1}^{T} \bar{x}_t \), and the expectation of the actual value is \( x = (1/T) \sum_{t=1}^{T} x_t \). The bias-variance framework can be decomposed as follows:

\[ E(\bar{x} - x)^2 = E(\bar{x} - E(\bar{x})) + E(E(\bar{x}) - x)^2 \]

\[ = E(\bar{x} - E(\bar{x}))^2 + (E(\bar{x}) - x)^2 \]  

(8)

where \( \text{Bias}^2(\bar{x}) \) indicates the prediction accuracy of the forecasting model and \( \text{Var}(\bar{x}) \) demonstrates the stability.

4.3. Process of Parameter Optimization. Selecting the appropriate parameter is very critical to improving the accuracy of model prediction; thus, the abovementioned FOA is used to optimize the parameters of the RBF, GRNN, and SVR.
models (as shown in Figure 6(a)). First, in the RBF model, the centers and widths \([\mu, \theta]\) of the basic functions should be substituted by the smell concentration judgment value \((S_i)\) of the FOA and other experiment parameters of RBF are shown in Table 2. The smoothing parameter \((\sigma)\) of the GRNN, the penalty parameter \((C)\), and loss function parameter \((\epsilon)\) of the SVR are also represented by \((S_i)\) of the FOA. After that, the offspring is entered into the three models, and the smell concentration value is calculated again. Then, smell concentration \((\text{Smell}_i)\), replacing \((S_i)\) with the smell concentration judgment function (also called the fitness function), is calculated; with the smaller value of fitness function, the better results will be found. Through the fruit fly’s random food searching using its sensitive sense of smell and flocking to the location of the highest smell concentration using its vision, the optimal parameters of the three models are obtained.

To test the effect of the model parameters optimized by the FOA, the four seasons of wind speed data were selected. The three criterions were employed to evaluate the performance of the three models optimized by the FOA. Results of the comparison are shown in Table 3 and Figure 6(b). It can be clearly observed that the FOARBF, FOAGRNN, and FOASVR consistently have the least statistical error as indicated by the MAE, RMSE, and IA. One can conclude that the FOA optimization can effectively improve the prediction performance of the traditional neural network model.
Start

Initialize the location of fruit fly swarm, population size, maximum iteration number

Every fruit fly searching for the food by osphresis

Calculate $S_i$ and $\text{Dist}_i$

Find and keep the maximal smell concentration value, and update the best location

Max iteration?

End

$P_j = \exp \left[ \frac{(X - X_j)^T (X - X_j)}{2\sigma^2} \right]$

$K(x, x_1) = \exp \left( -\frac{\sum_{i=1}^{M} (x_i - \mu_j)^2}{2\theta^2} \right)$

$f(x) = \min \frac{1}{2} ||W||^2 + C \sum_{i=1}^{k} \varepsilon \left( f(x_i) - y_i \right)$

$RBF$ $GRNN$ $SVR$

GRNN
FOAGRNN
SVR
FOASVR
RBF
FOARBFB

Figure 6: The procedures of RBF, GRNN, and SVR optimized by FOA.
4.4. The Process of Model Selection. Given the complexity and chaos of the original wind speed series, the tendency of wind speed is very difficult to directly predict by using the above-mentioned individual models. As such, the original wind speed datasets are decomposed into several IMFs and a residue $R(n)$ by EEMD, which make the raw datasets easier to simulate. The FOARBF, FOAGRNN, and FOASVR models are used to forecast each IMF and the residue $R(n)$ as the input nodes, hidden nodes, and output nodes of the three neural networks are set to 4, 9, and 1, respectively. The rolling operation method was used in this paper, and the wind speed data in four seasons were selected to test the proposed models.

The selection process of the hybrid model is shown in Figure 7 and its results are shown in Tables 4–7, and it can be clearly observed that each individual model exhibits the
best performance at a specific IMF. Nevertheless, no single model can perform best in all situations. For example, Table 4 shows the forecasting results in springtime and reveals that the FOARBF provides the best results at the IMF5 and IMF7.
The FOASVR, however, exhibits the lowest MAE and RMSE values among all individual models at IMF2, IMF3, and IMF6 while the lowest value at IMF4 and $R(n)$ is achieved by the FOAGRNN. The analysis of three other seasons can be seen in the Appendix.

4.5. Forecasting Results and Comparative Analysis. In the abovementioned process, the six independent IMFs and one residual decomposed by EEMD are predicted by three different models: FOARBF, FOAGRNN, and FOASVR. The optimal model corresponding to each IMF and $R(n)$ is then selected through model selection. In Step 3, each IMF is predicted by the selected optimal methods, and the final results are obtained by assembling the forecasting results of each IMF.

4.5.1. Forecasting Comparison Results. To evaluate the performance accuracy of the proposed hybrid model based on model selector, three single models and two benchmark models are employed to compare with the hybrid model. Single models include the FOARBF, FOAGRNN, and FOASVR, each of which is used for forecasting all of the signals decomposed by EEMD. Two benchmark models include persistence model and ARIMA model. The comparison results for forecasting ability are as shown in Table 8. Detailed analyses are elaborated as follows:

1. By comparing the hybrid model with the other five models, the lowest MAE and RMSE values are achieved by hybrid model. In particular, the IA values of the hybrid model were improved by 10.84%, 11.40%, 5.82%, 7.93%, and 3.04% on four seasons compared with the persistence model, ARIMA model, EEMD-FOARBF, EEMD-FOAGRNN, and EEMD-FOASVR.

2. When compared to benchmark model, the EEMD-FOARBF, EEMD-FOAGRNN, EEMD-FOASVR, and the hybrid model show optimal forecasting results according to MAE, RMSE, and IA, likely because EEMD technology is effective in improving the forecasting accuracy as a data preprocess step.

3. When compared to the EEMD-FOARBF, EEMD-FOAGRNN, and EEMD-FOASVR, the hybrid method also shows better prediction results, indicating that the hybrid method can take advantages of each individual model to obtain more complete information.

Above all, the proposed hybrid model has been verified as an effective approach for improving the forecasting performance through the analysis of the prediction results.

4.5.2. Tested with Bias-Variance Framework. Table 9 shows the results of the bias-variance test: the values of bias indicate the prediction accuracy of the forecasting model and values of variance demonstrate the stability. The results reveal the following:

1. The absolute values of the biases of the hybrid model are less than those of the other models, which indicates that the hybrid model has a higher accuracy in
wind speed forecasting. The variance results also show that the hybrid model is more stable.

(2) The results of bias and variance values of the EEMD-FOARBF, EEMD-FOAGRNN, EEMD-FOASVR, and hybrid model are less than the persistence model and ARIMA; this reveals EEMD and FOA are effective approaches for improving the accuracy and stability of forecasting models.

Thus, it is clear that the hybrid model has a higher accuracy and stability in wind speed forecasting, and it performs much better than individual models in forecasting.

5. Conclusions

Reliable and precise wind speed forecasting is vital for wind power generation systems. However, wind speed shows nonlinearity and nonstationarity, which pose great challenges to the task of predicting wind speed precisely. Regarding the currently available forecasting models, the single model applied for forecasting wind speed has limited capacity and is not suitable for all situations. The appropriate selection approach of the hybrid model can give full play to the strengths of each of the individual models and make each individual model perform in its specific manner. For these reasons, we proposed a hybrid model based on EEMD that combines three commonly used neural networks optimized by the FOA. The main contributions of this model are summarized as follows. (1) Due to the instability of wind series, EEMD technique is utilized as a preprocessing approach to decompose the original time series into a collection of relatively stationary subseries for forecasting. (2) To overcome the drawbacks of the unstable forecasting results of the RBF, GRNN, and SVR, the FOA optimization is applied to improve the prediction performance of the traditional forecasting model. (3) Because the IMF signals with different characteristics are hard to forecast by a single model, a model selection combining FOARBF, FOAGRNN, and FOASVR is proposed to further improve forecasting accuracy. The experimental results indicate that the proposed hybrid model has minimum statistical error in terms of MAE, RMSE, IA, and bias variance, and it proved that the proposed hybrid method performs better than single models and is superior to other hybrid models as well, such as the EEMD-FOARBF, EEMD-FOAGRNN, and EEMD-FOASVR. Based on the abovementioned analysis, we conclude that the proposed hybrid model can not only take full advantage of several single ANNs to improve prediction accuracy but also easily implement the task in wind parks.

Appendix

To further prove that the proposed hybrid model can select the best model for different cases, the forecasting results in other seasons can be seen in Tables 4–6. For example, Table 4 shows the experimental results from three single models in the summer. Among all the single models, when the FOARBF was applied, the value of IA was higher than those of the other methods at IMF2 and IMF6. At IMF4, IMF7, and $R(n)$, the FOAGRNN provides the optimal results. At other signals, the results from the FOASVR are the best. Table 5 shows the results in autumn. Among all the models, at IMF2, IMF3, IMF4, and IM6, the FOASVR performs the best while the FOAGRNN performs better than the others at IMF7 and $R(n)$. Meanwhile, the FOARBF provides the optimal results at other signals. The forecasting results of three single models in winter are presented in Table 6. At IMF6, IMF7, and $R(n)$, the most accurate results belong to the FOAGRNN. When the FOASVR is used, the results are more accurate from IMF2 to IMF4. Results show that the FOARBF only performs desirably at IMF5. From Tables 3–6, we find that FOASVR always performs well at high frequency signals, FOAGRNN works well at low frequency signals, and FOARBF usually provides optimal results at middle frequency signals. Consequently, no single model provides the best results for all of the signals, but each model has its strengths at special IMFs. Therefore, the best-suited model is chosen based on different conditions.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by the National Natural Science Foundation Project (41225018) and Arid Meteorology Research Fund (IAM201305).

References


Advances in Meteorology

State-of-the-art Methods and software tools for short-term prediction of wind energy production.


Submit your manuscripts at http://www.hindawi.com