Renewable Energy 113 (2017) 1345–1358

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction

Deyun Wang [a, b, c, *], Hongyuan Luo [a, b], Olivier Grunder [c], Yanbing Lin [a, b]

[a] School of Economics and Management, China University of Geosciences, Wuhan 430074, China
[b] Mineral Resource Strategy and Policy Research Center, China University of Geosciences, Wuhan 430074, China
[c] Université de Bourgogne Franche-Comté, UTBM, IRTES, Rue Thierry Mieg, 90010 Belfort Cedex, France

ARTICLE INFO

Article history:
Received 24 July 2016
Received in revised form
26 June 2017
Accepted 28 June 2017
Available online 29 June 2017

Keywords:
Multi-step ahead
Wind speed forecasting
Variational mode decomposition
Phase space reconstruction
Wavelet neural network

ABSTRACT

Accurate wind speed forecasting is crucial to reliable and secure power generation system. However, the intermittent and unstable nature of wind speed makes it very difficult to be predicted accurately. This paper proposes a novel hybrid model based on variational mode decomposition (VMD), phase space reconstruction (PSR) and wavelet neural network optimized by genetic algorithm (GAWNN) for multi-step ahead wind speed forecasting. In the proposed model, VMD is firstly applied to disassemble the original wind speed series into a number of components in order to improve the overall prediction accuracy. Then, the multi-step ahead forecasting for each component is conducted using GAWNN model in which the input-output sample pairs are determined by PSR technique. Finally, the ultimate forecast series of wind speed is obtained by aggregating the forecast result of each component. The proposed model is tested using two real-world wind speed series collected respectively in spring and autumn from a wind farm located in Xinjiang, China. The experimental results show that the proposed model outperforms all other comparison models including persistence method, PSR-BPNN, PSR-WNN, PSR-GAWNN and EEMD-PSR-GAWNN models adopted in this paper, which demonstrates that the proposed model has superior performances for multi-step ahead wind speed forecasting.

1. Introduction

Wind power, as a clean and renewable energy resource, has developed fast in the world. By the end of 2015, the total cumulative installed capacity from wind power around the world was approximately 432883 MW and increased by 17% compared to the previous year (369553 MW). With the rapid development of wind power industry, wind speed forecasting has already become a hot issue within the research field of power generation, due to its important role in energy generation planning, power grid integration and turbine maintenance scheduling [1]. However, wind speed forecasting has fully been shown to be an extremely difficult task because of the unstable and intermittent characteristics associated with wind energy [2]. Therefore, it is of great interest to develop relatively accurate wind speed forecasting models [3].

Over the past several decades, a great number of methods have been proposed in the literature to forecast the wind speed. These methods can be roughly categorized into the following four types: (a) numerical weather prediction (NWP) methods; (b) traditional statistical methods; (c) artificial intelligent (AI) methods; (d) hybrid methods [4]. NWP methods need to take into account the physical description of wind farm and therefore usually require abundant physical background knowledge and substantial computational resources [5]. In addition, wind speed is influenced by many complicated factors such as pressure, temperature, obstacles and roughness which are usually difficult to be determined in practical application. Therefore, it is very difficult to forecast wind speed using NWP methods, which has been confirmed by many researchers. For example, Soman et al. [6] presented an overview of comparative analysis of various wind speed forecasting techniques including NWP, statistical models, AI based models and hybrid techniques over different time scales, their results revealed that it is really a big challenge job to build a precise forecasting model using

* Corresponding author. School of Economics and Management, China University of Geosciences, Wuhan 430074, China.
E-mail address: wang.deyun@hotmail.com (D. Wang).

http://dx.doi.org/10.1016/j.renene.2017.06.095
0960-1481/

NWP methods. Consequently, many researchers turn to forecasting wind speed based upon the traditional statistical methods, AI based methods and hybrid methods.

Traditional statistical models are constructed based on the mature statistical equations to obtain the potential evolution rule from historical data [7]. The most commonly used traditional statistical models for wind speed forecasting include autoregressive model (AR), autoregressive moving average (ARMA), generalized autoregressive conditional heteroskedasticity (GARCH), Bayesian methods, Markov chain and Kalman filters methods. For instance, Schlink and Tetzlaff [8] applied AR for wind speed prediction at an airport, and the results illustrated that the intervals' width produced by AR were narrower than the intervals generated by persistence model. Lydia et al. [9] forecasted wind speed and power based on linear and non-linear ARMA models with and without external variables. Liu et al. [10] applied a hybrid ARMA-GARCH(-M) model to simulate the volatility and mean of wind speed. Li et al. [11] developed a robust two-step methodology to forecast wind speed based on three kinds of neural network models (back-propagation (BP) network, adaptive linear element network and radial basis function network) and Bayesian combination algorithm. Song et al. [12] employed a Markow-switching model for performing interval and point wind speed forecasting. Zuluaga et al. [13] presented an experimental comparison of three different variants of robust Kalman filtering for one-step ahead prediction of wind speed, and the results showed the superior performances of these proposed models.

As for the AI based methods, artificial neural networks (ANN), support vector regression (SVR) and least squares support vector machines (LSSVM) might be the most frequently used models for wind speed forecasting, and the empirical investigations have proven their superiority compared with the traditional linear models. For instance, Guo et al. [14] established a novel hybrid model for wind speed forecasting based on a BP neural network and the idea of eliminating seasonal effects from original wind speed datasets using seasonal exponential adjustment. Chang et al. [15] utilized ANN model for wind speed prediction, and the results indicated the good performance of the proposed model. Wang et al. [16] established a SVR model for preprocessing the irregular data existed in the actual wind speed series, and then calculated an absolute error between the forecast and observed values. If the absolute error is below a pre-specified threshold, the observation will be considered as outlier and then it will be replaced by the forecast value, otherwise the observation is accepted. Zhou et al. [17] developed a LSSVM based model for one-step ahead wind speed forecasting and obtained relatively accurate results compared to the persistence approach.

In order to further improve the forecast accuracy, some single technologies including fuzzy logic, SVR, ANN and some other optimization algorithms are systematically integrated and employed to predict wind speed for different applications [18–21]. However, the multiple frequency components existed in the wind speed series are always the challenging parts in forecasting, making the models which work on the original wind speed series cannot handle them appropriately. Fortunately, with the currently popular used concept of “divide and conquer” (or “decomposition and ensemble”) [22], a number of various decomposition-ensemble based hybrid models have been developed and widely used for prediction of wind speed, crude oil price and nuclear energy consumption in order to improve the overall forecast accuracy. For instance, Liu et al. [23] developed a new wind speed forecasting approach based on fast ensemble empirical model decomposition (FEEMD), genetic algorithm (GA), mind evolutionary algorithm (MEA) and ANN, and the results demonstrated that among all the involved methods, the proposed hybrid FEEMD-MEA-MLP model has the best forecasting performance. Wang et al. [24] proposed a wind speed forecasting method based upon improved EMD (EEMD) and optimized BP neural network (GA-BP) for on-line short term (1 h) and ultra-short term (10 min) wind speed forecasting, and computational results have shown the good performance of EEMD. Meng et al. [25] proposed a novel hybrid wind speed forecasting model based on crisscross optimization algorithm, wavelet packet decomposition and ANN, and the results based upon two wind speed series collected from a wind power observation station located in the Netherlands demonstrated the good performance of the proposed hybrid model. Wang et al. [26] developed a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm for multi-step electricity price forecasting, and the experimental results illustrated the superior performance of the proposed model.

Notable, multi-step ahead forecasting is much more important than one-step ahead forecasting in wind industry, because the multi-step ahead forecasting can effectively capture the dynamic behavior of wind speed in the future, which is crucial for improving the security and economic benefits of wind power system. However, the multi-step ahead forecasting may accumulate the forecast errors with increasing of the horizons, which leads to relatively large overall forecast error in real applications [27,28]. In order to improve the forecast accuracy, this paper proposes a novel hybrid model for multi-step ahead wind speed forecasting based on VMD whose superior performance has been justified in Refs. [29,30] and GAWNN. In the proposed model, the original wind speed series is firstly decomposed into a certain number of components using VMD. Then, the multi-step ahead forecasting for each component is conducted using GAWNN model in which the input-output sample pairs are determined by PSR technique. Finally, the forecast series of wind speed is obtained by aggregating the forecast result of each component. In contrast to aforementioned researches, there are two contributions or novelties in this paper: (1) VMD, as a competitive signal decomposition method, is firstly combined with PSR method and machine learning methods to tackle the instability and nonlinearity of the original wind speed series; (2) PSR method is specifically applied for determining the input-output sample pairs of GAWNN network in order to improve the overall forecast accuracy.

The remainder of this paper is organized as follows: Section 2 briefly describes the fundamental methods including VMD, PSR, wavelet theory, GA, ANN and GAWNN. Section 3 introduces the proposed hybrid model. Section 4 firstly presents the available data sets and error indicators used for experimentation, and then provides the forecast results of the proposed hybrid model and the comparative results. Finally, Section 5 provides the conclusions.

2. Methodology

The research methodology used in this study includes VMD, PSR, wavelet theory, GA, and BPNN. The brief description of those methods is stated as follows.

2.1. Variational mode decomposition

Variational mode decomposition (VMD) proposed by Dragomiretskiy and Zosso [31] in 2014, is a newly non-recursive signal processing technique, which is employed to adaptively decompose a real valued signal into a discrete set of band-limited sub-signals (known as modes) y_k owning specific sparsity properties. Each mode y_k generated by VMD technology can be compressed around a center pulsation w_k which is determined along with the decomposition process. The process of estimating the bandwidth of each mode includes the following three steps: (1) for each mode y_k,

applying the Hilbert transform to calculate the associated analytic signal so that a unilateral frequency spectrum can be obtained; (2) mixing with an exponential tuned to the respective estimated center frequency in order to shift the mode's frequency spectrum to baseband; (3) estimating the bandwidth for each mode y_k by using the H^1 Gaussian smoothness of the demodulated signal. Thus, the constrained variational problem can be presented as follows:

$$\begin{aligned}&\min_{\{y_k\},\{w_k\}}\left\{\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\otimes y_k(t)\right]e^{-jw_kt}\right\|_2^2\right\}\\&s.t.\quad\sum_{k=1}^{K}y_k=f(t)\end{aligned}\tag{1}$$

where $f(t)$ is the original main signal, and y_k is the kth component of signal $f(t)$; w_k, $\delta(t)$ and $\otimes$ represent the center frequency of y_k, the Dirac distribution and convolution operator, respectively; and k denotes the number of modes, t is time script.

Taking both penalty term and Lagrangian multipliers λ into consideration can convert the above constrained problem to the unconstrained one and make it easier to be addressed, which is shown as follows:

$$\begin{aligned}L(\{y_k\},\{w_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\otimes y_k(t)\right]e^{-jw_kt}\right\|_2^2\\+\left\|f(t)-\sum_{k=1}^{K}y_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{K}y_k(t)\right\rangle\end{aligned}\tag{2}$$

where α represents the balancing parameter of the data-fidelity constraint.

The augmented Lagrangian L is determined by Eq. (2) and its saddle point in a sequence of iterative sub-optimizations can be found through using the alternate direction method of multipliers (ADMM). According to ADMM optimization approach, it is assumed that updating y_k and w_k in two directions help realize the analysis process of VMD. Consequently, the solutions for y_k and w_k are expressed as follows:

$$\hat{y}_k^{n+1}=\frac{\hat{f}(w)-\sum_{i\neq k}\hat{y}_i(w)+\frac{\hat{\lambda}(w)}{2}}{1+2\alpha(w-w_k)^2}\quad,\quad w_k^{n+1}=\frac{\int_0^\infty w\left|\hat{y}_k^{n+1}(w)\right|^2dw}{\int_0^\infty\left|\hat{y}_k^{n+1}(w)\right|^2dw}\tag{3}$$

where $\hat{f}(w)$, $\hat{y}_i(w)$, $\hat{\lambda}(w)$ and $\hat{y}_k^{n+1}(w)$ denote the Fourier transforms of $f(t)$, $y_i(t)$, $\lambda(t)$ and $y_k^{n+1}(t)$, respectively, and n represents the number of iterations.

2.2. Phase space reconstruction

The phase space reconstruction (PSR) theory was first proposed by Takens in 1981 [32]. PSR is considered as a useful method to reconstruct the phase space of chaotic time series. A time series $x=\{x_i|i=1,2,\ldots,N\}$ can be reconstructed in a multi-dimensional phase space such that:

$$X(i)=\{x(i),x(i+\tau),\ldots,x(i+(m-1)\tau)\}\tag{4}$$

where, $i=1,2,\ldots,M$, $M=N-(m-1)\tau$. τ is the time delay parameter. m is the embedding dimension.

In the study, the C-C method is used to reconstruct the time series. The correlation integral for the embedded time series is defined in Eq. (5).

$$C(m,N,r,\tau)=\frac{2}{M(M-1)}\sum_{1\le i<j\le M}\theta(r-\|X(i)-X(j)\|)\tag{5}$$

where $\|X(i)-X(j)\|$ is the sup-norm, τ is the time delay parameter, $\theta(x)=\begin{cases}0, & x<0\\1, & x\ge 0\end{cases}$. The correlation integral is a cumulative distribution function, which means that the Euclidean distance between any two points will be less than r in the phase space. The reconstruction phase space can be shown as follows.

$$X=\begin{matrix}X(1)\\X(2)\\\vdots\\X(M)\end{matrix}=\begin{matrix}x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau)\\x(2) & x(2+\tau) & \cdots & x(2+(m-1)\tau)\\\vdots & \vdots & \vdots & \vdots\\x(M) & x(M+\tau) & \cdots & x(M+(m-1)\tau)\end{matrix}\tag{6}$$

where $i=1,2,\ldots,M, M=N-(m-1)\tau$. N is the number of data. τ is the time delay parameter. m is the embedding dimension.

A time series is divided into t disjoint sub-sequence, and the correlation integral $S\left(m,\frac{N}{t},r,\tau\right)$ of each sub-sequence can be defined as follows:

$$S(m,N,r,\tau)=\frac{1}{t}\sum_{s=1}^{t}\left[C_s\left(m,\frac{N}{t},r,\tau\right)-C_s^m\left(m,\frac{N}{t},r,\tau\right)\right]\tag{7}$$

The time series of relevance can be described by $S\left(m,\frac{N}{t},r,\tau\right)$, and the best embedding dimension m and the best delay time τ can be determined by the mean of $S\left(m,\frac{N}{t},r,\tau\right)$.

2.3. Hybrid GAWNN forecasting model

2.3.1. Wavelet theory

The theory of wavelet proposed by Morlet in 1980 appears to be more effective than the Fourier transform in processing non-stationary time series and is preferred to a windowed Fourier transform [33]. Fourier transform can provide frequency domain description for temporal processes of function or signal which is only available in the time domain. Fourier analysis gives frequency information which can only be obtained for the complete duration of the signal, and it does not provide any information about the local variations in time and requires the signal to be stationary. In order to well process the stationary signals, the Fourier transforms decompose the signals into a certain number of linear combinations of sines and cosines waves, however, the continuous wavelet transforms decompose the non-stationary signals into a number of linear combinations of wavelets.

The continuous wavelet transform is defined as follows:

$$W_f(a,b)=\frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty}f(t)\psi\left(\frac{t-b}{a}\right)dt\tag{8}$$

where a is the scaling function used to stretch or compress mother wavelet $\psi(t)$ that is related to the frequency of the signal; b is the translation function used to shift mother wavelet $\psi(t)$ to a time domain of the signal; $f(t)$ is the input signal. The mother wavelet $\psi(t)$ is defined as follows.

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\tag{9}$$

In this study, the Morlet function which is shown as follows is taken as the mother wavelet in the analysis process.

$$\psi(t) = Ce^{-t^2/2}\cos 1.75\,t \tag{10}$$

2.3.2. Genetic algorithm

Genetic algorithm (GA) proposed by Holland in 1975 is an intelligent global optimization algorithm based on the theory of natural selection. Because of good searching ability, GA has been widely applied in many practical fields and obtains satisfied results [34]. The mechanism of GA is illustrated as follows. GA begins with a pre-specified number of solutions or individuals denoted by population. In the population, each solution is called a chromosome (or individual) which denotes a feasible solution in the search space. Then, GA conducts the evolving process where the chromosomes are improved through a series of operations including selection, crossover and mutation. In each iteration, the chromosomes with lower fitness values are eliminated with a high probability. The operations including selection, crossover and mutation are repeated until the stop criteria is satisfied, and then the chromosome with the largest fitness value is decoded and adopted as the best solution.

2.3.3. Artificial neural networks

Artificial neural networks (ANNs) include a family of intelligent models that mimic the biological neural networks. ANNs are often used to approximate or estimate functions that depend on a large number of inputs. Since the distinguish performance of ANNs, they have been popularly used in many practical fields such as industrial engineering, economic forecasting and evaluating. The feed-forward back propagation network (BPNN) including one or more hidden layers is one of the ANNs, which has a relatively simple structure and thus can be realized easily. The BPNN network used in this study has a three-layer network consisting of an input layer, a hidden layer and an output layer. BPNN distinguishes itself by the presence of hidden layers whose computation nodes are correspondingly called hidden neurons. The function of hidden neurons is to connect the input and output of the network. Given a training set of input-output data, the most common learning rule for multilayer perceptron (MLP) neural networks is the back-propagation algorithm which involves the following two phases: the first one is a feed-forward phase in which the external input information at the input nodes is propagated forward to compute the output information signal at the output unit; the second one is a backward phase in which modifications to the connection weights are made based on the differences between the computed and observed information signals at the output units. In this study, a tangent sigmoid function is adopted as the neuron transfer function for BPNN.

2.3.4. Hybrid GAWNN forecasting model

The hybrid GAWNN model is based on the BPNN network, while the hidden neurons' transfer function is replaced by the Morlet wavelet function. In order to avoid falling into local optimum, the weight values of BPNN and the parameters of wavelet, a and b, are initially optimized through GA. Moreover, to solve the problem of slow convergence speed of the standard BPNN, this paper uses a dynamic learning rate mechanism to update the parameter values. Suppose that $X = (x_1, \ldots, x_m)$ and $Y = (y_1, \ldots, y_m)$ are the input and output vectors of the hybrid model and let l, W_{ik} and W_{kj} $(i = 1, 2, \ldots, m\ k = 1, 2, \ldots, l\ j = 1, 2, \ldots, n)$, $k = 1, 2, \ldots, l$, $j = 1, 2, \ldots, n)$ be the number of hidden neurons, weight values of the input layer and output layer, respectively, then the structure and flowchart of GAWNN forecasting model can be presented in Figs. 1 and 2, respectively.

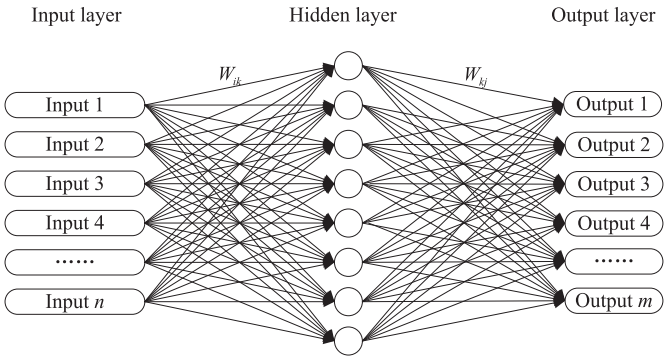


Fig. 1. The structure of WNN model.

Moreover, suppose that GAWNN model has a number of P studying samples which are $I = [X_1, \ldots, X_P], O = [Y_1, \ldots, Y_P]$, where $X_p = [x_{1p}, \ldots, x_{mp}], Y_p = [y_{1p}, \ldots, y_{np}]$, $p\ (= 1, \ldots, P)$ represents the p_{th} input-output sample pair. Let $Z_k = [z_{k1}, \ldots, z_{kp}]$ be the output of the k_{th} hidden neuron, then the detailed steps of GAWNN model can be described as follows.

Stage 1. Optimize the weight values of GAWNN model, i.e., w_{ik} and w_{kj};

Step 1: Initialize the probability of crossover operator p_c and mutation operator p_m. Set the index of generation $g = 0$. Code parameters w_{ik}, w_{kj} and a_k, b_k as one chromosome using real number coding method. Then generate a number of *popsize* chromosomes as an initial population $pop(0)$ by both random and artificial ways.

Step 2: Decode each chromosome and allocate its corresponding values to w_{ik}, w_{kj} and a_k, b_k. Calculate the network's output and the difference between the network's output and real values. Take the reciprocal of the above difference as the fitness function of GA, i.e.,

$$f = \frac{1}{\frac{1}{2}\sum_{p=1}^{P}\sum_{j=1}^{n}\left(y_{jp} - o_{jp}\right)^2} \tag{11}$$

Step 3: If the stop criterion is satisfied, go to Step 5;

Step 4: Repetition. Selection operation: select a pair of parent chromosomes from the current population for creation of the next generation by a way of “roulette wheel”; crossover operation: with probability p_c, apply the simple two-point crossover operators to these two parents; mutation operation: with probability p_c, apply the simple one-point mutation operator to generate a new chromosome.

Step 5: Stop the algorithm and choose the best chromosome in the population as the final solution of the problem. Decode the best chromosome and allocate its corresponding values to w_{ik}, w_{kj} and a_k, b_k.

Stage 2. GAWNN training process

Step 6: Calculate the output vector. Input the p_{th} sample, and obtain the output of the k_{th} hidden neuron through the following formula.

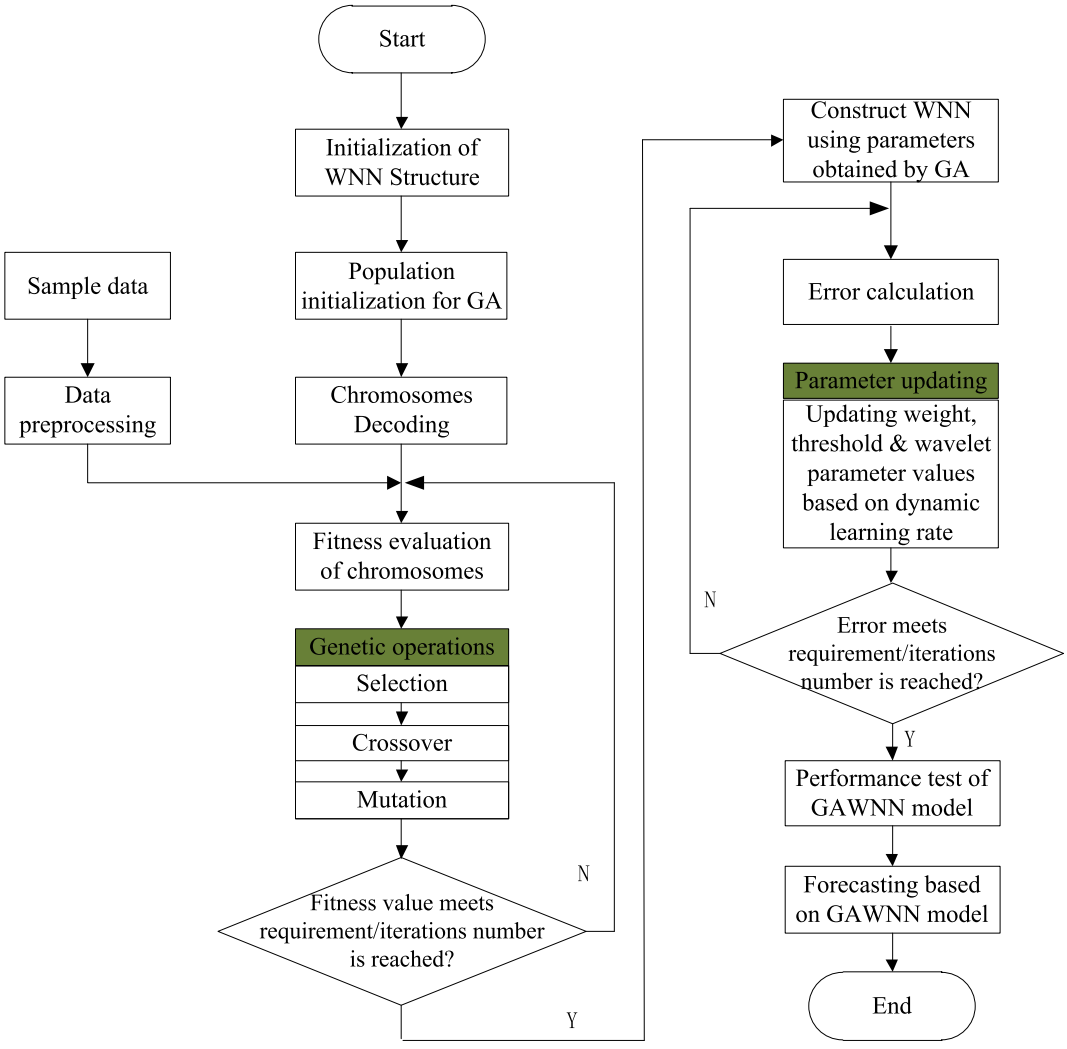


Fig. 2. The flowchart of GAWNN model.

$$Z_{kp} = \psi\left[\left(\sum_{i=1}^{m} w_{ik}x_{ip} - b_k\right) \Big/ a_k\right] \tag{12}$$

where ψ is the Morlet wavelet function, and $\psi(x) = \cos(1.75x)\exp\left(\frac{-x^2}{2}\right)$. For the p_{th} sample, calculate the output of the j neuron in the output layer through the following formula.

$$Z_{jp} = w_{j0} + \sum_{k=1}^{l} w_{jk}\left(\psi\left[\sum_{i=1}^{m}(w_{ik}x_{ip} - b_k) \Big/ a_k\right]\right) \tag{13}$$

where w_{j0} is the threshold value of the j_{th} neuron in the output layer.

Step 7: Update $w_{ik,}w_{kj}$ and a_k, b_k according to the following formula.

$$\begin{cases} w_{ik}(t+1) = w_{ik}(t) - \eta\dfrac{\partial E(t)}{\partial w_{ik}(t)} + \mu\Delta w_{ik}(t) \\ w_{kj}(t+1) = w_{kj}(t) - \eta\dfrac{\partial E(t)}{\partial w_{kj}(t)} + \mu\Delta w_{kj}(t) \\ a_k(t+1) = a_k(t) - \eta\dfrac{\partial E(t)}{\partial a_k(t)} + \mu\Delta a_k(t) \\ b_k(t+1) = b_k(t) - \eta\dfrac{\partial E(t)}{\partial b_k(t)} + \mu\Delta b_k(t) \end{cases} \tag{14}$$

Step 8: Training the network until finding a set of a, b, w which satisfy the following constrain: $E = \frac{1}{2}\sum_{p=1}^{P}\sum_{j=1}^{n}(y_{jp} - o_{jp})^2 < \xi$, where ξ is the pre-specified error, $O_p = [o_{1p}, \ldots, o_{np}]$ is the real value vector related to input sample X_p.

In the second stage of GAWNN model, in order to solve the problem of slow convergence speed of the standard BPNN, we use the following dynamic learning rate mechanism which includes the following three steps for updating the parameter values.

Step 1: In the training process of GAWNN model, if the total error denoted by ER_{t+1} in $(t+1)^{th}$ iteration is larger than that in t^{th}

iteration, and the increase rate of total error is larger than a pre-specified value ς (specified as 3% in this study), then the updating process is ignored and the learning rate is modified by $\eta_{t+1} = (1-\alpha)\eta_t$.

Step 2: If $ER_{t+1} < ER_t$, then update w_{ik}, w_{kj} and a_k, b_k. The learning rate is modified by $\eta_{t+1} = (1+\alpha)\eta_t$.

Step 3: If $ER_{t+1} > ER_t$, while the increase rate is less than ς, then update w_{ik}, w_{kj} and a_k, b_k, and keep the current learning rate value.

The learning rate updating process can be summarized as follows.

$$\eta_{t+1} = \begin{cases} (1+\alpha)\eta_t & ER_{t+1} \le ER_t \\ (1-\alpha)\eta_t & ER_{t+1} > ER_t(1+\varsigma) \\ \eta_t & ER_t < ER_{t+1}(t+1) \le ER_t(1+\varsigma) \end{cases} \tag{15}$$

According to the entropy function theory, for the minimization function, entropy function has faster convergence than mean square error. Thus, in the above learning rate updating process, the error function ER is set as follows.

$$ER = \sum_{p=1}^{P}\sum_{j=1}^{n}(z_{pj}\ln o_{pj} + (1-z_{pj})\ln(1-o_{pj})) \tag{16}$$

3. The VMD-PSR-GAWNN model

In this section, the hybrid VMD-PSR-GAWNN model is established based on the concept of "decomposition and ensemble" for multi-step ahead wind speed forecasting. The flowchart of the proposed model is shown in Fig. 3. The main structure of the hybrid model includes the following three steps:

Step 1: VMD is firstly utilized to decompose the actual wind speed series into a number of components with different frequencies which are respectively denoted by Mode1, Mode2, …, ModeN;

Step 2: The multi-step ahead forecasting for each component is conducted using GAWNN model in which the input-output sample pairs are determined by PSR technique.

Step 3: Finally, the ultimate forecast series of wind speed is obtained by aggregating the forecast result of each component.

The multi-step ahead forecasting scheme adopted in this paper is explained as follows. Since for different horizons, the sets of input-output sample pairs of GAWNN model determined by PSR technique are different, therefore, in order to simplify the explanation, we take the Mode2 generated by VMD in Case 1 as an example to illustrate the multi-step ahead forecasting scheme adopted in this paper. As mentioned in Section 4, two original wind speed series containing 2160 observations are adopted to evaluate the proposed forecasting model, and the 1^{st}-1440^{th} sampling points and the rest 720 sampling points are respectively adopted as the training set and testing set. As shown in Table 1, for Mode2, the parameters of τ and m determined by PSR are respectively 6 and 5. Based on the above information, the format of input-output sample pairs for conducting the multi-step ahead forecasting can be illustrated in Fig. 4.

Furthermore, as is shown in Fig. 4, we take six-step ahead forecasting as an example to clearly illustrate the multi-step ahead forecasting process. As shown in Fig. 4, in the training process, the input and output datasets of GAWNN model for six-step ahead forecasting are respectively $\{\{x_1, x_7, x_{13}, x_{19}, x_{25}\}, \{x_2, x_8, x_{14}, x_{20}, x_{26}\}, ..., \{x_{1410}, x_{1416}, x_{1422}, x_{1428}, x_{1434}\}\}$ and $\{x_{31}, x_{32}, ..., x_{1440}\}$, and in the testing process, the input and output datasets of GAWNN model are respectively $\{\{x_{1411}, x_{1417}, x_{1423}, x_{1429}, x_{1435}\}, \{x_{1412}, x_{1418}, x_{1424}, x_{1430}, x_{1436}\}, ..., \{x_{2128}, x_{2136}, x_{2142}, x_{2148}, x_{2154}\}\}$ and $\{x_{1441}, x_{1442}, ..., x_{2160}\}$.

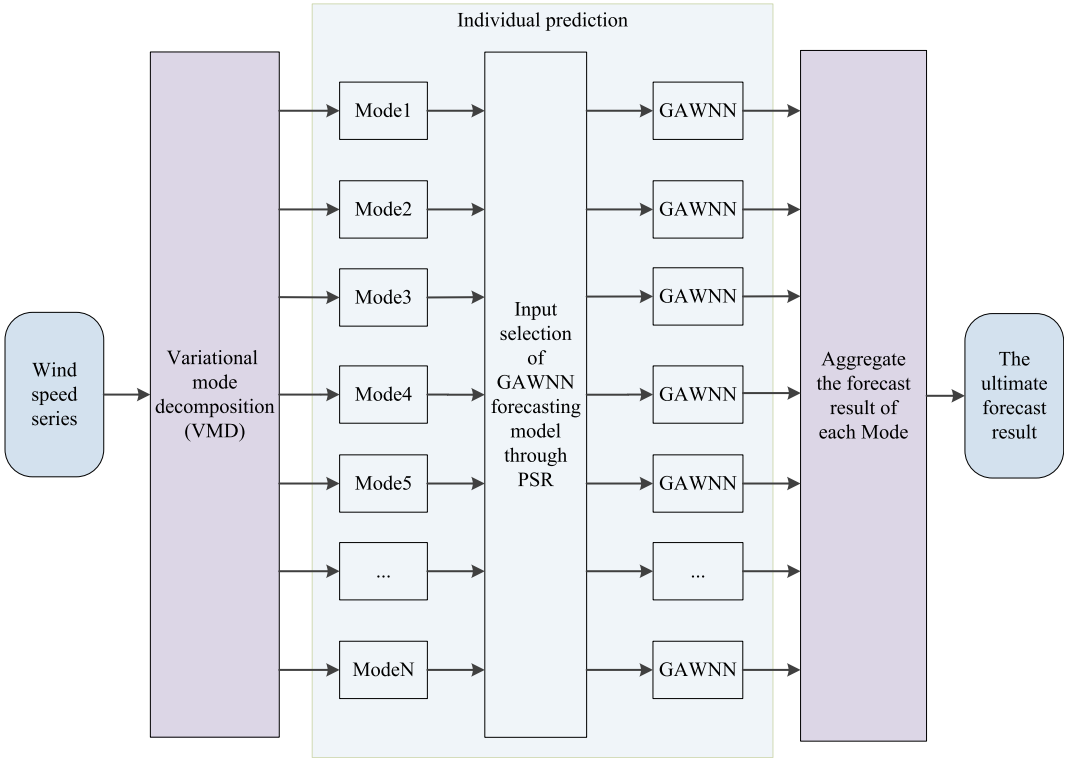


Fig. 3. The structure of VMD-PSR-GAWNN model.

Table 1
The parameters of PSR technique obtained by the C–C method for the Cases 1 and 2.

Case	Decomposed signals	Time window ϖ	Delay time τ	Embedding dimension m
Case 1	Mode1	24	18	3
	Mode2	21	6	5
	Mode3	16	3	7
	Mode4	18	3	7
	Mode5	19	4	6
	Mode6	19	4	6
	Mode7	21	3	8
	Mode8	24	3	9
	Mode9	18	5	5
Case 2	Mode1	24	16	3
	Mode2	21	7	4
	Mode3	20	3	8
	Mode4	17	3	7
	Mode5	18	3	7
	Mode6	20	4	7
	Mode7	18	4	6
	Mode8	21	5	6
	Mode9	19	4	6

4. Experimental results and comparative analysis

4.1. Experimental date description

This paper adopts two sets of hourly wind speed series collected respectively in spring (1st March 2014–29th May 2014) and autumn (1st September 2014–29th November 2014) seasons from a wind farm located in Xinjiang, China to evaluate the effectiveness of the proposed model, see Fig. 5. Due to the different climatic features in spring and autumn, these two data series have remarkably different fluctuation characteristics, and therefore, these two data series adopted in this paper can comprehensively and systematically evaluate the effectiveness and practicability of the proposed model. In addition, in each data set, the 1^{st}-1440^{th} (60 days) observations and 1441^{st}-2160^{th} (30 days) observations are respectively adopted as the training and testing sets. It should be noted that all of the prediction simulations are implemented in MATLAB R2010a.

4.2. Performance evaluation criteria

This paper adopts the following three popular performance measures to testify the efficiency of the proposed model: mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE). The performance measures of MAE, RMSE and MAPE are utilized to quantify the errors of forecast values, and the smaller they are, the better the prediction accuracy is. The computational formulas of these three performance measures are provided as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{x}(t) - x(t)| \tag{17}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{x}(t) - x(t))^2} \tag{18}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{x}(t) - x(t)}{x(t)}\right| \tag{19}$$

where n is the number of observed wind speed values, $\hat{x}(t)$ and $x(t)$ are respectively the forecast and observed values of wind speed at time t.

In addition, in order to compare the performances of different models, the persistence method illustrated in reference [35] is also adopted as a benchmark model, and the error difference between the persistence method and other forecasting models in this paper are calculated based on the following formula.

$$Dif = \frac{e_{per} - e_{other}}{e_{per}} \times 100\% \tag{20}$$

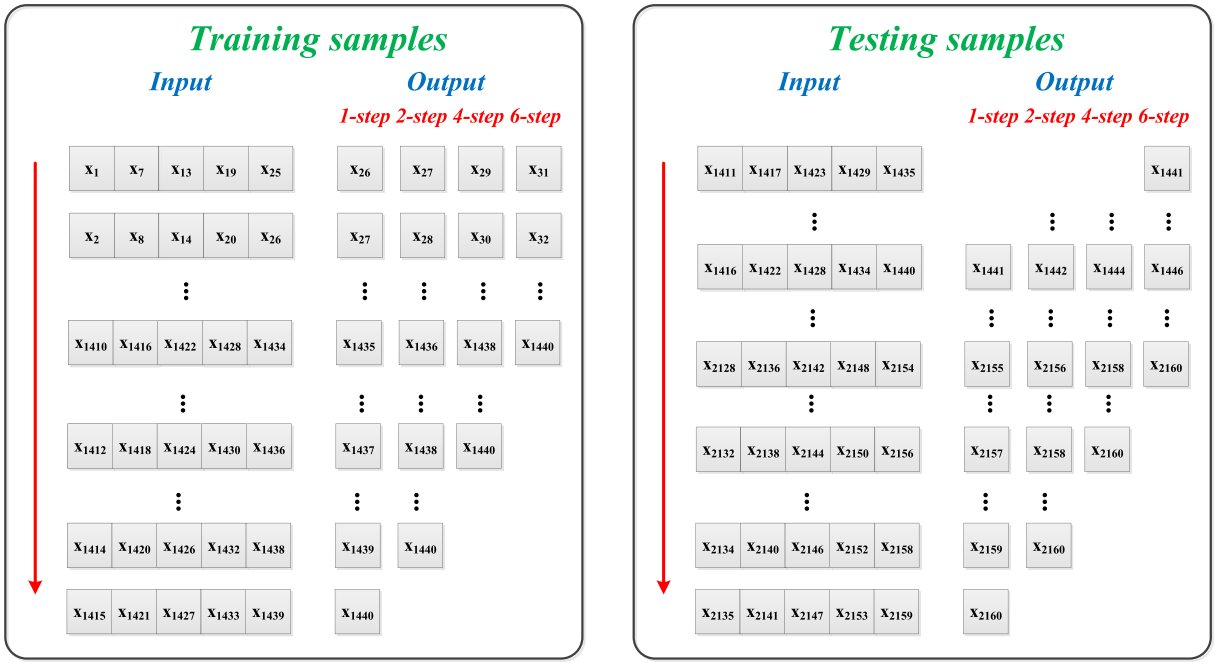


Fig. 4. The input and output data selection.

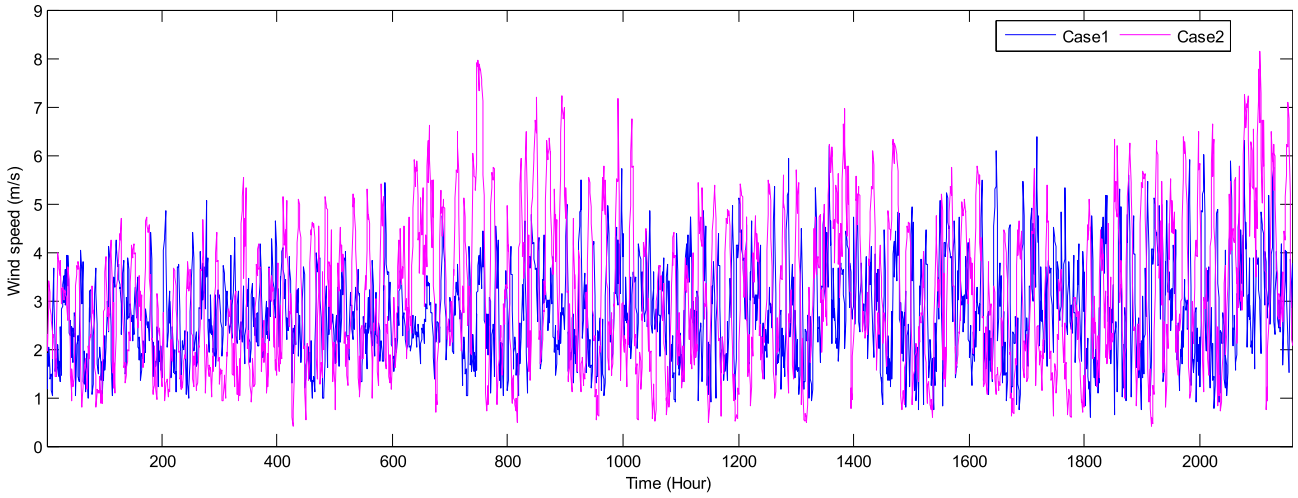


Fig. 5. Two original wind speed series.

where *Dif* represents the error difference between the persistence method and other forecasting models in the paper, e_{per} denotes the MAE, RMSE or MAPE value of the persistence method, and e_{other} represents the MAE, RMSE or MAPE value of other forecasting models in the paper.

4.3. Decomposed results by VMD

In order to improve the overall predicting performance, this paper firstly adopts VMD technique to respectively decompose the training sets (i.e., the first 1440 data points) of the two original wind speed series into a number of components, and the data decomposition results of these two wind speed series are listed in Figs. 6 and 7. It is obvious that each wind speed series is decomposed into 9 components which are respectively denoted by Mode1, Mode2, …, Mode9.

4.4. Parameter settings

After decomposition of the two wind speed series, the forecasting problem presented in this paper is transformed to the forecasting of each Mode generated by VMD. In this study, the GAWNN model is applied for the prediction of each Mode. As is shown in Figs. 6–7, each Mode owns 1440 data points which are applied for training the proposed forecasting model. The input-output sample pairs of GAWNN model are determined by PSR technique. The parameter settings of PSR technique including embedding dimension m and delay time τ are listed in Table 1. In this paper, GA algorithm is applied to optimize the parameters of WNN model for decreasing the randomness and instability of the

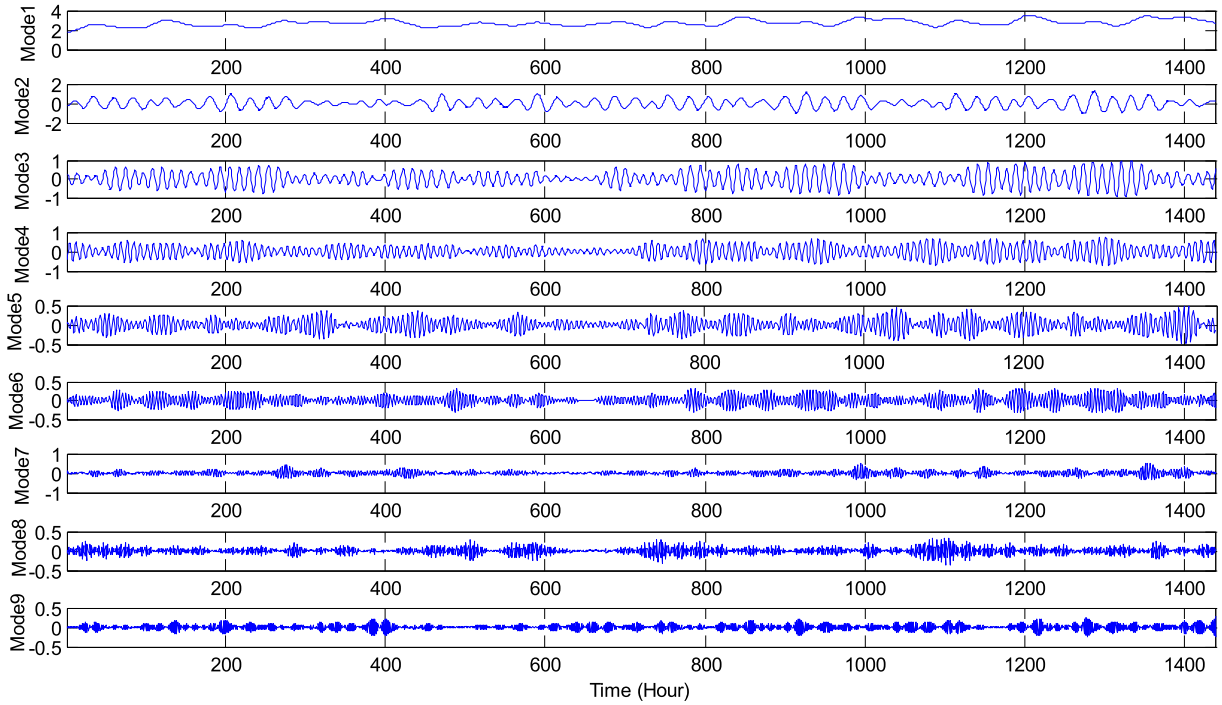


Fig. 6. The data decomposition results of wind speed series in Case1.

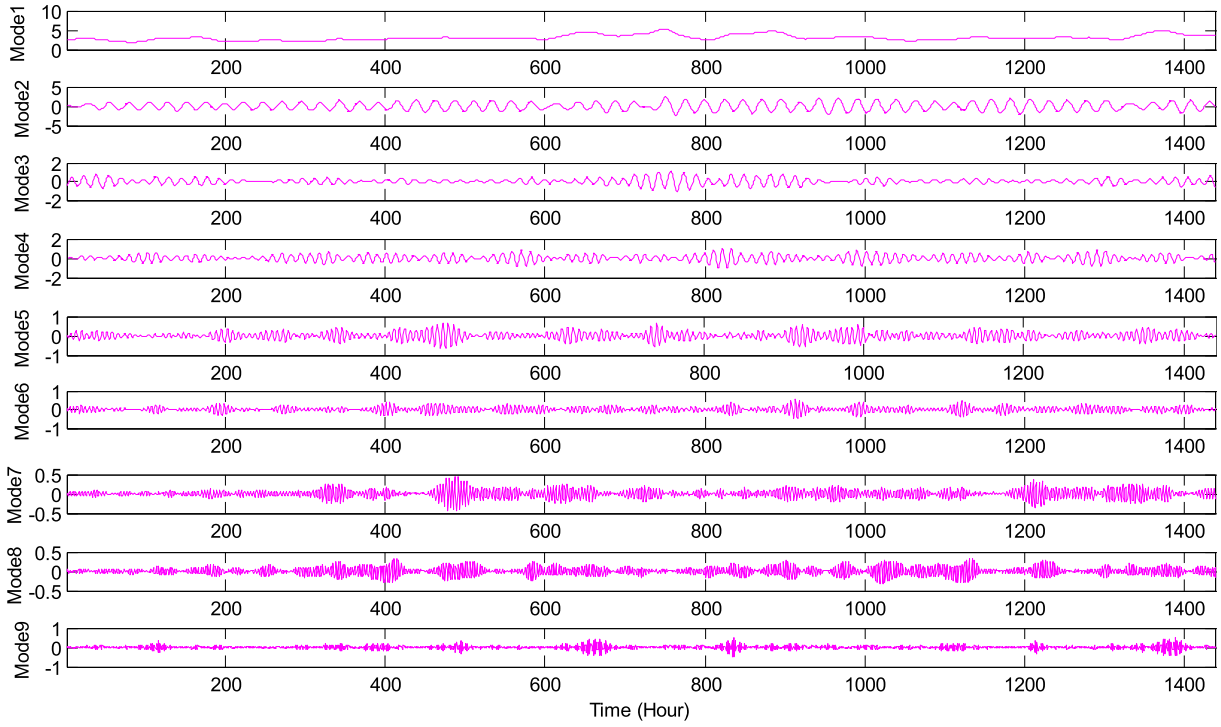


Fig. 7. The data decomposition results of wind speed series in Case2.

proposed model. The parameter settings of GA algorithm including the iterations (*Gen*), population size (*Pop*), crossover probability (p_c) and mutation probability (p_m) are listed as follows: $Gen = 100$, $Pop = 30$, $p_c = 0.7$ and $p_m = 0.1$.

4.5. Comparative analysis of different models

In order to demonstrate the advantages of VMD, the ensemble empirical mode decomposition (EEMD) proposed by Wu and Huang [36] in 2009 is adopted as a benchmark decomposition method. EEMD can overcome the mode mixing obstacle existed in empirical mode decomposition (EMD) method, and thus performs better then EMD in improving the overall forecast accuracy [37–39]. In this paper, the stopping iteration (SD_k) in EEMD is set as 0.2. In addition, to comprehensively evaluate the proposed VMD-PSR-GAWNN model, PSR-BPNN, PSR-WNN, PSR-GAWNN, EEMD-PSR-GAWNN models and persistence method are taken as the benchmark models.

4.5.1. Case 1

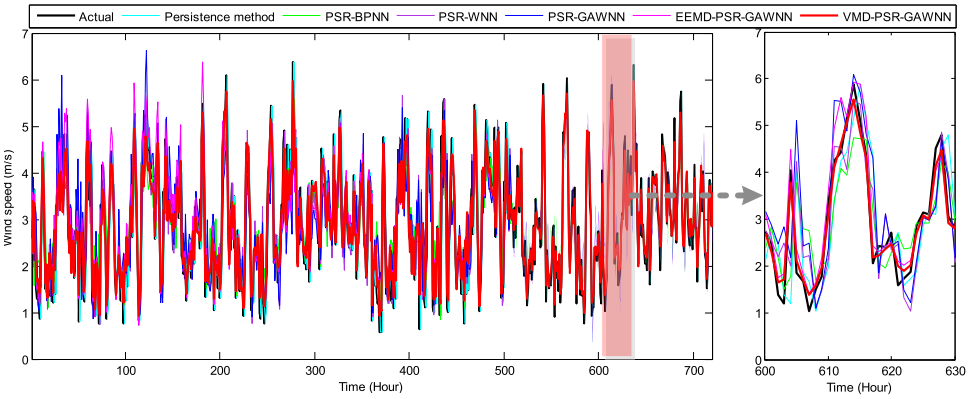
In Case 1, the original wind speed series is collected from spring (1st March 2014–29th May 2014) in a wind farm located in Xinjiang, China (see Fig. 5). The values of MAE, RMSE and MAPE of the proposed and comparison models are presented in Table 2 where the smallest value of each row is marked in boldface. As is shown in Table 2, the proposed VMD-PSR-GAWNN model owns the best performance compared with other comparison models. The conclusion can be further verified by the results presented in Fig. 8 which provides the forecasting performances of the proposed and other comparison models over horizons of one-step, two-step, four-step and six-step ahead. Therefore, based on the results shown in Table 2 and Fig. 8, it can be concluded that the proposed model outperforms all other comparison models for both one-step and multi-step ahead forecasting of wind speed in terms of the above three evaluation criteria.

In order to analyze the experimental results more clearly, the persistence method is adopted as a benchmark in this study, and the error difference between the persistence method and other

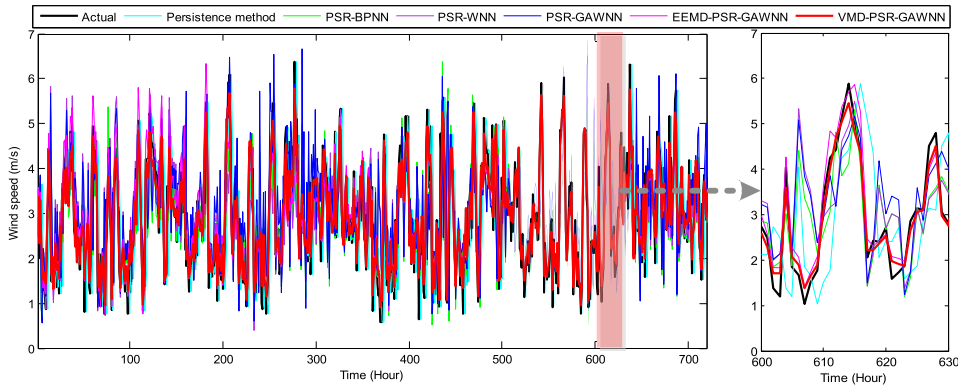
Table 2
Error comparison among different forecasting models in Case 1.

Prediction horizon	Index	Persistence method	PSR-BPNN	PSR-WNN	PSR-GAWNN	EEMD-PSR-GAWNN	VMD-PSR-GAWNN
1	MAPE(%)	31.422	28.564	26.901	26.132	13.714	**8.446**
	MAE	0.739	0.634	0.593	0.586	0.291	**0.187**
	RMSE	0.950	0.801	0.764	0.751	0.387	**0.235**
2	MAPE(%)	42.067	31.517	28.935	27.961	16.159	**9.133**
	MAE	0.985	0.729	0.651	0.629	0.366	**0.204**
	RMSE	1.285	0.967	0.833	0.782	0.437	**0.260**
4	MAPE(%)	54.329	37.401	33.627	32.593	18.426	**10.702**
	MAE	1.274	0.795	0.765	0.736	0.392	**0.238**
	RMSE	1.633	1.051	0.984	0.978	0.454	**0.301**
6	MAPE(%)	58.534	40.296	36.504	35.112	23.573	**12.022**
	MAE	1.380	0.913	0.788	0.761	0.462	**0.269**
	RMSE	1.716	1.101	1.063	1.018	0.603	**0.340**

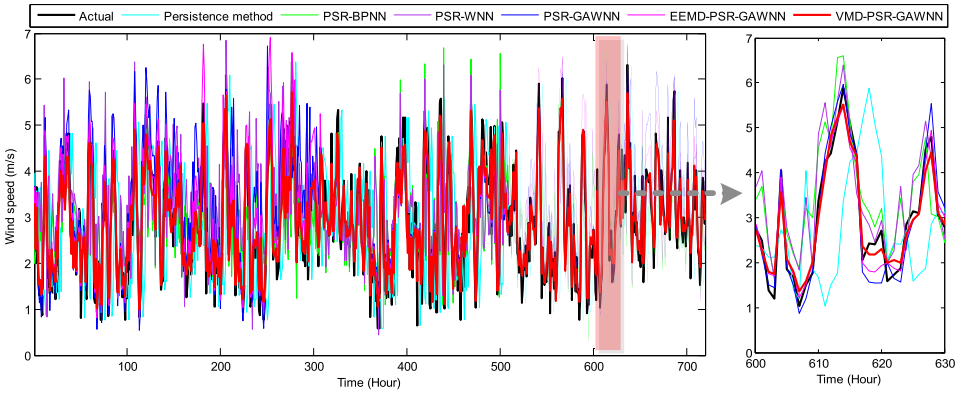
Note: The smallest value of each row is marked in boldface.



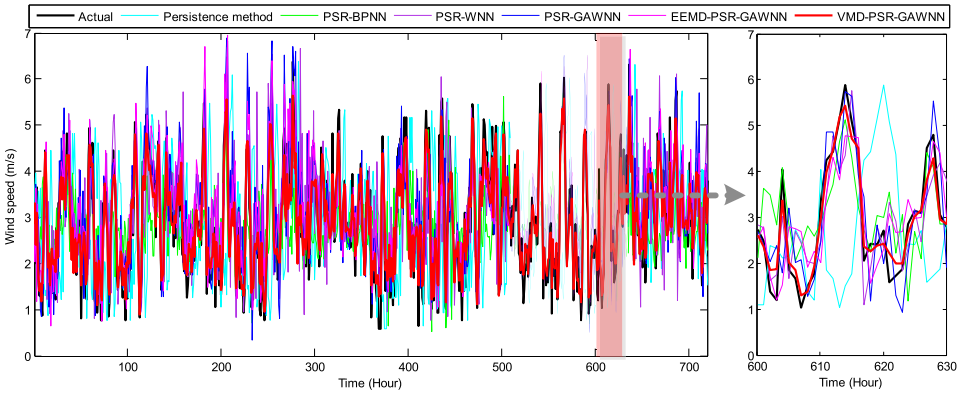
(a) One-step ahead forecasting.



(b) Two-step ahead forecasting.



(c) Four-step ahead forecasting.



(d) Six-step ahead forecasting.

Fig. 8. Multi-step ahead wind speed forecasting results in Case 1.

Table 3
Error reductions in comparison with the persistent model in Case 1.

Prediction horizon	Index	*Dif*				
		PSR-BPNN	PSR-WNN	PSR-GAWNN	EEMD-PSR-GAWNN	VMD-PSR-GAWNN
1	MAPE	9.10	14.39	16.84	56.36	**73.12**
	MAE	14.21	19.76	20.70	60.62	**74.70**
	RMSE	15.68	19.58	20.95	59.26	**75.26**
2	MAPE	25.08	31.22	33.53	61.59	**78.29**
	MAE	25.99	33.91	36.14	62.84	**79.29**
	RMSE	24.75	35.18	39.14	65.99	**79.77**
4	MAPE	31.16	38.10	40.01	66.08	**80.30**
	MAE	37.60	39.95	42.23	69.23	**81.32**
	RMSE	35.64	39.74	40.11	72.20	**81.57**
6	MAPE	31.16	37.64	40.01	59.73	**79.46**
	MAE	33.84	42.90	44.86	66.52	**80.51**
	RMSE	35.84	38.05	40.68	64.86	**80.19**

Note: The largest value of each row is marked in boldface.

forecasting models are calculated and presented in Table 3 where the largest value of each row is marked in boldface. From Table 3, it is obvious that the proposed VMD-PSR-GAWNN model obtains the largest improvement compared with other forecasting models, which indicates that the proposed model performs better than the persistence method and other comparison models.

In order to further investigate the influence of VMD technique, GA algorithm and wavelet theory on the proposed model, four following categories of comparisons are carried out in this experiment. The first category of comparison (*Comparison I*) is designed for testifying the advantages of VMD decomposition technique, and is conducted between the proposed VMD-PSR-GAWNN model and EEMD-PSR-GAWNN model. The second category of comparison (*Comparison II*) is designed for proving the positive effects of decomposition techniques, and is performed between the forecasting models embedded with decomposition techniques (VMD-PSR-GAWNN model and EEMD-PSR-GAWNN model) and the forecasting models without any decomposition techniques (PSR-GAWNN model). The third category of comparison (*Comparison III*) is designed for verifying the contribution of GA optimization algorithm on WNN model, and is conducted between the PSR-GAWNN model and PSR-WNN model. The fourth category of comparison (*Comparison IV*) is designed for demonstrating the positive effect of wavelet theory on ANN, and is conducted between the PSR-WNN model and PSR-BPNN model. The above four comparisons are conducted based on the forecast results shown in Tables 2 and 3

In *Comparison I*, through comparing VMD-PSR-GAWNN model with EEMD-PSR-GAWNN model, it is obvious that the forecast accuracy of VMD-PSR-GAWNN is superior to that of EEMD-PSR-GAWNN model for both one-step and multi-step ahead wind speed forecasting. Thus it can be concluded that the VMD decomposition technique is more effective than EEMD for improving the overall forecast accuracy. The reasons are mainly twofold: (1) VMD searches for a number of modes and their respective center frequencies, such that the band-limited modes reproduce the input signal exactly or in least-squares sense, thus VMD has the ability to separate components of similar frequencies compared with EEMD; (2) VMD is more robust to noisy data such as wind speed, $PM_{2.5}$ concentration and electricity price. Indeed, since each mode is updated by Wiener filtering in Fourier domain during the optimization process, the updated mode is less affected by noisy disturbances, and therefore VMD can be more efficiently for capturing the signal's short and long variations than EEMD [40]. In *Comparison II*, through comparing the forecasting models embedded with decomposition techniques (VMD-PSR-GAWNN model and EEMD-PSR-GAWNN model) and the forecasting model without any decomposition technique (PSR-GAWNN model), it is obvious that the decomposition techniques (VMD and EEMD) can effectively improve the overall forecasting ability of GAWNN model. In *Comparison III*, through comparing the PSR-GAWNN model and PSR-WNN model, it is obvious that GA algorithm has a positive effect on improving the forecasting performance of ANN. In *Comparison IV*, through comparing the PSR-WNN model and PSR-BPNN model, it is clear that WNN model has a superior performance compared with BPNN model, which illustrates that the forecasting ability of BPNN model is improved by integrating the wavelet theory.

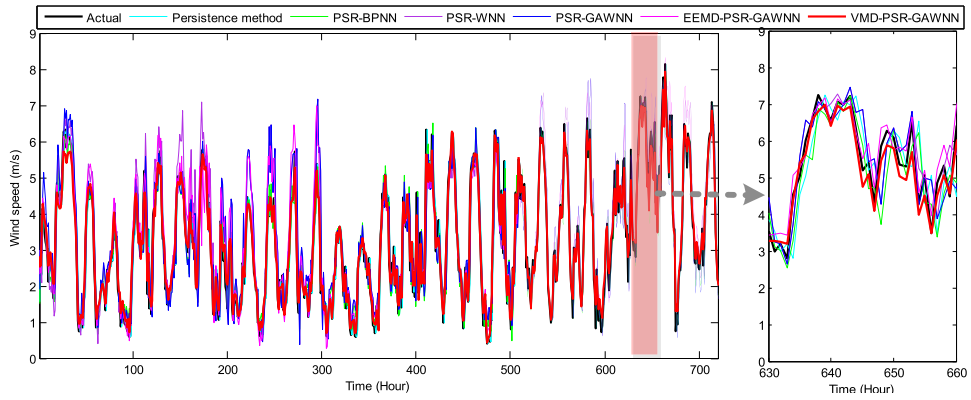
4.5.2. Case 2

In order to further confirm the advantages of the proposed forecasting model, Case 2 is conducted in this section. In Case 2, the original wind speed series is collected in autumn (1st September 2014–29th November 2014) from the same wind farm as in Case 1 (see Fig. 5). As is shown in Fig. 5, it is obvious that the original wind speed series in Case 2 fluctuates more than the data series in Case 1. The forecast results of all forecasting model adopted in this paper are presented in Fig. 9 and Tables 4 and 5. According to the forecast results of Case 2, similar conclusions to Case 1 can be obtained.

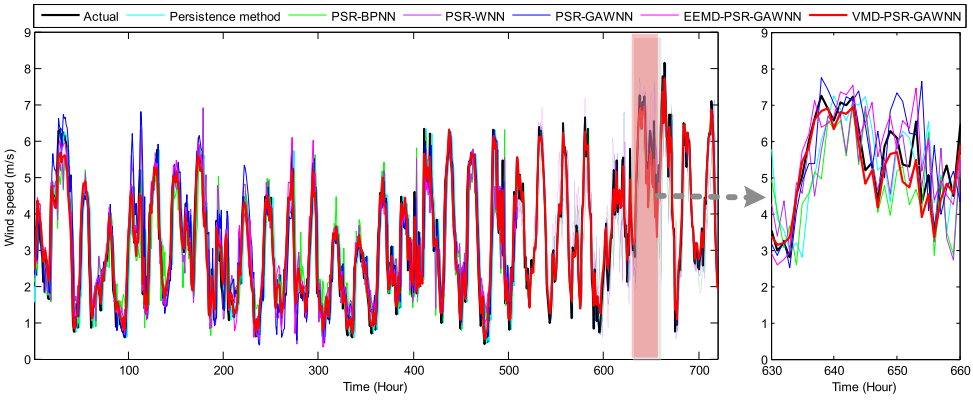
5. Conclusion

Accurate wind speed forecasting plays a vital role in power system operation, power grid security and electricity market. However, due to the intermittent and unstable nature of wind speed, it is very difficult to predict it accurately. This paper proposes a novel hybrid model (VMD-PSR-GAWNN) for multi-step ahead wind speed forecasting. In this forecasting model, VMD is firstly applied to decompose the original wind speed series into a number of independent components in order to improve the overall forecast accuracy. Then each component generated by VMD is forecasted using GAWNN model where the input-output sample pairs are determined by PSR technique. Finally, the ultimate forecast result of wind speed is obtained by adding up the forecast result of each component. Two real-world hourly wind speed series are adopted to testify the effectiveness of the proposed model. The experimental results indicate that the proposed model outperforms all other forecasting models including persistence method, PSR-BPNN, PSR-WNN, PSR-GAWNN and EEMD-PSR-GAWNN models in terms of MAE, RMSE and MAPE, which demonstrates that the proposed forecasting model is reliable and effective for both one-step and multi-step ahead wind speed forecasting.

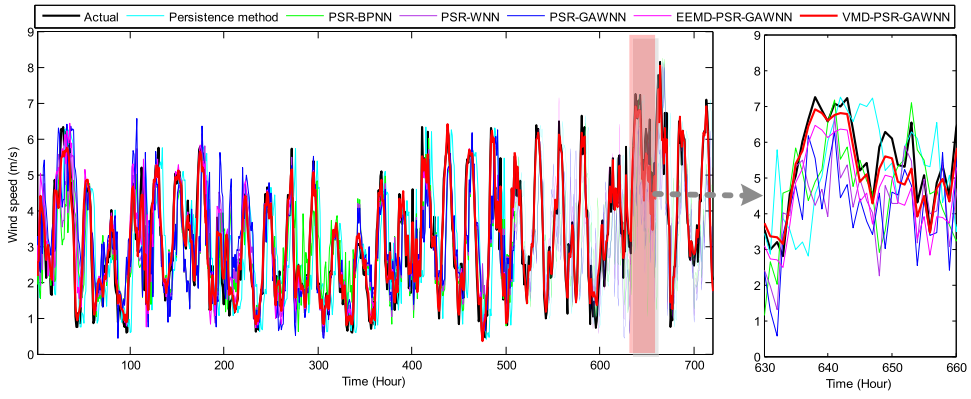
The purpose of this paper is to propose a precise forecasting model for multi-step ahead wind speed forecasting and verify its performance against other benchmark models. Based on the



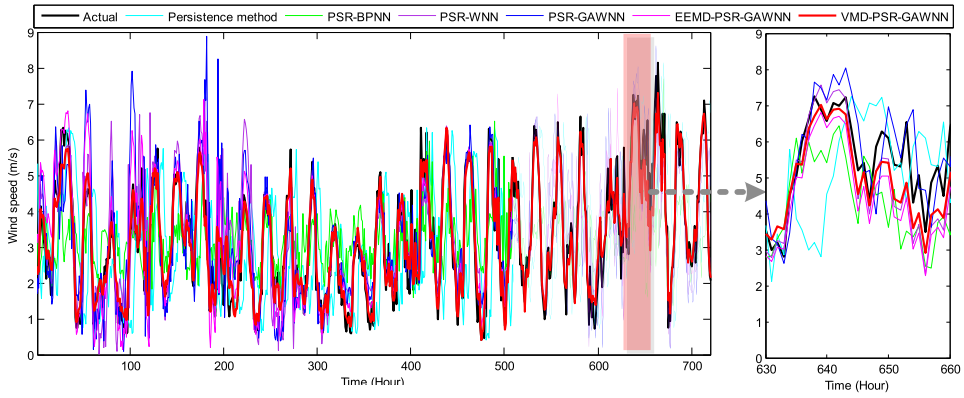
(a) One-step ahead forecasting.



(b) Two-step ahead forecasting.



(c) Four-step ahead forecasting.



(d) Six-step ahead forecasting.

Fig. 9. Multi-step ahead wind speed forecasting results in Case 2.

Table 4
Error comparison among different forecasting models in Case 2.

Prediction horizon	Index	Persistence method	PSR-BPNN	PSR-WNN	PSR-GAWNN	EEMD-PSR-GAWNN	VMD-PSR-GAWNN
1	MAPE(%)	22.408	21.779	21.174	19.628	12.874	**7.537**
	MAE	0.576	0.540	0.533	0.496	0.307	**0.180**
	RMSE	0.765	0.714	0.692	0.638	0.425	**0.234**
2	MAPE(%)	33.678	32.352	30.578	27.956	14.624	**8.469**
	MAE	0.850	0.772	0.749	0.689	0.384	**0.199**
	RMSE	1.091	1.001	0.916	0.807	0.481	**0.260**
4	MAPE(%)	52.583	42.825	38.156	36.153	17.152	**10.058**
	MAE	1.302	1.063	1.004	0.914	0.463	**0.234**
	RMSE	1.634	1.394	1.287	1.168	0.575	**0.299**
6	MAPE(%)	69.374	55.823	51.384	47.535	20.167	**11.441**
	MAE	1.663	1.219	1.187	1.139	0.521	**0.276**
	RMSE	2.040	1.501	1.363	1.352	0.662	**0.368**

Note: The smallest value of each row is marked in boldface.

Table 5
Error reductions in comparison with the persistent model in Case 2.

Prediction horizon	Index	*Dif*				
		PSR-BPNN	PSR-WNN	PSR-GAWNN	EEMD-PSR-GAWNN	VMD-PSR-GAWNN
1	MAPE	2.81	5.51	12.41	42.55	**66.36**
	MAE	6.25	7.47	13.89	46.70	**68.75**
	RMSE	6.67	9.54	16.60	44.44	**69.41**
2	MAPE	3.94	9.20	16.99	56.58	**74.85**
	MAE	9.18	11.88	18.94	54.82	**76.59**
	RMSE	8.25	16.04	26.03	55.91	**76.17**
4	MAPE	18.56	27.44	31.25	67.38	**80.87**
	MAE	18.36	22.89	29.80	64.44	**82.03**
	RMSE	14.69	21.24	28.52	64.81	**81.70**
6	MAPE	19.53	25.93	31.48	70.93	**83.51**
	MAE	26.70	28.62	31.51	68.67	**83.40**
	RMSE	26.42	33.19	33.73	67.55	**81.96**

Note: The largest value of each row is marked in boldface.

experimental results of Cases 1 and 2, it is obvious that the adopted two wind speed series with 2160 observations are enough to confirm the good forecast accuracy of our proposed model. However, it should be noted that the length of wind speed series is one of the most important uncertain factors influencing the performance of the proposed forecasting model, therefore, it is not appropriate to compare the proposed model's performance with other models that use much more complete datasets.

As mentioned above, the intermittent and unstable nature of wind energy makes the wind speed forecasting become a very difficult task. Therefore, there are still several research directions left for the future. For example, some meteorological factors such as atmospheric pressure, temperature, and precipitation may be integrated into the forecasting model in order to improve the forecast accuracy.

Acknowledgment

The authors would like to thank the editor and the two anonymous reviewers for their constructive comments on improving an early version of this paper. This research was supported by the National Natural Science Foundation of China (Grant No. 71301153); the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry of China; the Science Foundation of Mineral Resource Strategy and Policy Research Center, China University of Geosciences (Grant No. H2017011B); the Natural Science Foundation of Hubei Province (Grant No. 2015CFB497).

References

[1] A. Kusiak, Z. Zhang, A. Verma, Prediction, operations, and condition monitoring in wind energy, Energy 60 (1) (2013) 1–12.
[2] H.B. Azad, S. Mekhilef, V.G. Ganapathy, Long-term wind speed forecasting and general pattern recognition using neural networks, IEEE Trans. Sustain. 5 (2) (2014) 546–553.
[3] J.Z. Wang, Y. Wang, P. Jiang, The study and application of a novel hybrid forecasting model-A case study of wind speed forecasting in China, Appl. Energy 143 (2015) 472–488.
[4] H. Liu, H.Q. Tian, Y.F. Li, Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction, Appl. Energy 98 (2012) 415–424.
[5] C. Zhang, H. Wei, X. Zhao, T. Liu, K.A. Zhang, Gaussian process regression based hybrid approach for short-term wind speed prediction, Energy Convers. Manag. 126 (2016) 1084–1092.
[6] S.S. Soman, H. Zareipour, O. Malik, P. Mandal, A review of wind power and wind speed forecasting methods with different time horizons, in: North American Power Symposium (NAPS), IEEE, 2010, pp. 1–8.
[7] L.Y. Xiao, F. Qian, W. Shao, Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm, Energy Convers. Manag. 143 (2017) 410–430.
[8] U. Schlink, G. Tetzlaff, Wind speed forecasting from 1 to 30 minutes, Theor. Appl. Climatol. 60 (1998) 191–198.
[9] M. Lydia, S.S. Kumar, A.I. Selvakumar, G.E.P. Kumar, Linear and non-linear autoregressive models for short-term wind speed forecasting, Energy Convers. Manag. 112 (2016) 115–124.
[10] H. Liu, E. Erdem, J. Shi, Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed, Appl. Energy 88 (3) (2011) 724–732.
[11] G. Li, J. Shi, J.Y. Zhou, Bayesian adaptive combination of short-term wind speed forecasts from neural network models, Renew. Energy 36 (1) (2011) 352–359.
[12] Z. Song, Y. Jiang, Z.J. Zhang, Short-term wind speed forecasting with Markov-switching model, Appl. Energy 130 (5) (2014) 103–112.
[13] C.D. Zuluaga, M.A. Alvarez, E. Giraldo, Short-term wind speed prediction based on robust Kalman filtering: an experimental comparison, Appl. Energy 156 (2015) 321–330.
[14] Z.H. Guo, J. Wu, H.Y. Lu, J.Z. Wang, A case study on a hybrid wind speed forecasting method using BP neural network, Knowl. Base Syst. 24 (2011)

1048–1056.
[15] G.W. Chang, H.J. Lu, Y.R. Chang, Y.D. Lee, An improved neural network-based approach for short-term wind speed and power forecast, Renew. energy 105 (2017) 301–311.
[16] J. Wang, S. Qin, Q. Zhou, H. Jiang, Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China, Renew. energy 76 (2014) 91–101.
[17] L. Thiaw, G. Sow, S.S. Fall, M. Kasse, E. Sylla, S. Thioye, A neural network based approach for wind resource and wind generators production assessment, Appl. Energy 87 (5) (2010) 1744–1748.
[18] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, P.S. Dokopoulos, A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation, IEEE Trans. Energy Convers. 19 (2) (2004) 352–361.
[19] L. Thiaw, G. Sow, S.S. Fall, M. Kasse, E. Sylla, S. Thioye, A neural network based approach for wind resource and wind generators production assessment, Appl. Energy 87 (5) (2010) 1744–1748.
[20] Q.H. Hu, S.G. Zhang, Z.X. Xie, J.S. Mi, J. Wan, Noise model based ν-support vector regression with its application to short-term wind speed forecasting, Neural Netw. 57 (2014) 1–11.
[21] O.B. Shukur, M.H. Lee, Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA, Renew. Energy 76 (2015) 637–647.
[22] L. Tang, L. Yu, S. Wang, J.P. Li, S.Y. Wang, A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting, Appl. Energy 93 (5) (2012) 432–443.
[23] H. Liu, H. Tian, X. Liang, Y. Li, New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, mind evolutionary algorithm and artificial neural networks, Renew. Energy 83 (2015) 1066–1075.
[24] S.X. Wang, N. Zhang, L. Wu, Y.M. Wang, Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method, Renew. Energy 94 (2016) 629–636.
[25] A.B. Meng, J.F. Ge, S.Z. Chen, Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm, Energy Convers. Manag. 114 (2016) 75–88.
[26] D.Y. Wang, H.Y. Luo, O. Grunder, Y.B. Lin, H.X. Guo, Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm, Appl. Energy 190 (2017) 390–407.
[27] X.Y. Yang, X. Yang, S.Y. Chen, Wind speed and generated power forecasting in wind farm, Proc. Chin. Soc. Electr. Eng. 25 (11) (2005) 1–5.
[28] Z.H. Guo, J. Wu, H.Y. Lu, J.Z. Wang, A case study on a hybrid wind speed forecasting model using BP neural network, Knowl. Base Syst. 24 (2011) 1048–1056.
[29] A.A. Abdoos, A new intelligent method based on combination of VMD and ELM for short term wind power forecasting, Neurocomputing 203 (2016) 111–120.
[30] S.A. Lahmiri, Variational mode decompoisition approach for analysis and forecasting of economic and financial time series, Expert Syst. Appl. 55 (2016) 268–273.
[31] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2014) 531–544.
[32] F. Takens, Detecting strange attractors in turbulence, vol. 898, Springer Berlin Heidelberg, 1981, pp. 366–381.
[33] M. Kucuk, N. Agiralioglu, Wavelet regression techniques for stream flow predictions, J. Appl. Stat. 33 (9) (2006) 943–960.
[34] J.H. Holland, Adaptation in natural and artificial systems, University of Michigan Press, 1975.
[35] Z.Y. Su, J.Z. Wang, H.Y. Lu, G. Zhao, A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting, Energy Convers. Manag. 85 (2014) 443–452.
[36] Z.H. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (1) (2008) 1–41.
[37] W.C. Wang, K.W. Chau, L. Qiu, Y.B. Chen, Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition, Environ. Res. 139 (2015) 46–54.
[38] J.L. Zhang, Y.J. Zhang, L. Zhang, A novel hybrid method for crude oil price forecasting, Energy Econ. 49 (2015) 649–659.
[39] X. Jiang, L. Zhang, X. Chen, Short-term forecasting of high-speed rail demand: a hybrid approach combining ensemble empirical mode decomposition and gray support vector machine with real-world applications in China, Transp. Res. Part C Emerg. Technol. 44 (4) (2014) 110–127.
[40] S. Lahmiri, A variational mode decompoisition approach for analysis and forecasting of economic and financial time series, Expert Syst. Appl. 55 (2016) 268–273.