

Short-Term Wind Speed Prediction Using Signal Preprocessing Technique and Evolutionary Support Vector Regression

Jujie Wang¹ · Yaning Li²

Published online: 28 November 2017
© Springer Science+Business Media, LLC, part of Springer Nature 2017

Abstract Short-term wind speed prediction is beneficial to guarantee the safety of wind power utilization and reduce the cost of wind power generation. As a kind of the powerful artificial intelligent algorithms, support vector regression (SVR) has been successfully employed in solving forecasting problems. However, due to the intrinsic complexity and multi-patterns of wind speed fluctuations, it is regarded as one of the most challenging applications for wind speed prediction. To alleviate the influence of complexity and capture these different patterns, this study proposes a novel approach named SIE–WDA–GA–SVR for short-term wind speed prediction, which applies the seasonal information extraction (SIE) and wavelet decomposition algorithm (WDA) into hybrid model that integrates the genetic algorithm (GA) into SVR. First, the proposed approach uses SIE to decompose the original wind speed into seasonal and trend components, and the seasonal indices are calculated by SIE. Second, the proposed approach uses WDA to decompose the trend component into both the approximate and the detailed scales. Third, the proposed approach uses GA–SVR to forecast the approximated and detailed scales, respectively. Then, the prediction values of the trend component can be obtained by integrating the prediction values of the approximated scale into the prediction values of the detailed scale. By integrating the seasonal indices into the prediction values of trend component, we can obtain the final forecasting results of the original wind speed. Moreover, the partial autocorrelation function is used to determine the number of input dimension for the SVR, and the GA is used to select the parameters of the SVR. Four real wind speed datasets are used as test samples to verify the proposed approach. Experimental results indicate that the proposed approach outperforms other benchmark models in four statistical error measures, and can improve the forecasting accuracy of wind speed.

✉ Jujie Wang
wjj0913@126.com

¹ Climate and Weather Disasters Collaborative Innovation Center, School of Economics and Management, Nanjing University of Information Science and Technology, Nanjing 210044, China

² College of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China

Keywords Wind speed prediction · Seasonal information extraction · Wavelet decomposition algorithm · Support vector regression · Hybrid model

1 Introduction

As one kind of the most promising green energy, wind power is rapidly growing around the world [1]. In the past 3 years, it has been doubled for the accumulative installed capacity of wind energy generation. It is estimated that in 2020, approximately 12% of total world electricity demands will be supplied from wind power [2]. With the large scale penetration of wind energy in the power grids, a number of challenges have been posed because of the intermittent and stochastic natures of wind speed fluctuations [3,4]. These uncertainties of wind speed fluctuations can put the system reliability and power quality at risk [5]. In order to deal with the problem and improve the utilization efficiency of wind energy, accurate wind speed prediction is indispensable [6].

In general, there are two main categories for wind speed prediction in the time scales: short-term prediction and long-term prediction [7]. The time scales of short-term prediction are minutes, hours and days, and the time scales of long-term prediction are months and years. It is very important to improve the short-term prediction precision for guaranteeing the safety of power grid and reduce the cost of wind power generation [8]. It is beneficial to improve the long-term prediction precision for planning the windmills [9]. In this study, we mainly focus on the prediction issue of short-term wind speed.

Extensive efforts have been devoted for enhancing the prediction ability of short-term wind speed in recent years. According to existing literatures, there are two main categories for the short-term wind speed forecasting [10–12]. The first category can be referred to conventional statistical models, such as autoregressive moving average model (ARMA), Kalman filter, stochastic model, Markov chain, and so on [13–17]. These approaches try to use the historical wind speed data to construct prediction models, and the statistical regularities of wind speed fluctuations are described. They have been successfully used to predict the short-term wind speed. For example, Lalarukh and Yasmin used the ARMA model to predict the short-term wind speed in Quetta, Pakistan [13]. Louka et al. [14] used the Kalman filter technique to predict the short-term wind speed. Bivona et al. [16] presented a stochastic model to forecast the short-term wind speed. Shamshad et al. [17] presented the Markov chain models to forecast the short-term wind speed. All empirical results indicated that these conventional statistical models are suitable for forecasting the short-term wind speed. However, the prediction performance of these models will be worse if the nonlinear features of wind speed fluctuations are obvious. In other words, these statistical models may be insufficient to capture the hidden nonlinear features in wind speed.

To overcome the drawbacks of the conventional statistical models, artificial intelligence (AI) techniques with powerful nonlinear self-learning capacities, including artificial neural networks (ANNs), support vector regressions (SVRs), fuzzy logic methods, etc., have become increasingly popular for wind speed forecasting [18–23]. For example, Cadenas and Rivera [18] used the ANN to forecast the wind speed in the region of La Venta, Oaxaca, Mexico, and the empirical results indicated that the ANN could increase the prediction ability of short-term wind speed. Li and Shi [19] applied three ANNs to the hourly wind speed forecasting. In Ref. [20], Flores et al. used a back-propagation neural network (BPNN) to forecast the wind speed, and the conclusion indicated that the BPNN could increase the accuracy of wind speed forecasting. Zhou et al. [21] applied the SVR with fine tuning parameters to forecast

the wind speed, and the conclusion indicated that the proposed model outperformed the persistence model for predicting the short-term wind speed. Hu et al. [22] presented a control algorithm based on v -support vector regression and augmented Lagrange multiplier method for wind speed forecasting, and the results indicated that the proposed control algorithm could increase the forecasting effectiveness. Kavousi-Fard et al. [23] proposed a fuzzy-based prediction interval for wind power prediction. More introductions to AI methods can be found in [24–31].

Even though AI methods (e.g., ANN, SVR, and fuzzy logic) provide a great deal of promise, they also suffer from a number of shortcomings such as the time wasting, slow convergence, local minima, and the risk of model over-fitting [32,33]. In order to overcome these drawbacks, intelligent optimization algorithms, mainly including genetic algorithm (GA) and particle swarm optimization (PSO), have been successfully used to optimize the model parameters of AI and enhance the prediction performance of them [34–36]. For instance, Hu et al. proposed a short-term traffic flow hybrid forecasting method based on PSO and SVR. Experimental results showed that the hybrid method could get accurate forecasting results than individual models [34]. Cao and Parry [35] examined the relative effectiveness of hybrid model based on ANN and GA in forecasting future earnings per share. Gu et al. [36] proposed a housing price forecasting based on GA and SVR.

As a climate-driven renewable resource, the seasonal variations and trend variations of wind speed fluctuations are two most commonly encountered phenomena. Generally speaking, two kinds of variations in wind speed are mutual penetration, and the seasonal information of wind speed is often neglected in the most existing researches. The phenomena will cause large deviation in wind speed prediction. According to Zhang and Qi [37], seasonal information extraction (SIE) can extremely reduce the prediction error in many seasonal time series. The SIE can decompose the seasonal time series into seasonal and trend components, and can help extract the seasonal information and make forecasting more efficient. Therefore, In order to increase the accuracy of wind speed forecasting, this study uses the SIE technique to extract the seasonal information from wind speed.

On the other hand, the multi-patterns of wind speed fluctuations are the other commonly encountered phenomena because of wind being a climate-driven green energy which is influenced by many meteorological parameters. Thus, accurate wind speed forecasting is a difficult task if these forecasting models are directly constructed by the original wind speed. In order to improve the predictive accuracy, it is necessary to consider and analyze the multi-patterns characteristics of wind speed fluctuations. Thus, a multi-patterns decomposition technique is indispensable to construct a suitable wind speed prediction model [38]. As a relatively novel multi-patterns decomposition technique, wavelet decomposition algorithm (WDA) can decompose a complicated multi-patterns signal into an approximate part associated with low frequency and a detailed part associated with high frequency, which can show the local and global dynamic properties of a signal at specific timescales [39,40]. It has been widely applied in the prediction issues [41]. Thus, this study tends to construct a suitable forecasting model for wind speed using WDA.

Based on the above consideration, this study proposes a new method named SIE–WDA–GA–SVR for forecasting the short-term wind speed, which applies the SIE and WDA into hybrid model that integrates the GA into the SVR. First, the proposed approach uses SIE to decompose the original wind speed into seasonal and trend components, and the seasonal indices are calculated by SIE. Second, the proposed approach uses WDA to decompose the trend component into both the approximate and the detailed scales. Third, the proposed approach uses GA–SVR to forecast the approximated and detailed scales, respectively. Then, the prediction values of the trend component can be obtained by integrating the prediction

values of the approximated scale into the prediction values of the detailed scale. By integrating the seasonal indices into the prediction values of trend component, we can obtain the final forecasting results of the original wind speed. Moreover, the partial autocorrelation function (PACF) is used to determine the number of input dimension for the SVR, and the genetic algorithm (GA) is used to select the parameters of the SVR. Four real wind speed datasets are used as test samples to verify the proposed approach. Experimental results indicate that the proposed SIE–WDA–GA–SVR model outperforms other benchmark models in four statistical error measures, and can improve the prediction ability of the short-term wind speed.

In addition, the main contribution of this paper can be summarized as follows:

- (1) Considering the seasonal variation and multi-patterns characteristics of wind speed fluctuations, a comprehensive signal preprocessing technique is proposed to extract the useful information of wind speed fluctuations.
- (2) Instead of determining the input and output relationship of wind speed series by experience way, this study employs partial autocorrelation function (PACF) to find the lag length of wind speed series and determine the number of input dimension for the SVR.
- (3) A novel approach named SIE–WDA–GA–SVR is proposed to predict the short-term wind speed. The proposed strategy can provide higher prediction accuracy compared with the traditional methods.

The rest of this paper is organized as follows. The formulation process of the novel SIE–WDA–GA–SVR model can be described in Sect. 2. Section 3 presents the different error criteria and numerical results obtained from four real datasets. Finally, the conclusions and future researches are summarized in Sect. 4.

2 Proposed Approach

This section presents a novel approach named SIE–WDA–GA–SVR for wind speed forecasting, which feeds SIE and WDA into hybrid model that combines GA and SVR. The proposed approach is briefly described as follows, and the flowchart is shown in Fig. 1.

- Step 1: The original wind speed are decomposed into seasonal and trend components, and calculate the seasonal indices by SIE.
- Step 2: The trend component can be decomposed into both approximate part associated with low frequency and detailed part associated with high frequency by WDA.
- Step 3: By employ the PACF, we can find the input and output relation of the wind speed series, and determine the number of the input dimension of the SVR model for both approximate part and detailed part.
- Step 4: Train the SVR model based on the optimal parameters obtained from GA.
- Step 5: Forecast both low frequency and high frequency using the constructed SVR models.
- Step 6: By the sum of the prediction values of both approximate part and detailed part, we can get the forecasting values of the trend component.
- Step 7: By aggregating the prediction values of seasonal component to the prediction values of trend component, we can obtain the final forecasting result of the wind speed.

2.1 Seasonal Information Extraction (SIE)

As a climate-driven renewable resource, the seasonal variations and trend variations of wind speed are two most commonly encountered phenomena. The SIE technique can decompose the original wind speed datasets into seasonal and trend components, and calculate the

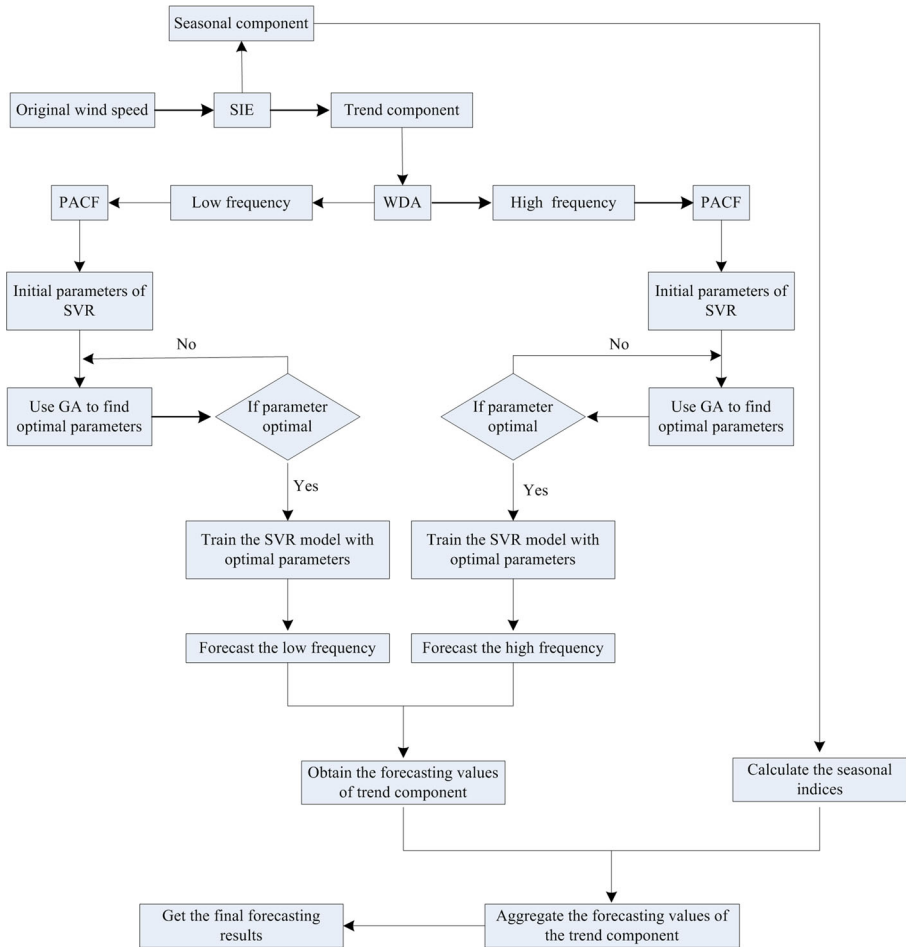


Fig. 1 SIE-WDA-ESVR flowchart

seasonal indices. Generally, addition and multiplication operations are used to generate composite SIE models with seasonal and trend components. According to Zhang and Qi [37], the multiplicative decomposition of the SIE technique is widely used to extract the seasonal information in real datasets. Therefore, the multiplicative composite model is adopted to extract the seasonal information of original wind speed in this study. The concrete process of the algorithm can be described as follows [42]:

Assuming that $T = m \times l$, and m and l denote the number of cycles and the number of data items in each cycle, respectively. Let x_t denote the wind speed at time t ($t = 1, 2, \dots, T$), $S_{.j}$ and Tr_{ij} represent the seasonal and trend components, respectively. Then, x_{ij} denotes the j -th datum of the i -th cycle ($i = 1, 2, \dots, m; j = 1, 2, \dots, l$), and

$$x_{ij} = Tr_{ij} \times S_{.j}, \tag{1}$$

Then, the seasonal index $S_{.j}$ can be obtained by

$$S_{.j} = x_{ij} / Tr_{ij}. \tag{2}$$

Because the trend component Tr_{ij} is unknown, it is need to be approximate by the average of x_{ij} in each cycle.

In fact, the average of the i -th cycle can be derived as follows:

$$\bar{x}_i = (x_{i1} + x_{i2} + \dots + x_{il})/l \quad (i = 1, 2, \dots, m). \tag{3}$$

If S_{ij} denotes the normalization data for items x_{ij} , then

$$S_{ij} = \frac{x_{ij}}{\bar{x}_i} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, l). \tag{4}$$

Then, $S_{.j}$ can be defined as follows:

$$S_{.j} = \frac{S_{1j} + S_{2j} + \dots + S_{mj}}{m} \quad (j = 1, 2, \dots, l). \tag{5}$$

This definition of $S_{.j}$ conforms to the normalization process and is demonstrated as follows:

$$\sum_{j=1}^l S_{.j} = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^l S_{ij} = \frac{1}{m} \sum_{i=1}^m \left(\sum_{j=1}^l x_{ij}/\bar{x}_i \right) = \frac{1}{m} \sum_{i=1}^m l = l. \tag{6}$$

Then, the trend component can be obtained as follows:

$$Tr_{ij} = \frac{x_{ij}}{S_{.j}} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, l). \tag{7}$$

Considering the cycle influence of wind speed data, in this paper, $l = 24$ is as a cycle and $m = [T/l]$.

2.2 Wavelet Decomposition Algorithm (WDA)

As a relatively novel signal processing technique, wavelet decomposition algorithm (WDA) can decompose a complex signal into both the low and high frequencies. The low and high frequencies which are associated with the approximate part and detailed part, respectively, can show the local and global dynamic properties of the signal at specific timescales [39,40]. So the WDA has been widely used for signal decomposition and complex data processing [43–48]. In this section, the WDA is used for wind speed decomposition, and the brief introduction of the WDA is described as follows.

In general, the WDA can be classed into two categories: continuous wavelet decomposition (CWD) and discrete wavelet decomposition (DWD). Let $\psi(t)$ denote a mother wavelet function, the definition of the CWD can be described as follows [49]:

$$CWT_x^\psi(b, a) = \phi_x^\psi(b, a) = \frac{1}{\sqrt{|a|}} \int x(t) \cdot \psi^\bullet \left(\frac{t-b}{a} \right) dt \tag{8}$$

where $\psi^\bullet(t)$ denotes the complex conjugate of $\psi(t)$, and a and b denote the scale and translational parameters, respectively. When $a = 1/2^s$ and $b = k/2^s$, a discrete version of Eq. (8) can be described as follows:

$$DWT_x^\psi(k, s) = \phi_x^\psi \left(\frac{k}{2^s}, \frac{1}{2^s} \right) = \int_{-\infty}^{\infty} x(t) \cdot \psi^\bullet \left(\frac{t - k/2^s}{1/2^s} \right) dt, \tag{9}$$

where s and k meet the following constrained form

$$\begin{cases} T = 2^s + k \\ 0 \leq s \leq \log_2 T, \quad s \in Z \\ k \in Z \end{cases}$$

In general, the mother wavelet function needs to be set at the beginning of WDA. Due to the high computing efficiency of the daubechies wavelet filters of order 3 (db3), it is widely applied in data analysis issues. Considering the advantages of db3, this study selects it as a mother wavelet function to generate a set of wavelet basis functions $\{\psi(t)\}_{s,k}$ by the scale and translational transformations of the mother wavelet. In general, the number of these basis functions depends on the length of the signal. In this study, $\log_2 T$ wavelet basis functions can be used to decompose the complex wind speed signal into both the low and high frequencies which can show the local and global dynamic properties of the wind speed at specific timescales. The low frequency is associated with the approximate part and can reveal the trend of wind speed, and the high frequency is associated with the detailed part and tends to be related to exogenous variables effect. For more detail information about WDA, please refer to [43–49].

2.3 Support Vector Regression (SVR)

As a novel machine learning technique, SVR has been superior in minimizing the expected error of a learning machine and reducing the problem of over-fitting [50]. This algorithm has been widely applied in prediction issues [21,22,24,34,36]. The notion of an SVR model can be briefly described as follows.

Given a data set $\{X_i, y_i\}_{i=1}^{sn}$, where $X_i \in R^n$ is the input vector, $y_i \in R$ is the actual output value, and sn is the sample number. The basic idea of the SVR is to map the input vector space into a higher dimensional feature space via a nonlinearly mapping $\varphi(X_i)$, and find a linear function of the higher dimensional feature space to show the nonlinear relationship between input data and output data. The linear function $f(X_i)$ named SVR function can be described as

$$f(X_i) = \omega^T \varphi(X_i) + b, \tag{10}$$

where ω and b are the coefficients.

The values of coefficients ω and b can be estimated by minimizing the following penalty function $R(C, \varepsilon)$

$$R(C, \varepsilon) = \frac{1}{2} \|\omega\|^2 + C \cdot \frac{1}{sn} \sum_{i=1}^{sn} |y_i - f(X_i)|_\varepsilon, \tag{11}$$

where C is the penalty parameter, ε is the non-sensitivity coefficient and denotes the radius of the tube located around the regression function $f(X_i)$, and

$$|y_i - f(X_i)|_\varepsilon = \begin{cases} 0, & |y_i - f(X_i)| \leq \varepsilon \\ |y_i - f(X_i)| - \varepsilon, & \text{otherwise} \end{cases}$$

By introducing two slack variables ξ_i and ξ_i^* , the infeasible constraints of the optimization problem Eq. (11) can be transformed into the following constrained form

$$\begin{aligned} \min(\omega, b, \xi, \xi^*) &= \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{sn} (\xi_i + \xi_i^*) \\ \text{subject to} &\begin{cases} y_i - \omega^T \varphi(X_i) - b & \leq \varepsilon + \xi_i^* \\ -y_i + \omega^T \varphi(X_i) + b & \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* & \geq 0 \end{cases} \end{aligned} \tag{12}$$

Let α_i and α_i^* denote the Lagrange multipliers. By using the Lagrange equation, the maximal dual function can be described as

$$\begin{aligned} \max(\alpha_i, \alpha_i^*) &= \sum_{i=1}^{sn} y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^{sn} (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^{sn} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(X_i, X_j) \\ \text{subject to } \sum_{i=1}^{sn} (\alpha_i - \alpha_i^*) &= 0 \text{ and } \alpha_i, \alpha_i^* \in [0, C]. \end{aligned} \tag{13}$$

By exploiting the optimality constraints, the SVR function can be obtained as follows,

$$f(x_j) = \sum_{i=1}^{sn} (\alpha_i - \alpha_i^*)k(X_i, X_j) + b, \tag{14}$$

where $k(X_i, X_j)$ is the kernel function. In general, the Gaussian radial basis function (RBF) is the most frequently adopted for SVR modeling. In this study, the RBF function is also chosen as the kernel function of SVR model, which is defined as follows:

$$k(X_i, X_j) = \exp\left(-\|X_j - X_i\|^2 / 2\sigma^2\right). \tag{15}$$

where σ denotes the width of the RBF. Thus, in the modeling process of SVR, three parameters need to be chosen including the penalty parameter C , the width of the RBF σ , and the non-sensitivity coefficient ε . In this study, the GA is used to determine the appropriate parameter values of the SVR model.

2.4 The Forecasting Model of GA–SVR

In the modeling process of SVR, the choice of three parameters will influence the performance of the model including the penalty parameter C , the width of the RBF σ , and the non-sensitivity coefficient ε . It is a challenge task to select the appropriate parameters of the SVR model [21,34]. Genetic algorithm (GA) based on the theories of natural selection mechanisms and Darwin’s main principle, has been successfully applied in optimization problems [35,36]. In this study, GA is used to optimize three parameters of the SVR and the partial autocorrelation function (PACF) is used to determine the number of the input dimension for the SVR. The structure of the proposed GA–SVR model for wind speed prediction is shown in Fig. 2 and the operational process of the model can be described as follows.

- Step 1: Divide the wind speed data into the training samples and test samples. The training samples are used to train the forecasting model, and the test samples are used to evaluate the performance and effectiveness of the forecasting model.
- Step 2: Determine the number of the input dimension for the prediction model. Generally, the number of the input dimension can influence the performance of AI techniques directly. As a special kind of AI techniques, the forecasting precision of SVR is also affected by the number of the input dimension. It is a challenging task to select the appropriate number of the input dimension for the SVR model. The partial autocorrelation function (PACF) is used to determine the number of the input dimension for the SVR model.
- Step 3: Randomly generate an initial population of the chromosomes. In the modeling process of the GA–SVR model, three parameters need to be chosen including the penalty parameter C , the width of the RBF σ , and the non-sensitivity coefficient ε . In this

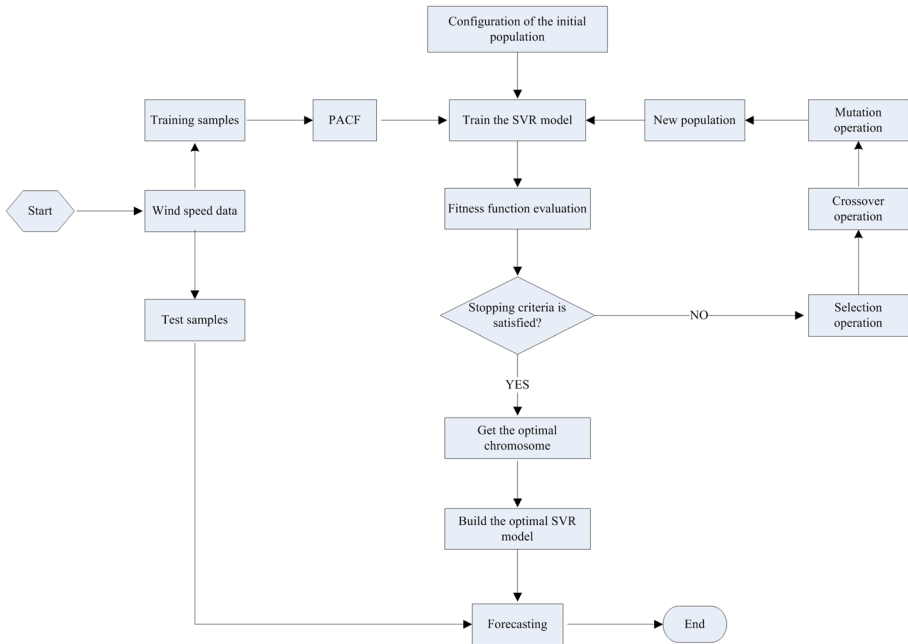


Fig. 2 The structure of the proposed GA-SVR model

study, the size of the initial population is set to 20 chromosomes and each chromosome is consisted of three segments corresponding three parameters of the SVR model.

- Step 4: Calculate the fitness function. The fitness function is used to evaluate the optimal chromosome. In this study, the fitness function is defined as $2/\sum_{i=1}^{ts} (y_i - \hat{y}_i)^2$, where ts denotes the number of training samples, y_i and \hat{y}_i represent the actual value and validation value at time i , respectively.
- Step 5: Generate a new population of the chromosomes by selection operation, crossover operation and mutation operation. In this study, the roulette wheel is used for selecting the excellent chromosomes to reproduce. The probability of crossover operation and the rate of mutation operation are set to 0.8 and 0.01, respectively.
- Step 6: If stopping criteria have not been met, return to Step 4.

2.5 Partial Autocorrelation Function (PACF)

In time series analysis issues, the correlation between a variable series and its different lags can be measured by partial autocorrelation function (PACF) method. Inspired by it, this study also uses the PACF method to determine the number of the input dimension for SVR model instead of the experience way. The brief introduction of the PACF is described as follows [51, 52].

If $x_t(t = 1, 2, \dots, T)$ is the wind speed at time t and γ_k denotes the covariance at lag k , then we can get the estimation value $\hat{\gamma}_k$ of γ_k as follows:

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), \quad k = 0, 1, \dots, L \tag{16}$$

where \bar{x} is the average value of time series x_t , T is the data size, and L is the maximum lag. The choice of L depends on the length of the data. In general, $L = T/4$.

If ρ_k denotes the autocorrelation function (ACF) at lag k , then we can get the estimation value $\hat{\rho}_k$ of ρ_k as follows:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \tag{17}$$

If $\beta_{k,k}$ denotes the PACF at lag k , then the estimation value $\hat{\beta}_{k,k}$ of the $\beta_{k,k}$ can be derived as follows:

$$\begin{aligned} \hat{\beta}_{1,1} &= \hat{\rho}_1 \\ \hat{\beta}_{k+1,h} &= \hat{\beta}_{k,h} - \hat{\beta}_{k+1,k+1} \dots \hat{\beta}_{k+1,k-h+1} \quad (h = 1, 2, \dots, k) \\ \hat{\beta}_{k+1,k+1} &= \frac{\hat{\rho}_{k+1} - \sum_{h=1}^k \hat{\rho}_{k+1-h} \hat{\beta}_{k,h}}{1 - \sum_{h=1}^k \hat{\rho}_h \hat{\beta}_{k,h}} \end{aligned} \tag{18}$$

where $k = 1, 2, \dots, L$.

To assess the significance of autocorrelation between lags, the confidence intervals have been widely adopted. In this study, the 95% confidence interval is employed to determine the optimal lags of wind speed for all models. The definition can be described as follows:

$$\begin{aligned} r_{0.95}^+ &= + \frac{2}{\sqrt{T}} \\ r_{0.95}^- &= - \frac{2}{\sqrt{T}} \end{aligned} \tag{19}$$

where T is the data size, $r_{0.95}^+$ and $r_{0.95}^-$ denote the upper and lower critical values, respectively. If $\hat{\beta}_{k,k} \in (r_{0.95}^-, r_{0.95}^+)$, then x_{t-k} is one of input variable. Otherwise, it is not.

3 Experimental Design and Comparison Results

3.1 Evaluation Criteria

In order to evaluate the performance of all involved prediction models, four error measures are adopted including the mean absolute error (MAE), root mean-square error (RMSE), mean absolute percentage error (MAPE) and standard deviation (SD). These error measures are defined as follows:

$$MAE = \frac{1}{fs} \sum_{i=1}^{fs} |e_i|, \tag{20}$$

$$RMSE = \sqrt{\frac{1}{fs} \sum_{i=1}^{fs} e_i^2}, \tag{21}$$

$$MAPE = \frac{1}{fs} \sum_{i=1}^{fs} \left| \frac{e_i}{y_i} \right|, \tag{22}$$

$$SD = \sqrt{\frac{1}{fs} \sum_{i=1}^{fs} (e_i - \mu)^2} \tag{23}$$

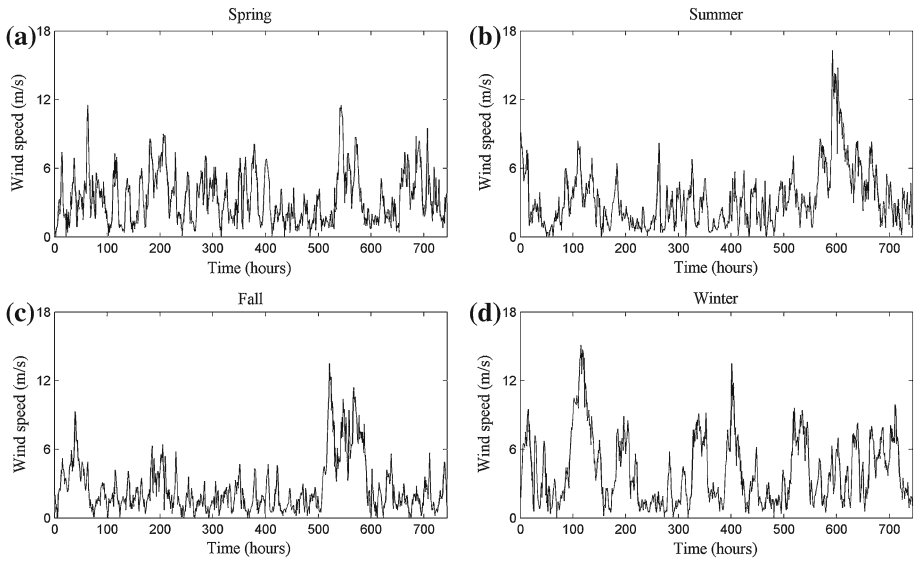


Fig. 3 Four wind speed Cases: **a** Spring, **b** Summer, **c** Fall, and **d** Winter

Table 1 Statistical measures of four cases

Cases	Mean (m/s)	SD (m/s)	Maximum (m/s)	Median (m/s)	Minimum (m/s)
Spring	3.32	2.21	11.50	2.80	0.00
Summer	3.22	2.49	16.30	2.65	0.00
Fall	2.53	2.24	13.50	1.80	0.00
Winter	4.02	2.94	15.10	3.20	0.00

where $e_i = y_i - \hat{y}_i$, $\mu = \frac{1}{fs} \sum_{i=1}^{fs} e_i$, and fs denotes the number of forecasting samples. y_i and \hat{y}_i represent the actual value and forecasting value of wind speed at time i , respectively.

3.2 Datasets

The mean hourly wind speed datasets of wind farm in the province of Gansu, China, are collected to evaluate the proposed model. In order to further verify the generalization ability of the proposed model, four wind speed Cases in May 2010, August 2010, October 2010, and January 2011 are randomly selected as the four seasons in a year. Each dataset has 744 wind speed records. In the modeling process, the top 80% of each dataset (about 600 wind speed records) is called as the training dataset which is used to train the proposed model, and the remaining 20% of each dataset (about 144 wind speed records) is called as validation dataset which is used to evaluate the performance and effectiveness of the proposed model. Figure 3 shows four real mean hourly wind speed datasets. Table 1 shows the statistical measures results of four Cases.

3.3 SIE of Wind Speed Datasets

As a climate-driven renewable resource, the seasonal variations and trend variations of wind speed are two most commonly encountered phenomena. In this study, The SIE is used to

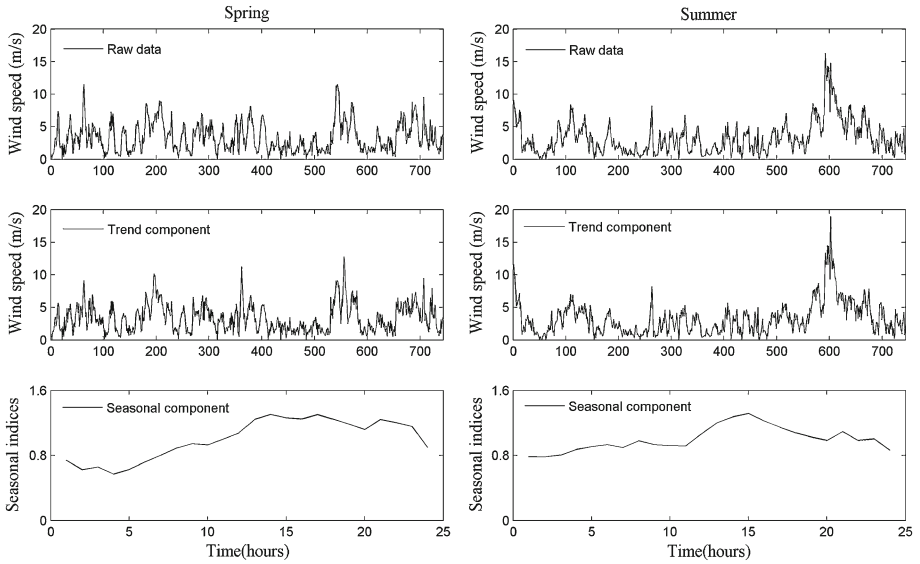


Fig. 4 SIE results in Spring and Summer

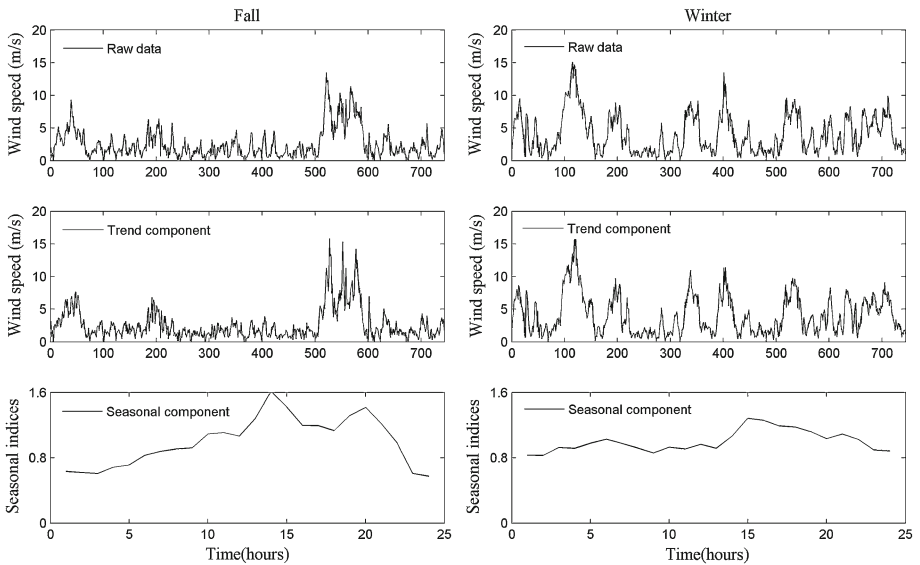


Fig. 5 SIE results in Fall and Winter

decompose the seasonal time series into seasonal and trend components, and extracts the seasonal information of wind speed fluctuations. Figures 4 and 5 show the SIE process of four wind speed datasets. From Figs. 4 and 5, we can see that the seasonal and trend variations of each Case can be obtained by SIE.

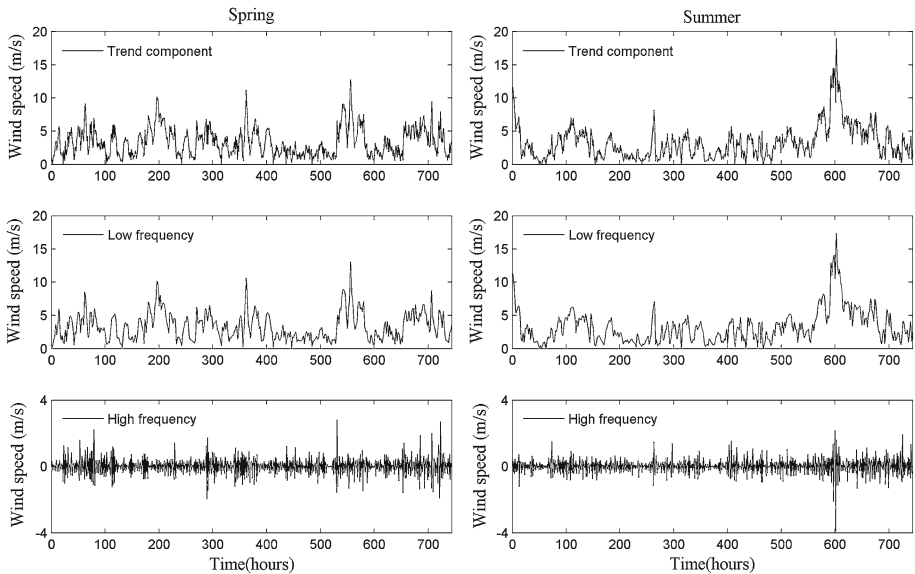


Fig. 6 WDA results in Spring and Summer

3.4 WDA of Trend Components

Due to the intrinsic complexity and multi-patterns of wind speed fluctuations, this study adopts the WDA to decompose the complex trend components of original wind speed signals into both the low and high frequencies. The low and high frequencies which are associated with the approximate part and detailed part, respectively, show the local and global dynamic properties of the wind speed at specific timescales. Figures 6 and 7 show the decomposition process of four Cases by WDA. From Figs. 6 and 7, we observe that the complex trend component of each wind speed signal has been decomposed into both low frequency and high frequency, which are used to establish the corresponding SVR model.

3.5 The Modeling Process of GA-SVR

3.5.1 Input Structure Determination

The number of the input dimension for the SVR has an important influence for forecasting performance. If we ignore the relationship between the wind speed series, this will lead to a bad prediction performance and a slow convergence speed. In order to enhance the prediction ability of the SVR model, we adopt the PACF method to extract the relation of wind speed series, and determine the number of the input dimension for the SVR model. The plots of PACF against the lag length in four Cases can be shown in Fig. 8. The number of the input dimension of each SVR is obtained by the plots of PACF against the lag length, and they can be shown in Table 2. The sample pairs of data can be determined according to the prediction horizon and the number of the input dimension for the model. In this study, we mainly focus on the one-step ahead wind speed prediction, and the prediction horizon is set as one. As an example, the sample pairs of the low frequency in Spring can be shown in Fig. 9. As is shown

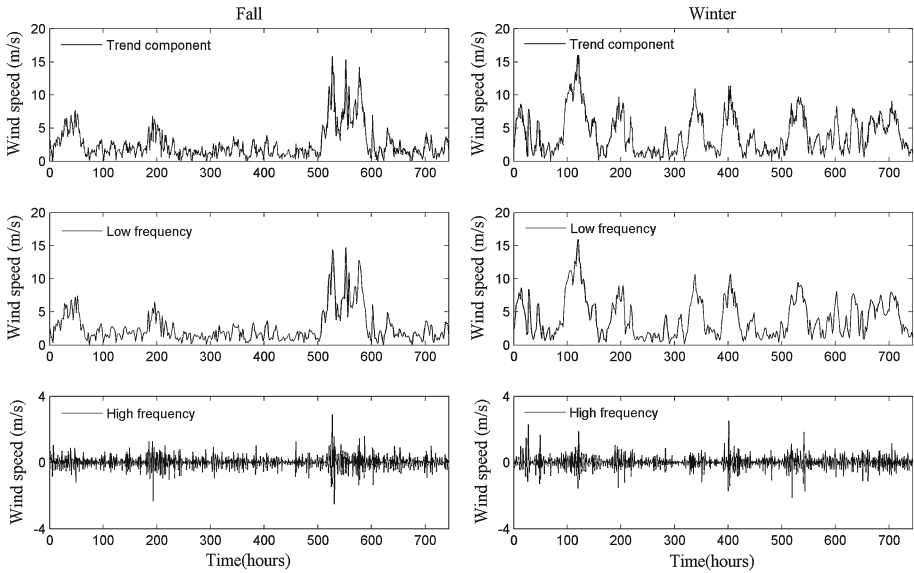


Fig. 7 WDA results in Fall and Winter

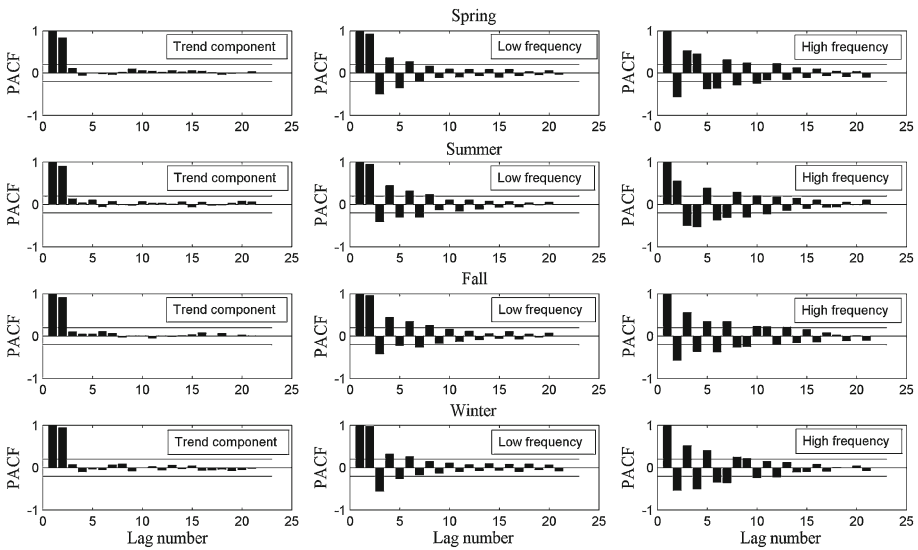


Fig. 8 Plots of PACF against the lag length in four seasons

in Fig. 9, 738 sample pairs can be obtained for the low frequency in Spring. The sample pairs for other wind speed data can be also got in a similar way.

3.5.2 Model Parameters Determination

Three parameters of the SVR models need to be chosen including the penalty parameter C , the width of the RBF σ , and the non-sensitivity coefficient ϵ . In this study, the GA is used to

Table 2 The number of the input dimension for the SVR model by the PACF

Cases	Datasets			Cases	Datasets		
	Trend component	Low frequency	High frequency		Trend component	Low frequency	High frequency
Spring	2	6	12	Fall	2	8	13
Summer	2	8	12	Winter	2	6	12

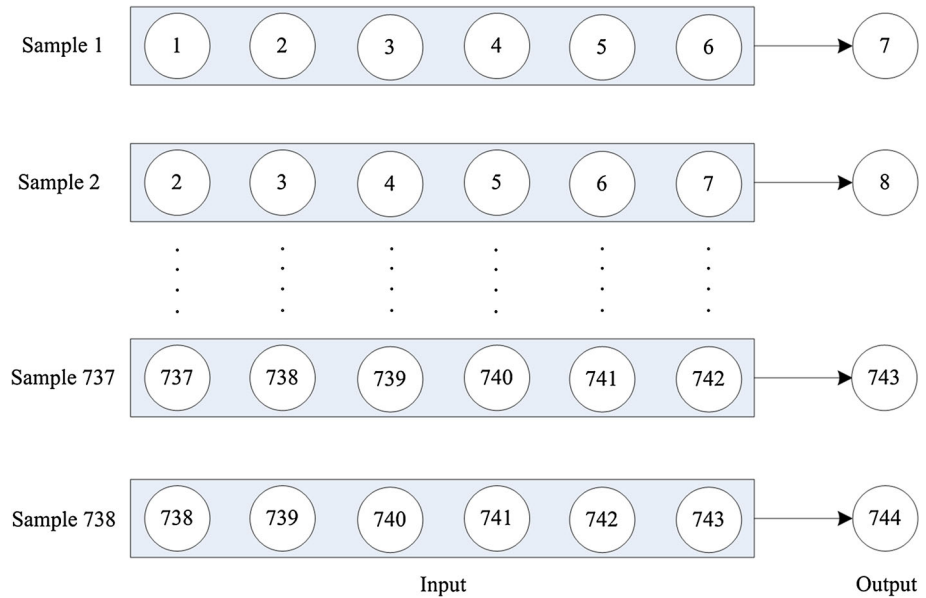


Fig. 9 The sample pairs of the low frequency in Spring

determine the appropriate parameter values of the SVR model. In the modeling process, the sample pairs of data first are determined according to the prediction horizon and the number of the input dimension for model. Then, these sample pairs are randomly partitioned into training set (80%) and validation set (20%). The training set is used to train the GA–SVR, and the validation set is used to test the prediction performance of this model. Note that the training set and validation set are randomly re-partitioned during each simulation, and they are different from each other in multiple simulations. Each simulation can establish a prediction model and obtain a set of model parameters. In this study, the mean of these parameters for the repeated 30 times simulations is used to establish the appropriate model for obtaining the stable result of wind speed prediction. Table 3 shows the final parameters of GA–SVR models for each wind speed series.

3.6 Final Prediction Results of Wind Speed

According to the corresponding SVR model by building in this previous section, both low frequency and high frequency can be predicted. Then, the prediction values of trend component can be obtained by sum the prediction values of both low frequency and high frequency.

Table 3 The final parameters of GA–SVR models for each wind speed series

Cases	Optimal parameters			Cases	Optimal parameters		
	<i>C</i>	σ	ϵ		<i>C</i>	σ	ϵ
Spring				Fall			
Trend component	4.32	1.25	0.00002	Trend component	6.52	1.44	0.00003
Low frequency	4.14	1.17	0.00002	Low frequency	5.59	1.72	0.00003
High frequency	3.81	2.54	0.00005	High frequency	2.95	2.16	0.00004
Summer				Winter			
Trend component	7.21	2.31	0.00006	Trend component	9.76	2.76	0.00001
Low frequency	8.32	2.01	0.00005	Low frequency	8.57	2.65	0.00002
High frequency	5.78	3.19	0.00001	High frequency	6.04	3.12	0.00004

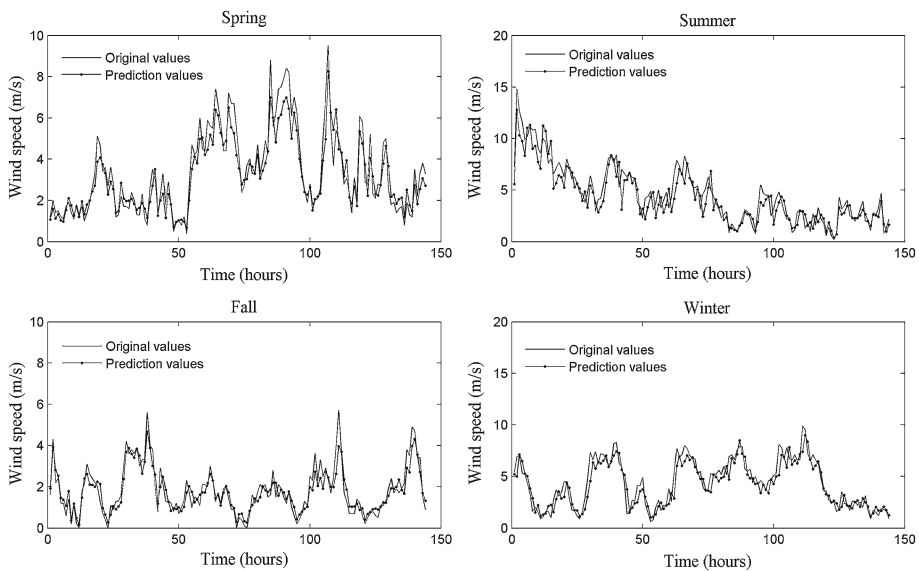


Fig. 10 The final forecasting results of the proposed SIE–WDA–ESVR model

By aggregating the prediction values of seasonal component to the prediction values of trend component, we can obtain the final forecasting results of the wind speed. The final forecasting results of the proposed SIE–WDA–GA–SVR model can be shown in Fig. 10. From Fig. 10, we can see that the prediction values of the proposed approach can approximately describe the characteristics of four wind speed datasets.

3.7 Model Comparisons

In wind speed forecasting issues, the back-propagation neural network (BPNN) which is a benchmark predictor, is often selected as a reference to assess the other forecasting methods. Generally, a novel model is first compared with the BPNN to assess the forecasting ability of it. In the modeling process of the BPNN model or the BPNN part of the hybrid model, the hidden nodes number is determined by Kolmogorov theorem, and the logsig and purelin

Table 4 The comparisons results of these different models

Cases	Errors	Models				
		BPNN	SVR	SIE–BP	SIE–SVR	SIE–WDA–GA–SVR
Spring	MAE	1.3401	1.0221	0.9477	0.8963	0.5472
	RMSE	1.8600	1.3916	1.2735	1.1834	0.6872
	MAPE	0.4331	0.4014	0.3528	0.3189	0.1764
	SD	1.7703	1.3828	1.2432	1.1834	0.6479
Summer	MAE	2.0523	1.1538	1.1229	0.9592	0.8849
	RMSE	2.6778	1.7315	1.5843	1.2258	1.1126
	MAPE	0.6058	0.4345	0.3696	0.2938	0.2412
	SD	2.5266	1.7315	1.5459	1.2254	1.0605
Fall	MAE	1.1360	0.9388	0.5639	0.4637	0.3516
	RMSE	1.4365	1.1411	0.7391	0.6294	0.4372
	MAPE	0.8531	0.7065	0.6513	0.3293	0.3043
	SD	1.0723	0.8781	0.7390	0.6288	0.4349
Winter	MAE	1.0272	0.9727	0.9275	0.8025	0.6876
	RMSE	1.3748	1.2484	1.2445	1.0582	0.9115
	MAPE	0.2924	0.2849	0.2402	0.2516	0.1998
	SD	1.3577	1.2145	1.0709	1.0400	0.8841

functions are selected as the activation functions of both hidden layer and output layer, respectively. The learning velocity is set as the default values 0.01.

In this study, four models of BPNN, SVR, SIE–BP (hybrid SIE and BPNN) and SIE–SVR are selected as the benchmarks to evaluate the forecasting ability of the proposed SIE–WDA–GA–SVR model. The number of the inputs dimension for all models is determined by PACF. The comparisons results of these models are shown in Table 4. From Table 4, we can clearly observe that four statistical errors of the proposed model are the minimum compared with other benchmark models. These results can be explained below. First, the proposed model has the better prediction performance compared with the BPNN and SVR. This result indicates that the proposed approach can fully capture the seasonal information and different patterns of wind speed fluctuations. In addition, the proposed model can considers all linear and non-linear structures of wind speed fluctuations, and has the higher prediction precision compared SIE–BP and SIE–SVR. Thus, it is concluded that the SIE–WDA–GA–SVR model can enhance the forecasting ability of wind speed and is an effective approach.

4 Conclusions

To guarantee the security of wind energy utilization and lower the cost of wind power generation, it is very essential to enhance the prediction ability of wind speed fluctuations. In consideration of the intrinsic complexity and multi-patterns features of the wind speed, a novel SIE–WDA–GA–SVR model is proposed for forecasting the short-term wind speed, which feeds SIE and WDA into hybrid model that combines GA and SVR. The performance of the SIE–WDA–GA–SVR model is comprehensively evaluated using four real prediction cases of wind speed, and compared with a number of benchmark algorithms and baselines. Experimental results indicate that the proposed approach outperforms other benchmark mod-

els in four statistical error measures, and is effective to improve the forecasting accuracy of wind speed.

Acknowledgements This research was supported by the National Natural Science Foundation of China (Grant No. 71501101), the Natural Science Foundation of Jiangsu Province (Grant No. BK20150928), the Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (Grant No. 2015SJB063), the Qing Lan Project, the National Natural Science Foundation of China (Grant No. 1170011208, 91546117, and 61502242), the National Social Science Fund of China (Grant No. 16ZDA047), the Startup Foundation for Introducing Talent of NUIST (S8113097001), the Project Funded by the Flagship Major Development of Jiangsu Higher Education Institutions, the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions, and the Top-notch Academic Programs Project of Jiangsu Higher Education Institutions.

References

1. Ma L, Luan SY, Jiang CW, Liu HL, Zhang Y (2009) A review on the forecasting of wind speed and generated power. *Renew Sustain Energy Rev* 13:915–20
2. Bigdeli N, Afshar K, Gazafroudi AS, Ramandi MY (2013) A comparative study of optimal hybrid methods for wind power prediction in wind farm of Alberta, Canada. *Renew Sustain Energy Rev* 27:20–9
3. Abdelkafi A, Masmoudi A, Krichen L (2013) Experimental investigation on the performance of an autonomous wind energy conversion system. *Int J Electr Power Energy Syst* 44(1):581–90
4. Liu H, Tian HQ, Chen C, Li YF (2013) An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization. *Electr Power Energy Syst* 52:161–73
5. Georgilakis PS (2008) Technical challenges associated with the integration of wind power into power systems. *Renew Sustain Energy Rev* 12(3):852–63
6. Foley AM, Leahy PG, Marvuglia A, McKeogh EJ (2012) Current methods and advances in forecasting of wind power generation. *Renew Energy* 37(1):1–8
7. Wang JJ, Zhang WY, Wang JZ, Han TT, Kong LB (2014) A novel hybrid approach for wind speed prediction. *Inform Sci* 273:304–18
8. Kavasseri RG, Seetharaman K (2009) Day-ahead wind speed forecasting using f-ARIMA models. *Renew Energy* 34:1388–93
9. Mohandes M, Halawani T, Rehman S, Hussain AA (2004) Support vector machines for wind speed prediction. *Renew Energy* 29:939–47
10. Morales JM, Mínguez R, Conejo AJ (2010) A methodology to generate statistically dependent wind speed scenarios. *Appl Energy* 87(3):843–55
11. Fadare DA (2010) The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. *Appl Energy* 87(3):934–42
12. Wang JJ, Zhang WY, Li YN, Wang JZ, Dang ZL (2014) Forecasting wind speed using empirical mode decomposition and Elman neural network. *Appl Soft Comput* 23:452–9
13. Lalarukh K, Yasmin ZJ (1997) Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Sol Energy* 61(1):23–32
14. Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Pytharoulis I, Kallos G (2008) Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J Wind Eng Ind Aerod* 96:2348–62
15. Malmberg A, Holst U, Holst J (2005) Forecasting near-surface ocean winds with Kalman filter techniques. *Ocean Eng* 32:273–91
16. Bivona S, Bonanno G, Burlon R, Gurrera D, Leone C (2011) Stochastic models for wind speed forecasting. *Energy Convers Manage* 52:1157–65
17. Shamshad A, Bawadi MA, Hussin WMAW, Majid TA, Sanusi SAM (2005) First and second order Markov chain models for synthetic generation of wind speed time series. *Energy* 30:693–708
18. Cadenas E, Rivera W (2009) Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. *Renew Energy* 34:274–8
19. Li G, Shi J (2010) On comparing three artificial neural networks for wind speed forecasting. *Appl Energy* 87:2313–20
20. Flores P, Tapia A, Tapia G (2005) Application of a control algorithm for wind speed prediction and active power generation. *Renew Energy* 30(4):523–36
21. Zhou JY, Shi J, Li G (2011) Fine tuning support vector machines for short-term wind speed forecasting. *Energy Convers Manage* 52:1990–8

22. Hu QH, Zhang SG, Xie ZX, Mi JS, Wan J (2014) Noise model based v -support vector regression with its application to short-term wind speed forecasting. *Neural Netw* 57:1–11
23. Kavousi-Fard A, Khosravi A, Nahavandi S (2016) A new fuzzy-based combined prediction interval for wind power forecasting. *IEEE Trans Power Syst* 31(1):18–26
24. Santamaría-Bonfil G, Reyes-Ballesteros A, Gershenson C (2016) Wind speed forecasting for wind farms: a method based on support vector regression. *Renew Energy* 85:790–809
25. Noorollahi Y, Jokar MA, Kalhor A (2016) Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. *Energy Convers Manage* 115:17–25
26. Dowell J, Pinson P (2016) Very short-term probabilistic wind power forecasts by sparse vector autoregression. *IEEE T Smart Grid* 7(2):763–70
27. Ghorbani MA, Khatibi R, FazeliFard MH, Naghipour L, Makarynskyy O (2016) Short-term wind speed predictions with machine learning techniques. *Meteorol Atmos Phys* 128:57–72
28. Men ZX, Yee E, Lien FS, Wen DY, Chen YS (2016) Short-term wind speed and power forecasting using an ensemble of mixture density neural networks. *Renew Energy* 87:203–11
29. Heineremann J, Kramer O (2016) Machine learning ensembles for wind power prediction. *Renew Energy* 89:671–9
30. Sun W, Liu MH (2016) Wind speed forecasting using FEEMD echo state networks with RELM in Hebei, China. *Energy Convers Manage* 114:197–208
31. Doucoure B, Agbossou K, Cardenas A (2016) Time series prediction using artificial wavelet neural network and multi-resolution analysis: application to wind speed data. *Renew Energy* 92:202–11
32. Wang SX, Zhang N, Wu L, Wang YM (2016) Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. *Renew Energy* 94:629–36
33. Meng AB, Ge JF, Yin H, Chen SZ (2016) Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. *Energy Convers Manage* 114:75–88
34. Hu WB, Yan LP, Liu KZ, Wang H (2016) A Short-term traffic flow forecasting method based on the hybrid PSO-SVR. *Neural Process Lett* 43:155–72
35. Cao Q, Parry ME (2009) Neural network earnings per share forecasting models: a comparison of backward propagation and the genetic algorithm. *Decis Support Syst* 47:32–41
36. Gu JR, Zhu MC, Jiang LGY (2011) Housing price forecasting based on genetic algorithm and support vector machine. *Expert Syst Appl* 38:3383–6
37. Zhang GP, Qi M (2005) Neural network forecasting for seasonal and trend time series. *Eur J Oper Res* 160:501–14
38. Wang JJ (2014) A hybrid wavelet transform based short-term wind speed forecasting approach. *Sci World J* 914127:1–12
39. Cohen A, Daubechies I, Vial P (1993) Wavelets on the interval and fast wavelet transform. *Appl Comput Harmon A* 1(1):54–81
40. Gencay R, Selcuk F, Whitcher B (2001) Differentiating intraday seasonalities through wavelet multi-scaling. *Phys A* 289(3–4):543–56
41. Cannas B, Fanni A, See L, Sias G (2006) Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. *Phys Chem Earth* 31(18):1164–71
42. Guo ZH, Wu J, Lu HY, Wang JZ (2011) A case study on a hybrid wind speed forecasting method using BP neural network. *Knowl-Based Syst* 24:1048–56
43. Donoho DL, Johnstone IM (1998) Minimax estimation via wavelet shrinkage. *Ann Stat* 26:879–921
44. Chang SG, Yu B, Vetterli M (2000) Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Trans Image Process* 9:1522–31
45. Ramsey JB (2002) Wavelets in economics and finance: past and future. *Stud Nonlinear Dyn Econom* 6:1–27
46. Li T, Li Q, Zhu S (2003) A survey on wavelet applications in data mining. *Sigkdd Explor* 4:49–68
47. Hussain MS, Reaz MBI, Mohd-Yasin F, Ibrahimy MI (2009) Electromyography signal analysis using wavelet transform and higher order statistics to determine muscle contraction. *Expert Syst* 26:35–48
48. Lu CJ (2010) Integrating independent component analysis-based de-noising scheme with neural network for stock price prediction. *Expert Syst Appl* 37:7056–64
49. Goswami JC, Chan AK (1999) Fundamentals of wavelets: theory, algorithms, and applications. Wiley, Hoboken, pp 149–52
50. Vapnik VN (1999) An overview of statistical learning theory. *IEEE Trans Neural Netw* 10:988–999
51. Wang H, Zhao W (2009) ARIMA model estimated by particle Swarm optimization algorithm for consumer price index forecasting. *lecture notes in computer science. Artif Intell Comput Intell* 5855:48–58
52. Guo ZH, Zhao WG, Lu HY, Wang JZ (2012) Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renew Energy* 37:241–9