Energy 93 (2015) 1456–1466

Contents lists available at ScienceDirect

Energy

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

Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression

Jianming Hu [a], Jianzhou Wang [b,*]

[a] School of Mathematics & Statistics, Lanzhou University, Lanzhou 730000, China
[b] School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

ARTICLE INFO

Article history:
Received 17 December 2014
Received in revised form
17 July 2015
Accepted 13 October 2015
Available online 19 November 2015

Keywords:
Gaussian process regression model
Wind speed forecasting
Empirical wavelet transforms (EWT)

ABSTRACT

Short-term wind speed forecasting plays a major role in wind energy plant operations and the integration of wind power into traditional grid systems. This paper proposes a hybrid model, which is composed of the EWT (empirical wavelet transform), PACE (partial auto-correlation function) and GPR (Gaussian process regression) method, for short-term wind speed prediction. In this proposed approach, the EWT is employed to extract meaningful information from the wind speed series by designing an appropriate wavelet filter bank, and the GPR simulates the internal uncertainties and dynamic features of the wind speed time-series using inputs identified by the PACF. The hybrid GPR model can offer point predictions and interval estimations of future wind speed. Additionally, this study adopts a moving window approach in the prediction process to deal adequately with the training data set, thereby adapting to the time-varying characteristic of the wind speed. The proposed hybrid model was validated with real mean half-hour wind speed data and hourly wind speed data. The computational results show that the suggested hybrid model favorably improves point wind speed forecasts in comparison with other models and provides satisfactory interval wind speed prediction.

1. Introduction

Along with the rapid development and utilization of wind energy, installed wind capacity is drastically rising, and the global wind industry has boomed in recent years [1]. However, due to the intermittent and stochastic nature of wind sources, the integration of wind power into traditional grid systems poses many challenges to the wind industry, including energy generation planning and turbine maintenance scheduling, and electrical grid system operations, including the management of the variability of wind power generation and interconnection standards [1,2].

To mitigate the aforementioned problems caused by the integration of wind energy into power systems, accurate dynamic wind speed prediction by reliable methods and models is becoming increasingly important and urgently needed. Short-term wind speed prediction is an important, representative and effective way to obtain accurate information, which is instrumental in the planning of economic load dispatch and load increment/decrement decisions made with respect to the management of a significant amount of wind power [3]. Numerous methods and approaches have been proposed in the existing literature, and the proposed models can be grouped into two categories: point forecasting of the wind speed and interval forecasting.

At present, the majority of the models in the literature have focused on point wind speed forecasting. These methods include physical modeling methods NWP (numerical weather prediction), time series models, soft computing approaches and hybrid models. The NWP models carry out predictions for weather conditions, including wind speed, through simulations of fluid dynamics and thermodynamics equations [4,5]. The models generally do well in extrapolating future wind conditions and trends. Time series models are widely used tools in the forecasting field, including short-term wind speed forecasting; examples include ARMA (autoregressive moving average) [6], ARIMA (autoregressive integrated moving average) [7], FARIMA (fractional autoregressive integrated moving average) [8], exponential smoothing techniques [9] and grey predictors [10]. These time series models can yield good forecasts under the conditions for which time series possess linearity.

Soft computing methods possess good capabilities, such as good adaptability and strong generality, and they have been extensively

* Corresponding author. Tel.: +86 931 8914050; fax: +86 931 8912481.
E-mail address: wjz@lzu.edu.cn (J. Wang).

http://dx.doi.org/10.1016/j.energy.2015.10.041
0360-5442/

used by researchers for wind speed forecasting. Among the soft computing methods, artificial neural networks, including the Elman neural network [11], BPNN (back-propagation neural network) [12,13], recurrent neural networks [14], RBFNN (radial basis function neural network) [15] and multilayer perceptron neural network [16] are among the most extensively used approaches for the prediction of wind speed. SVM (support vector machine) [17–19] is another type of soft computing method. A SVM can be generalized to deal with high number of dimensions, small samples, and complex nonlinear problems.

In recent years, hybrid approaches to wind speed and wind power forecasting have been increasingly investigated to pursue more accurate and stable forecasts. Several types of wind forecasting studies with hybrid models have been presented, including the hierarchical integration of NWP and statistical models and the combination of candidate statistical models and modeling algorithms. In the integration of NWP and statistical models, statistical methods are used to adjust the numerical wind speed predictions to provide significant forecast improvement [5]. The combinations of candidate statistical models and modeling algorithms produce wind speed predictions through integrating individual statistical models by the same or different modeling algorithms; examples include data preprocessing-based approaches [19], parameter optimization based approaches [20] and the weighting-based approaches [21].

Researchers proposed the aforementioned models for point or deterministic wind speed or power forecasts. For optimal management of wind power production and distribution, it is also important to obtain information on the uncertainty of potential future events in addition to single-valued predictions. Information about the uncertainty of projections can be conveyed by prediction intervals in a practical and visual way [23]. Interval predictions of wind speed can meet the needs of various applications, including power system operations [24,25] where risks and uncertainties must be quantified. For this reason, researchers have conducted considerable research in interval wind speed or power prediction considering uncertain information. The approaches proposed in the literature include probability-distribution functions [26,27] Bayesian theory and structural break modeling [28], copula estimator [29], fuzzy inference [22], Bayesian model averaging model [30].

To provide satisfactory and reliable point forecasts and interval estimations for future wind speeds, this paper proposes a hybrid forecasting approach using hybrid prediction. The proposed hybrid constructor incorporates the EWT (empirical wavelet transform), PACF (partial autocorrelation function) and GPR (Gaussian process regression) method, which performs the distribution prediction of the future wind speed. The prediction intervals, quantiles and single-valued predictions can be inferred from the obtained predictive set of forecasts depending upon the specific interest. The developed hybrid method is examined using mean half-hour wind speed data and hourly wind speed data. The simulation results reveal that the hybrid forecasting method outperformed other popular algorithms.

The remainder of this paper is organized as follows. Section 2 discusses the contributions of the proposed model. Section 3 introduces the required individual models and describes the developed hybrid model. In Section 4, the wind speed predictions and advantages of the developed strategy are analyzed and discussed through comparisons with other benchmark models. Finally, the study's conclusions are presented in Section 5.

2. Contribution

As mentioned in Section 1, hybrid approaches utilize the advantages of each individual model and combine them through the same or different modeling algorithms, thereby generating more accurate and reliable wind speed predictions, to some extent.

Based on the aforementioned forecasting methodology, this article integrates two already existing models and algorithms for wind speed forecasting. The main contributions of this study with respect to those offered by other studies in the same area of research can be summarized as follows:

(1) The hybrid approach is proposed not only to generate point and interval wind speed forecasts but also to tackle the characteristics of wind speed series. The proposed approach extracts meaningful information from a short-term wind speed series to eliminate disturbing factors, tackles the uncertainties in the wind speed series with the latent function, and then models the behavior of the wind speed. Finally, the hybrid approach provides the uncertainty associated with the future wind speed.

(2) In the pretreatment of the short-term wind speed series, the WT (wavelet transform) or EMD (empirical mode decomposition) techniques are commonly used to eliminate noise in the time series. However, the WT lacks the ability of self-adaptive data processing and the wavelet basis and parameters need to be specified beforehand, while the EMD is sensitive to noise and sampling and lacks mathematical theory. The EWT method remedies the drawbacks of the aforementioned decomposition methodologies to some extent. It can adaptively represent the processed signal by automatically generating the adaptive wavelet and then decomposing the signal into a finite number of modes.

(3) GPR is employed to generate probabilistic wind speed forecasts. GPR is a principled and practical probabilistic approach, which is advantageous in the interpretation of model predictions. It possesses very good adaptability and strong generality to deal with high dimensions and small samples in complex nonlinear problems. Compared with neural networks and SVM, GPR is easy to implement, self-adaptive to enable superior parameter estimation, and flexible enough to make nonparametric inferences.

(4) The GPR is based on Bayesian inference that solves this issue by incorporating prior domain knowledge of wind speed characteristics and by specifying prior distributions for the model parameters. In addition, a Bayesian forecasting model can produce a predictive sample of the wind speeds. The predictive sample provides more information than a classical point forecast, including credible intervals and quantile.

(5) Wind speed is characteristically time-varying, which should be exhibited by the wind speed forecasting model. Thus, this study employs a moving window method in these prediction processes, which allows the proposed hybrid model to adaptively respond to wind speed changes and to better reflect the actual forecasting environment.

3. The hybrid EWT-GPR model

The hybrid constructor composed of the EWT and GPR model is proposed for point wind speed forecasting as well as interval estimation. In the suggested hybrid model, the EWT is employed to extract different modes from the wind speed series by designing an appropriate wavelet filter bank, and the GPR model performs the forecasting task with the appropriate input determined by the PACF. To adapt quickly to the dynamic features of the wind speed time-series, a moving window approach is adopted. Both the point and interval wind speed forecast can be inferred from the predictive probability distribution of the future wind speed obtained by the proposed approach. Fig. 1 shows the general structure of the proposed hybrid wind speed forecasting method.

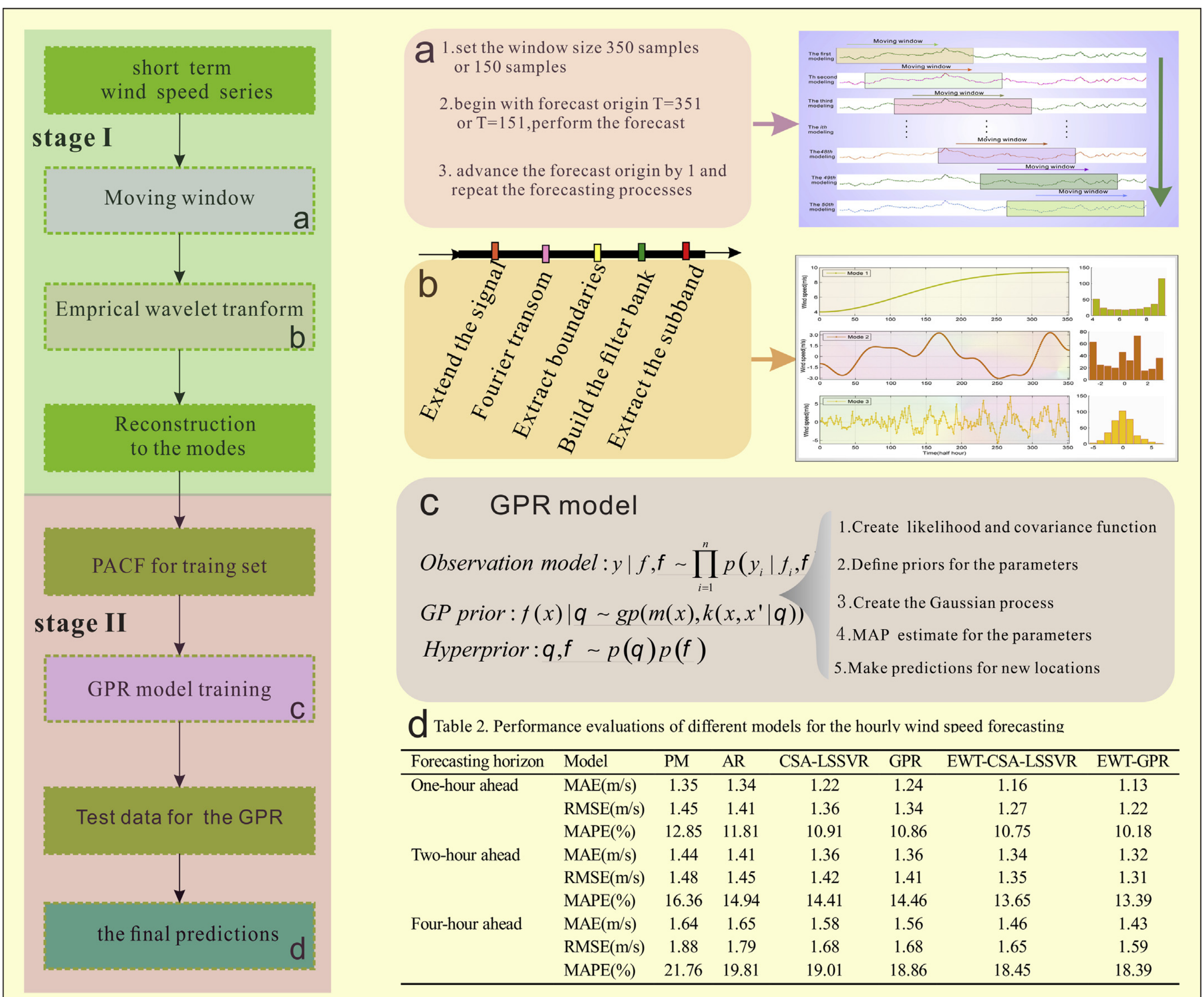
The hybrid approach mainly contains two stages; the main tasks in each stage are described as follows:

Stage I: Data preprocessing. The EWT algorithm is utilized to divide the original short-term wind speed series into several independent modes and one residual series. Specifically, the wind speed series is firstly extended by mirroring, then the extended signal is computed using a Fourier transform; by segmenting the Fourier spectrum based on the detected maxima, the empirical wavelet filter bank is established when the condition (3) (see the Subsection 3.1) is satisfied. Finally, the filter bank is used to filter the signal to extract each subseries. After the decomposition to the original wind speed series, the modes and one residual are obtained. The decomposed modes possess different meanings at any point, e.g., the low-frequency component represents the main features of the raw data series. The decomposition extracts the meaningful information from the original wind speed series, thereby making preparations for the forecasting. Thus, the residual series is discarded because the residual is small and can be regarded as an uncorrelated white noise series; the decomposed modes are aggregated into the new data series.

Stage II: Forecasting. The PACF is employed to first identify the partial autocorrelation of the new data reconstructed by the meaningful modes. Then the appropriate inputs determined from the PACF are used to construct the Gaussian likelihood. By applying gradient-based optimization methods to the Gaussian likelihood, the maximum a posterior estimates of the parameters can be obtained. Finally, the established GPR is henceforth utilized to predict the distribution of the future wind speed with the test set in different forecasting horizons; the predicted mean wind speed can be obtained as well as the estimated interval of wind speed.

3.1. Empirical wavelet transform (EWT)

The EWT (empirical wavelet transform) proposed by Jérôme Gilles [31] identifies and extracts the different intrinsic modes of a time series. The algorithm relies on robust preprocessing for peak detection, then performs spectrum segmentation based on detected maxima, and constructs a corresponding wavelet filter bank.



Forecasting horizon	Model	PM	AR	CSA-LSSVR	GPR	EWT-CSA-LSSVR	EWT-GPR
One-hour ahead	MAE(m/s)	1.35	1.34	1.22	1.24	1.16	1.13
	RMSE(m/s)	1.45	1.41	1.36	1.34	1.27	1.22
	MAPE(%)	12.85	11.81	10.91	10.86	10.75	10.18
Two-hour ahead	MAE(m/s)	1.44	1.41	1.36	1.36	1.34	1.32
	RMSE(m/s)	1.48	1.45	1.42	1.41	1.35	1.31
	MAPE(%)	16.36	14.94	14.41	14.46	13.65	13.39
Four-hour ahead	MAE(m/s)	1.64	1.65	1.58	1.56	1.46	1.43
	RMSE(m/s)	1.88	1.79	1.68	1.68	1.65	1.59
	MAPE(%)	21.76	19.81	19.01	18.86	18.45	18.39

Fig. 1. The overall framework of the hybrid EWT-GPR model.

The empirical wavelets can be defined as bandpass filters on each Λ_n, where Λ_n denotes each segment $\Lambda_n = [w_{n-1}, w_n]$, and $\cup_{n=1}^{N} \Lambda_n = [0, \pi]$. $\forall n > 0$ The empirical scaling function and the empirical wavelets are defined by equations (1) and (2), respectively:

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & if|\omega| \le (1-\gamma)\omega_n \\ cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}|\omega| - \omega_n + \tau_n\right)\right] & if(1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n \\ 0 & otherwise \end{cases} \tag{1}$$

and

$$\hat{\varphi}_n(\omega) = \begin{cases} 1 & if(1+\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_{n+1} \\ cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_{n+1}}|\omega| - \omega_{n+1} + \tau_{n+1}\right)\right] & if(1-\gamma)\omega_{n+1} \le |\omega| \le (1-\gamma)\omega_{n+1} \\ sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}|\omega| - \omega_n + \tau_n\right)\right] & if(1-\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_n \\ 0 & otherwise \end{cases} \tag{2}$$

The function $\beta(x)$ is an arbitrary $C^k([0,1])$ function and satisfies the following properties.$\beta(x) = \begin{cases} 0 & if \ x \le 0 \\ 1 & if \ x \le 0 \end{cases}$ and $\beta(x) + \beta(1-x) = 1, \forall \ x \in [0,1]$

When

$$\gamma < \min_n \left(\frac{\omega_{n+1} - \omega_n}{\omega_{n+1} + \omega_n}\right), \tag{3}$$

then the set $\{\phi_1(t), \{\varphi_n(t)\}_{n=1}^{N}\}$ is a tight frame of $L^2(R)$.

The detail coefficients are obtained from the following inner products between the signal and the empirical wavelets:

$$\begin{aligned} w_f^{\varepsilon}(n,t) = \langle f, \varphi_n \rangle = \int f(\tau)\overline{\varphi_n(\tau - t)}d\tau \\ = f(\tau)\left(\overline{\hat{\varphi}_n(\tau - t)}\right)^{\vee} \end{aligned} \tag{4}$$

and the approximation coefficients are calculated by the following inner products with the scaling function

$$\begin{aligned} w_f^{\varepsilon}(0,t) = \langle f, \phi_1 \rangle = \int f(\tau)\overline{\phi_1(\tau - t)}d\tau \\ = f(\tau)\left(\overline{\hat{\phi}_1(\tau - t)}\right)^{\vee} \end{aligned} \tag{5}$$

The reconstruction is obtained by

$$\begin{aligned} f(t) = w_f^{\varepsilon}(0,t)*\phi_1(t) + \sum_{n=1}^{N} w_f^{\varepsilon}(n,t)*\varphi_n(t) \\ = \left(\widehat{w_f^{\varepsilon}}(0,\omega)*\phi_1(\omega) + \sum_{n=1}^{N} \widehat{w_f^{\varepsilon}}(n,\omega)*\varphi_n(\omega)\right)^{\vee} \end{aligned} \tag{6}$$

Following this formalism, the empirical mode is given by

$$f_0(t) = w_f^{\varepsilon}(0,t)*\varphi_1(t) \tag{7}$$

$$f_k(t) = w_f^{\varepsilon}(k,t)*\varphi_k(t) \tag{8}$$

Because the EWT borrows the framework of the classic wavelet transform for decomposing and reconstructing the signals. Numerous studies have demonstrated that three-level decompositions can achieve good forecasting results for non-stationary time series, such as electricity load series, price series [32], wind power series [33] and wind speed series [34]; therefore this study performs three-level decomposition, similar to the literature [34].

3.2. Gaussian process regression model

GP (Gaussian processes) are powerful tools for probabilistic modeling [35]. They can be used to define prior distributions over latent functions in hierarchical Bayesian models. The prior over function is defined implicitly by the mean and covariance functions, which determine the smoothness and variability of the function. The inference can then be conducted directly in the function space by evaluating or approximating the posterior process.

The models can be summarized as:

$$\text{Observation model}: \ y \Big| f, \phi \sim \prod_{i=1}^{n} p(y_i|f_i, \phi) \tag{9}$$

$$\text{GP prior}: \ f(x)|\theta \sim gp(m(x), k(x, x'|\theta)) \tag{10}$$

$$\text{Hyperprior}: \ \theta, \phi \sim p(\theta)p(\phi) \tag{11}$$

where $m(x)$ and $k(x, x'|\theta)$ denote the mean and covariance functions, ϕ and θ are the parameters of the covariance function and the parameters of the observation model, respectively. The covariance function is the crucial ingredient in the GP predictors as it encodes the prior assumptions on the latent function, such as the smoothness and scale of the variation. A function of input pairs is a valid covariance function as long as the covariance matrices that it produces are symmetric and positive semi-definite. An example of a stationary covariance function is the squared exponential

$$k(x_i, x_j|\theta) = \sigma_{se}^2 \exp\left(\sum_{k=1}^{d} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{2l_k^2}\right) \tag{12}$$

where $\theta = \{\sigma_{se}^2, l_1 \cdots l_k\}$ σ_{se}^2 is a magnitude parameter that scales the overall variation of the unknown function, and l_k is a length-scale parameter that governs how fast the correlation decreases as the distance increases in the input dimension k.

A GPR model can be constructed as follows

$$y_i = f(x_i) + \varepsilon_i, i = 1, \cdots n \tag{13}$$

Where $\varepsilon_i \sim N(0, \sigma^2), i = 1, \cdots n, f(x)$ is given a GP prior. Assuming the predicted values $\tilde{f}$ with new input $\tilde{x}$ to the latent function, the joint prior for latent variables f and $\tilde{f}$ is

$$\begin{bmatrix} y \\ \tilde{f} \end{bmatrix} \bigg| x, \tilde{x}, \theta \sim N\left(0, \begin{bmatrix} K_{ff} + \sigma_n^2 I & K_{f\tilde{f}} \\ K_{\tilde{f}f} & K_{\tilde{f}\tilde{f}} \end{bmatrix}\right) \tag{14}$$

where $K_{f\tilde{f}} = k(x, \tilde{x}|\theta)$ and $K_{\tilde{f}\tilde{f}} = k(\tilde{x}, \tilde{x}|\theta)$. By defining the marginal distribution of $\tilde{f}$ as $p(\tilde{f}|\tilde{x}, \theta) = N(\tilde{f}|0, K_{\tilde{f}\tilde{f}})$, the conditional distribution of $\tilde{f}$ given f is

$$\tilde{f}\bigg| y, x, \tilde{x}, \theta \sim N\left(m(\tilde{x}|\theta), k\left(\tilde{x}, \tilde{x}'\bigg|\theta\right)\right) \tag{15}$$

where the mean function $m(\tilde{x}|\theta) = k(x, x|\theta)(K_{ff} + \sigma_n^2 I)^{-1} y$ and covariance function $k(\tilde{x}, \tilde{x}'|\theta) = k(\tilde{x}, \tilde{x}'|\theta) - k(\tilde{x}, x|\theta)(K_{ff} + \sigma_n^2 I)^{-1} k(x, \tilde{x}'|\theta)$, which define the conditional distribution of the latent function

3.3. Partial autocorrelation function

The PACF (partial autocorrelation function) plays a very important role in recognizing the non-stationary of wind speed series and determining the lag orders.

First we determine the ACVF (autocovariance function) $\gamma(\cdot)$ of the causal ARMA (p, q) process defined by

$$\phi(B)X_t = \theta(B)Z_t, \quad \{Z_t\} \sim WN\left(0, \sigma^2\right), \tag{16}$$

where $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$. The causality assumption implies that

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{17}$$

where $\sum_{j=1}^{\infty} \psi_j z_j = \theta(z)/\phi(z), \ |z| \le 1$.

Based on the above, we can obtain

$$\gamma(h) = E(X_{t+h}X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|}. \tag{18}$$

The PACF (partial autocorrelation function) of an ARMA process $\{X_t\}$ is the function $\alpha(\cdot)$ defined by the equations.

$\alpha(0) = 1$ and $\alpha(h) = \phi_{hh}, \ h \ge 1$, where ϕ_{hh} is the last component of

$\phi_h = \Gamma_h^{-1} \Upsilon_h, \ \Gamma_h = [\gamma(i-j)]_{i,j=1}^h$, and $\Upsilon_h = [\gamma(1), \gamma(2), \ldots, \gamma(h)]'$.

4. Case study

In this section the efficacy of the proposed approach is demonstrated with a case study and four subsections: collection of data, evaluation criteria of forecasting performance, simulation and comparison and discussion, which are presented sequentially.

4.1. Collection of data

In this study, wind speed data from a wind farm location in China are used to demonstrate the effectiveness and reliability of the proposed hybrid EWT-GPR forecasting approach. Because the GPR models show advantages for addressing small samples and nonlinear problems, this study samples two small wind speed data sets: 422 mean hourly wind speed observations and 222 mean half-hour wind speed observations. In our simulation, the size of moving window are set as 350 samples and 150 samples (small samples)for the hourly wind speed series and half-hour wind speed series, respectively. The samples in the moving window are used to train the EWT-GPR model; as the window moves forward, the corresponding model is established for the hourly forecasting and half-hour forecasting. The process of modeling repeats 72 times with the movement of the fixed-number window.

With respect to the direction of the wind and its impact on the forecasting process, the correction analysis was performed to illustrate the relationship between the wind speed and wind direction. The test results are shown as follows.

Table 1 shows that the correlation coefficient is 0.15 and p-value is 0.662 under the Spearman rank correlation test, which means that there exists weak relationship between wind direction and wind speed, i.e., the direction of the wind has weak impact on the wind speed forecasting. Therefore, the wind direction was not taken into our model.

4.2. Evaluation indices for forecasting performance

To evaluate the generation capacities of the proposed hybrid approach, three statistical indices are utilized to measure the forecasting accuracy. These indices are the MAE (mean absolute error), RMSE (root mean square error) and MAPE (mean absolute percent error), for which small values indicate high forecast performance. These indices are defined as follows:

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \left| p_t^{true} - p_t^{forecast} \right| \tag{19}$$

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{p_t^{true} - p_t^{forecast}}{p_t^{true}} \right| \tag{20}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left| p_t^{true} - p_t^{forecast} \right|^2} \tag{21}$$

where p_t^{true} is the observed value for the time period t and $p_t^{forecast}$ is the predicted value for the corresponding period. The MAE reveals how similar the predicted values are to the observed values, whereas the RMSE measures the overall deviation between the predicted values and the observed values. The MAPE is a unit-free measure of accuracy for the predicted wind series and is sensitive to small changes in the data.

Table 1
The results of correlation test.

Test type	Correlation coefficient	p-value (two tail test)
Spearman test	0.15	0.662

4.3. Simulation

The simulation submits the wind series to the proposed hybrid approach to obtain the predicted outcomes. This subsection considers the hourly wind speed forecasting as an example to present the entire simulation process.

4.3.1. Data preprocessing

It can be seen from Fig. 3 (The top subplot) that the original wind speed series has large fluctuations. As mentioned in Subsection 3.1, three-level decomposition was used because it can describe wind speed series in a meaningful way. With the EWT algorithm, the three uncorrelated modes were extracted from the wind speed series (shown in Fig. 2), and one residual was obtained after the extraction. Prior to the operation of forecasting the future wind speed, the residual was ignored, which played a role in cleaning the noisy information or removing noisy data from the wind speed series. All of the three modes were reconstructed by the EWT into the new series.

4.3.2. Forecasting

After the data preprocessing, the reconstructed data series was first processed into the input of the forecasting engine with the PACF. As can be seen in Fig. 4, there exists high partial autocorrelations in the reconstructed wind speed series. Lag 3 of the PACF is significant after the elimination of internal correlation. Therefore, the values from lag-1 to lag-3 were selected as the input to train the GPR model for the future wind speed forecasting. The available inputs of the GPR model were linearly normalized in the range [0, 1] to overcome the saturation phenomenon. Then, the implementations of the Bayesian inference were conducted with the defined prior for the hyperparameters of the GPR model in advance. Finally, by employing the established GPR model, the forecasting results of the combined series were achieved. As the moving window moved forward, the aforementioned two stages were repeated until the last sample of the data set was processed.

As for the half-hour wind speed forecasting, the foregoing two operations were implemented on the half-hour wind speed series. The corresponding decomposition and reconstruction results are presented in Figs. 6 and 7, respectively, and the half-hour ahead forecasting results are shown in Fig. 5 with the one-hour ahead forecasting results for the hourly wind speed series.

4.4. Comparisons and discussion

To facilitate the analysis and discussions of the suggested hybrid model, four other models for short-term wind speed forecasting were employed for comparison with the proposed model and the assessment of the prediction performance in this subsection. The established models were the AR (auto-regression model), the single GPR model, the CSA-LSSVR model (the LSSVR (Least Squared Support Vector Regression) model optimized by CSA (coupled simulated annealing)), and a hybrid model named EWT-CSA-LSSVR. The selection of the LSSVR model from soft computing methods was based on the consideration that the LSSVR model can offer reliable results with high computing speed. The LSSVR and GPR models can be treated as soft computing algorithms, but they are parametric and nonparametric model, respectively. In addition, the PM (persistence method) was employed as a benchmark to compare with the other tools.

Table 2 shows the evaluation results obtained from the PM, AR, CSA-LSSVR, EWT- CSA-LSSVR and EWT-GPR models with respect to the mean hourly wind speed prediction. It can be readily seen from Table 2 that the proposed hybrid approach outperforms the other

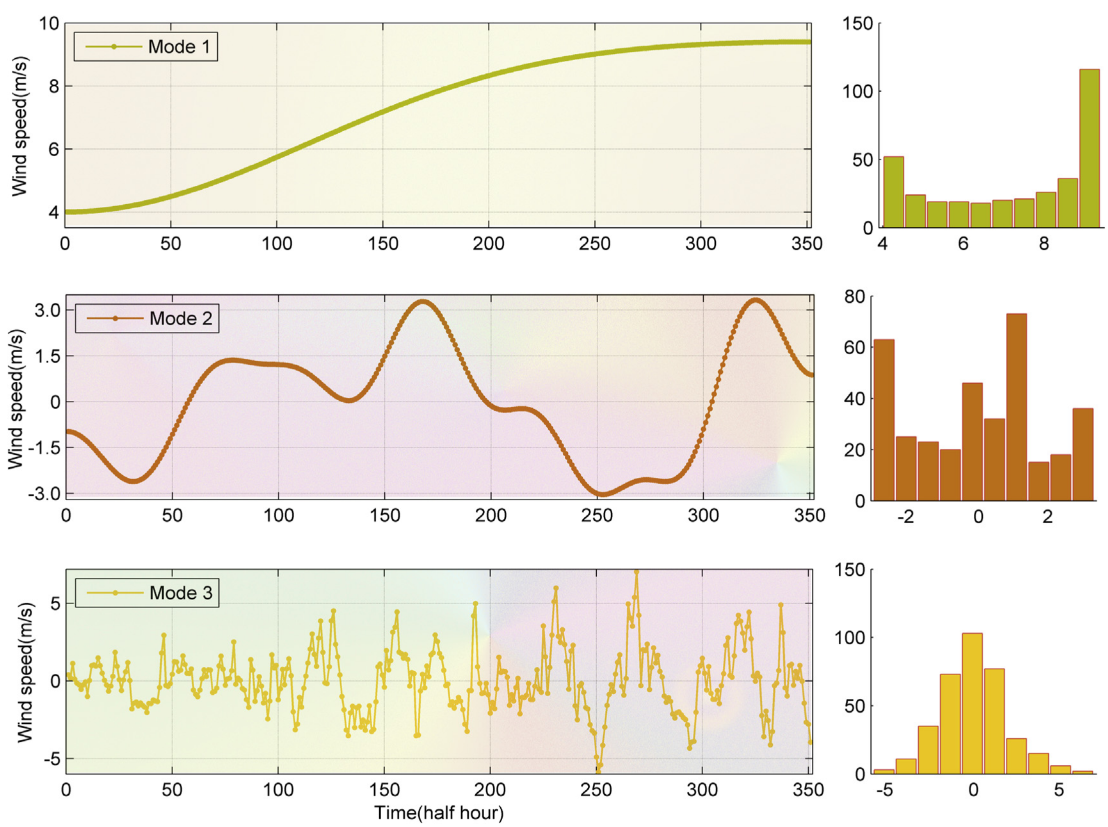


Fig. 2. The decomposed subseries of the hourly data set by EWT.

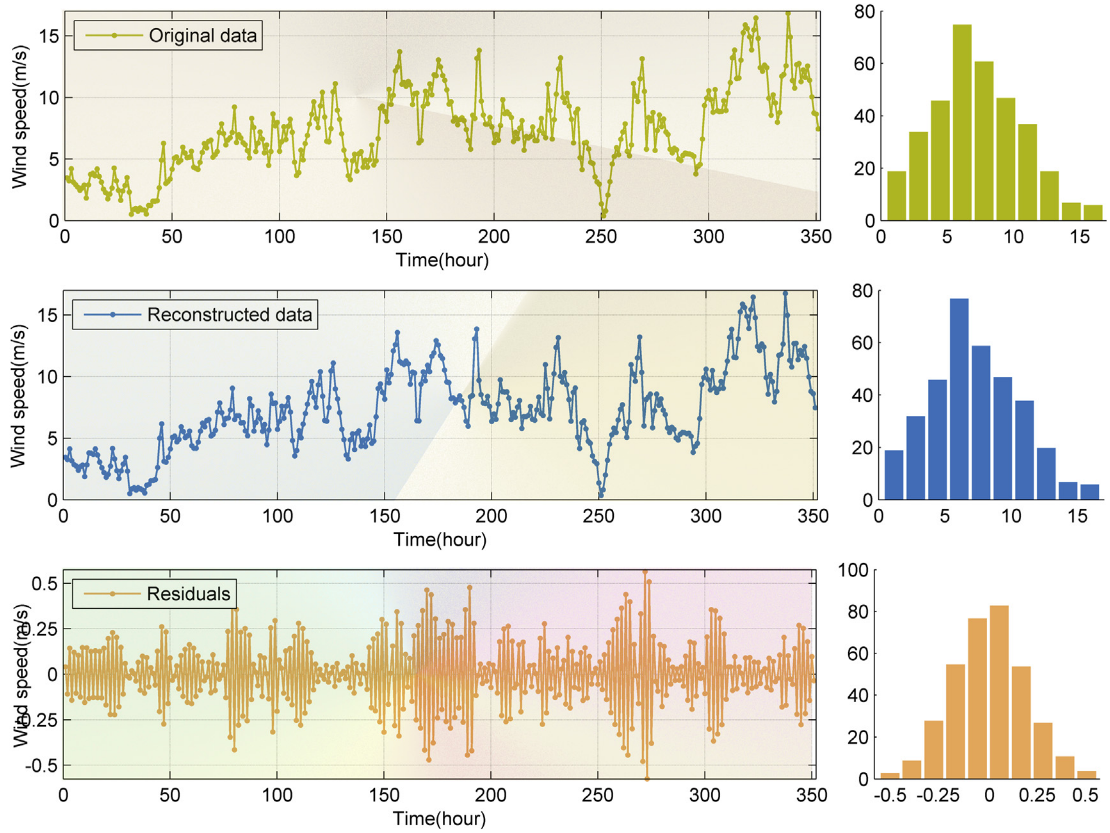


Fig. 3. Comparison of the original hourly wind speed series and the reconstructed wind and residuals series.

models in terms of the three forecasting evaluation indices (MAE, RMSE and MAPE).

More detailed analyses were performed. The proposed hybrid approach performed better than the other models for different forecasting horizons. For example, the developed method outperforms the other models in one-hour ahead forecasting with lower RMSE value of 1.22 m/s for one-hour ahead forecasting in contrast to 1.45 m/s, 1.41 m/s, 1.36 m/s, 1.34 m/s and 1.27 m/s for the PM, AR, CSA-LSSVR, GPR and EWT-CSA-LSSVR models, respectively. A low MAPE value of 13.39% was obtained by the proposed constructor for the two-hour ahead forecasting, while the PM, CSA-LSSVR, EWT- CSA-LSSVR and EWT-GPR methods resulted in higher MAPE values (16.36%, 14.94%, 14.41%, 14.46% and 13.65%), respectively. The comparison of the predictions shows that the integration of the proposed data preprocessor and the forecasting engine is a good choice for short-term wind speed prediction.

To further evaluate the established approaches, the models were submitted to further analyses. First, because the PM is generally used as a benchmark to baseline forecasting performance, other models were compared with it. Table 2 shows that the AR, CSA-

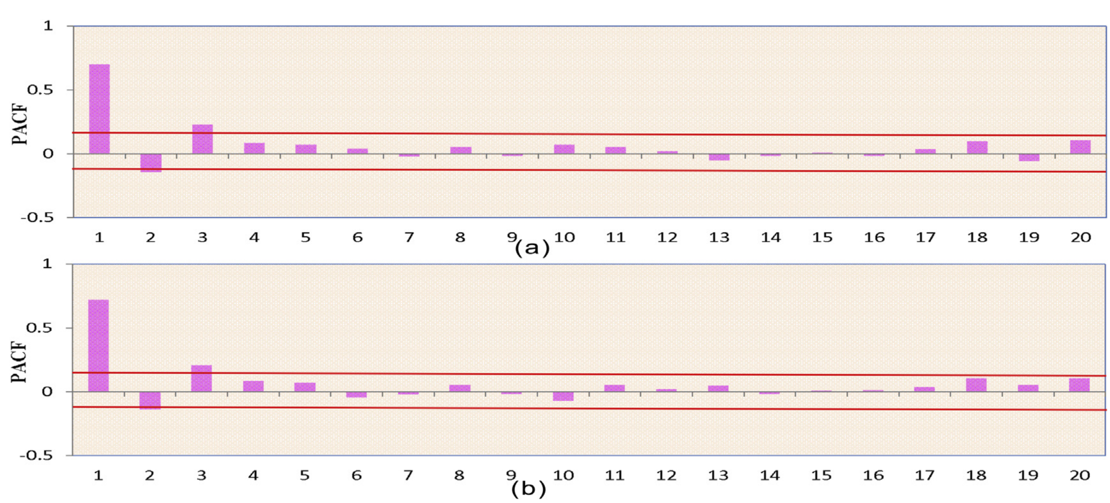


Fig. 4. The PACF values of the samples. (a) is for the hourly wind speed series and (b) is for the half-hour wind speed series.

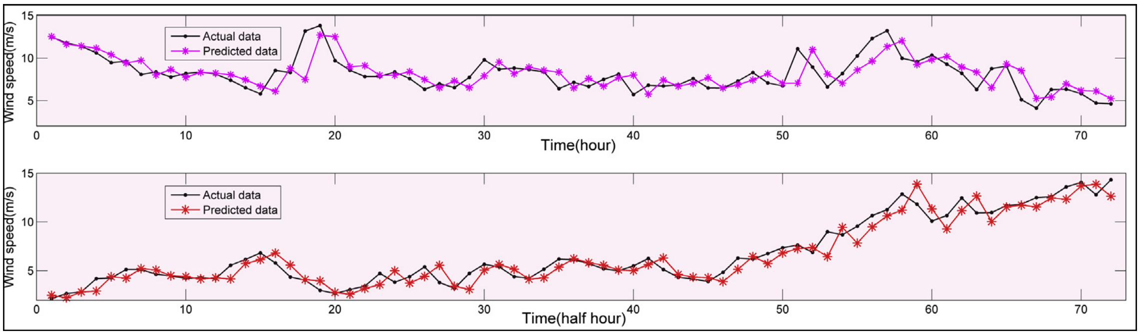


Fig. 5. Comparison between the actual wind speed series and the forecasting results of the proposed model. (The top subplot is the one-hour ahead forecasting and the bottom subplot is the half-hour ahead forecasting.)

LSSVR, GPR and hybrid models performed better than the PM model. Second, the single GPR achieved the same or better performance as CSA-LSSVR, which reveals that nonparametric GPR had nearly the same capacity as the powerful parametric LSSVR model for short term wind speed forecasting. In addition, when compared with the CSA-LSSVR model, the EWT-CSA-LSSVR model had smaller evaluation indices (MAE, RMSE and MAPE). Likewise, the EWT-GPR model had smaller evaluation indices compared to the GPR model. The foregoing comparisons demonstrate that the preprocessing methods were effective in boosting the forecasting accuracy of short-term wind speed prediction.

Similar to the prediction to the hourly wind speed series, the prediction results for the half-hour wind speed series also demonstrated the capability of the EWT-GPR method to reliably and accurately predict wind speed compared to the other models. For example, the forecasting results listed in Table 3 show that the EWT-GPR model had the lowest RMSE value (0.78 m/s) and the lowest MAPE value (13.68%) for half-hour ahead wind speed forecasting. In comparison, the PM, AR, CSA-LSSVR, GPR and EWT-CSA-LSSVR approaches resulted in higher RMSE values of 0.84 m/s, 0.82 m/s, 0.80 m/s, 0.81 m/s and 0.80 m/s and higher MAPE values of 15.83%, 14.70%, 14.25%, 14.54% and 14.05%, respectively. In the forecasting for the half-hour wind speed series, the individual GPR performed a little worse than the CSA-LSSVR model, but the hybrid EWT-GPR produced much better forecasting results than those produced by the hybrid EWT-CSA-LSSVR model. This demonstrates that the EWT algorithm is more helpful to the GPR model than CSA-LSSVR approach in improving the wind speed forecasting accuracy. In addition, it can be seen from Fig. 5 that the predicted values from the EWT-GPR approach are consistent with the actual half-hour wind speed series. The detailed comparison proves that the proposed EWT-GPR method has a strong capacity to model the random

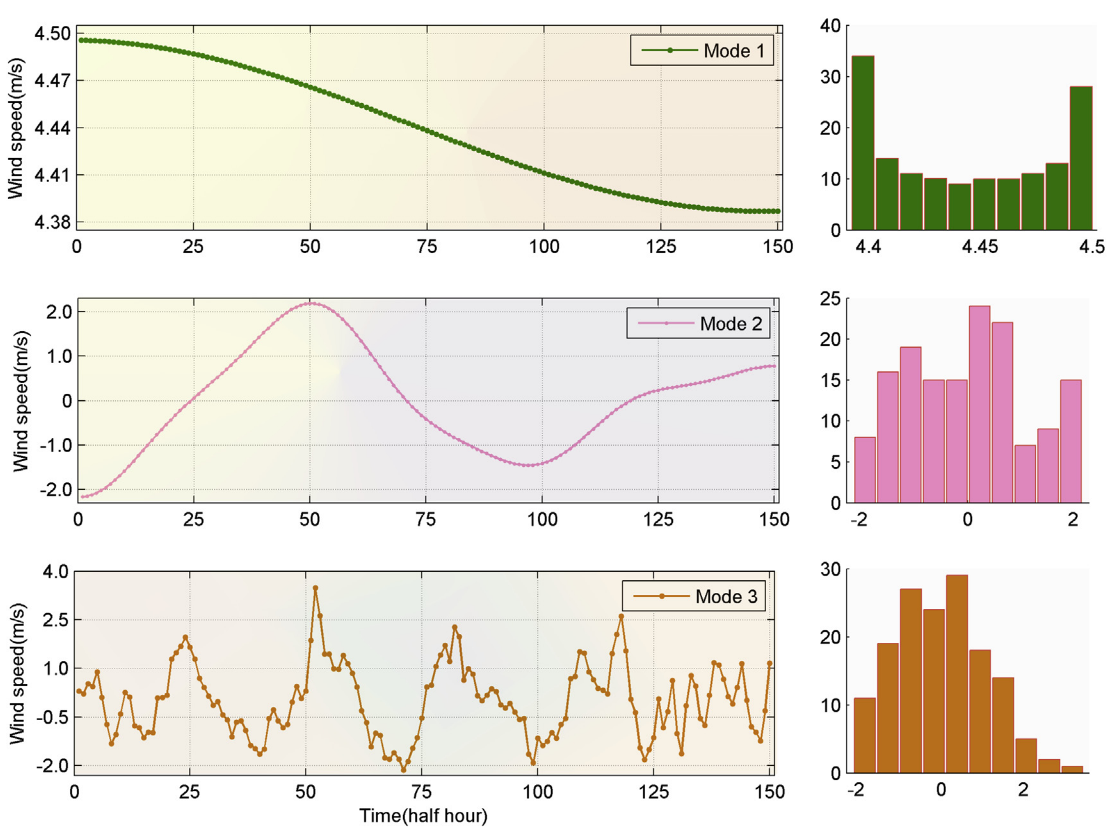


Fig. 6. The decomposed subseries of the half-hour data set by EWT.

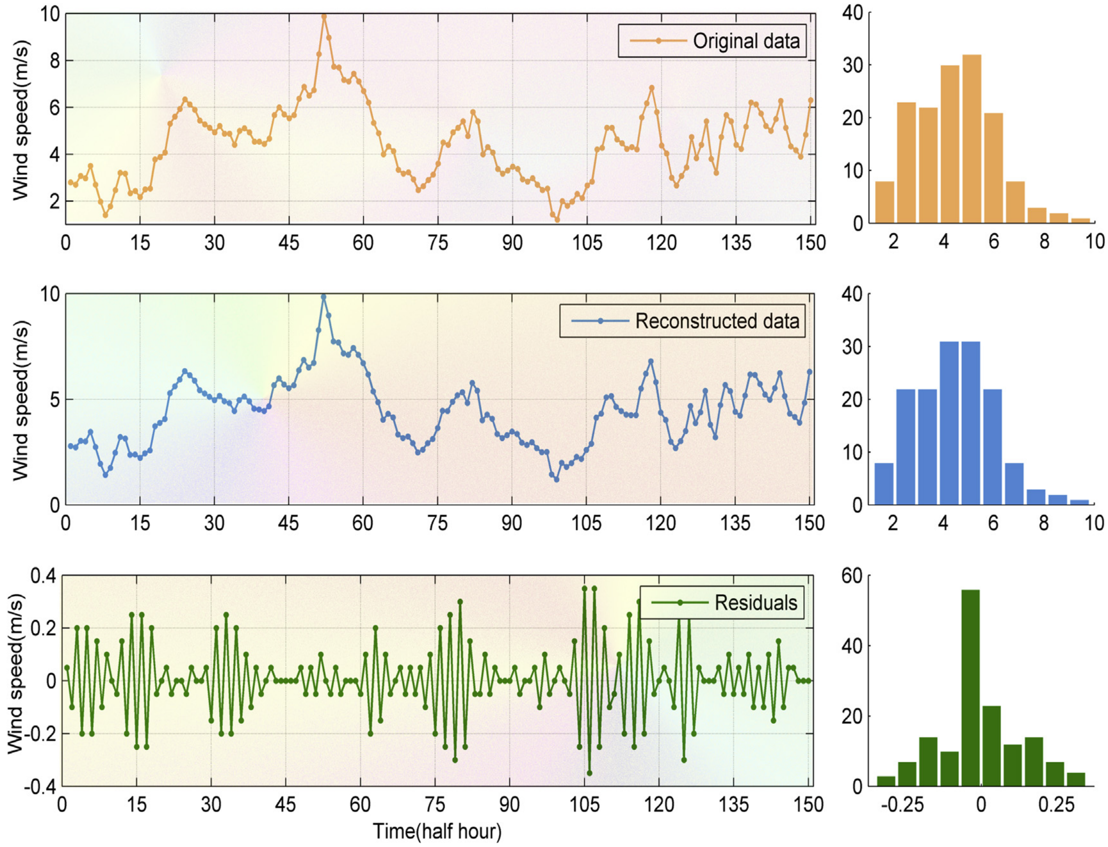


Fig. 7. Comparison of the original half-hour wind speed series and the reconstructed wind and residuals series.

Table 2
Performance evaluations of different models for the hourly wind speed forecasting.

Forecasting horizon	Model	PM	AR	CSA-LSSVR	GPR	EWT-CSA-LSSVR	EWT-GPR
One-hour ahead	MAE (m/s)	1.35	1.34	1.22	1.24	1.16	1.13
	RMSE (m/s)	1.45	1.41	1.36	1.34	1.27	1.22
	MAPE (%)	12.85	11.81	10.91	10.86	10.75	10.18
Two-hour ahead	MAE (m/s)	1.44	1.41	1.36	1.36	1.34	1.32
	RMSE (m/s)	1.48	1.45	1.42	1.41	1.35	1.31
	MAPE (%)	16.36	14.94	14.41	14.46	13.65	13.39
Four-hour ahead	MAE (m/s)	1.64	1.65	1.58	1.56	1.46	1.43
	RMSE (m/s)	1.88	1.79	1.68	1.68	1.65	1.59
	MAPE (%)	21.76	19.81	19.01	18.86	18.45	18.39

dynamics and forecast the wind speed with satisfactory and reliable performance.

The above analysis focuses on the forecasting accuracy of the single-valued predications. Obviously, the PM, AR and EWT-CSA-LSSVR models can only offer point estimation of wind speed given an appropriate input. In addition, during the modeling process of these parametric models, modeling biases exist because parametric models are, at best, only an approximation of the true stochastic dynamics that generate a given dataset [36]. Because the GPR model is a powerful, nonparametric, probabilistic approach, it

Table 3
Performance evaluations of different models for the half-hour wind speed forecasting.

Forecasting horizon	Model	PM	AR	CSA-LSSVR	GPR	EWT-CSA-LSSVR	EWT-GPR
half-hour ahead	MAE (m/s)	0.72	0.69	0.65	0.66	0.64	0.62
	RMSE (m/s)	0.84	0.82	0.80	0.81	0.80	0.78
	MAPE (%)	15.83	14.70	14.25	14.54	14.05	13.68
One-hour ahead	MAE (m/s)	0.82	0.79	0.76	0.77	0.74	0.71
	RMSE (m/s)	0.94	0.91	0.89	0.91	0.87	0.85
	MAPE (%)	19.83	18.70	17.25	17.54	17.05	16.68
Two-hour ahead	MAE (m/s)	1.12	1.07	1.01	1.03	0.94	0.88
	RMSE (m/s)	1.31	1.25	1.12	1.11	1.05	0.98
	MAPE (%)	25.83	23.57	21.25	21.53	19.75	19.68

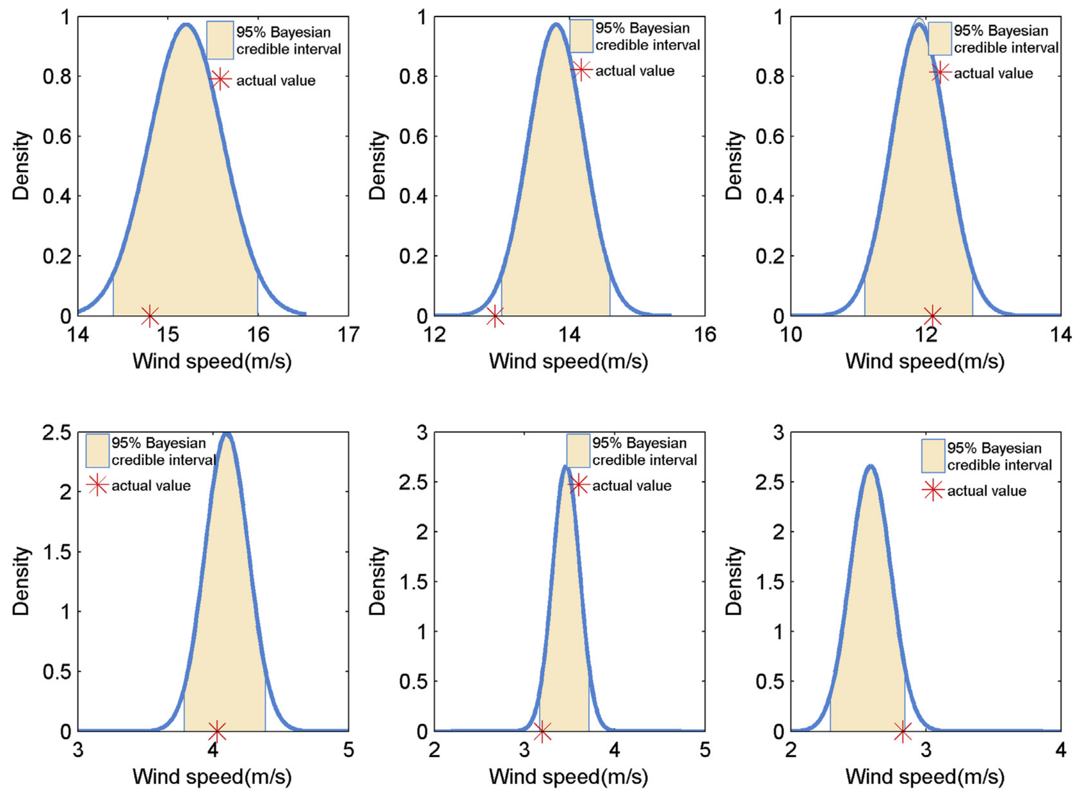


Fig. 8. The predictive pdf of the wind speed for the first, fifth and 20th forecasts. (The top three plots are the predictive pdf of the hourly wind speed with respect to one-hour ahead forecasting, and the below three plots are the predictive pdf of the half-hour wind speed with respect to half-hour ahead forecasting.)

offers a probabilistic predictive distribution and avoids the shortcomings of parametric models. Hence, the hybrid GPR model can provide the most likely value and the probability information corresponding to the forecast based on the predictive probability distribution. Fig. 8 shows the uncertainties of the one-step ahead wind speed forecasts modeled with Gaussian probability density functions and the 95% confidence interval of the one-step ahead wind speed forecast; and Fig. 9 shows the one-step ahead forecasting of the most likely values (mean values) obtained from the suggested model as well as the corresponding 95% confidence intervals. It can be seen from Fig. 9 that most of the actual observations fall in the confidence intervals. Such interval estimation can be incorporated into the decision-making process for production scheduling and control. The GPR forecasts can help reduce or minimize penalties from imbalance charges that are the result of deviations in scheduled output. Such information can also reduce the significant opportunity costs of being too conservative in bidding output into a forward market, due to the uncertainty of availability.

5. Conclusions

With the rapid growth of wind energy, accurate and reliable methods and techniques for short-term wind speed forecasting are urgently needed. Owing to the effect of various environmental factors, wind speed data present high fluctuations, autocorrelation and stochastic volatility, making it difficult to forecast wind speed using a single model. This paper proposed a hybrid forecasting

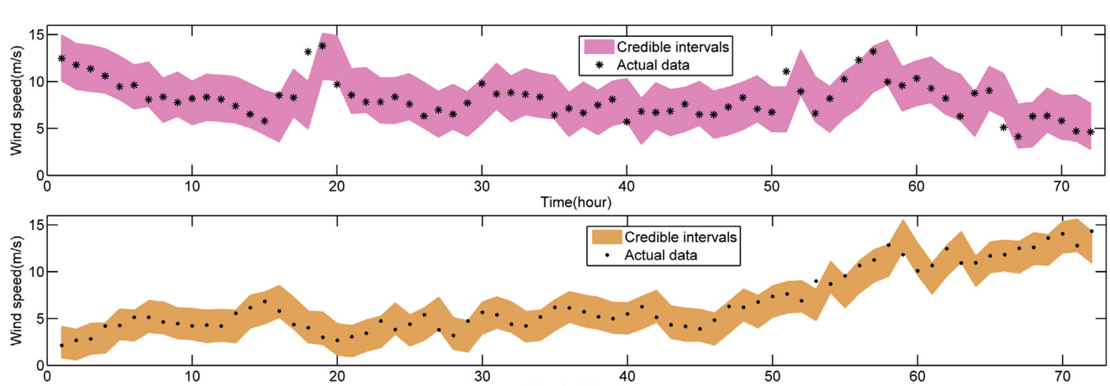


Fig. 9. The 95% confidence intervals of the forecast against the actual data. (The top subplot is the one-hour ahead forecasting for hourly wind speed series and the bottom subplot is the half-hour ahead forecasting for half-hour wind speed series.)

approach based on the concept of hybrid prediction, which integrates the concepts of EWT and GPR. The suggested hybrid model allows prior information about the wind speeds to be incorporated into the model, because the GPR is developed based on Bayesian theory. For the prediction, a moving window is adopted to deal with training data set, thus adapting to the time-varying nature of the wind speed. In addition to point predictions, the hybrid model gives predictive distributions and prediction intervals. The effectiveness of the proposed EWT-GPR model was demonstrated with real mean half-hour wind speed data and hourly wind speed data. The two forecasting test results that were obtained suggest that the developed hybrid wind speed forecasting approach based on the GPR model integrated with the EWT algorithm has the ability to yield good wind speed predictions, as well as interval estimation.

Acknowledgments

This research was supported by the National Natural Science Foundation of China under Grant 71171102/G0107.

References

[1] World Wind Energy Association. World wind energy report. 2013. half-year, http://www.wwindea.org/webimages/Half-year_report_2013.pdf.
[2] Zhang Zhao-Sui, Sun Yuan-Zhang, Cheng Lin. Potential of trading wind power as regulation services in the California short-term electricity market. Energy Policy 2013;59:885–97.
[3] Bian XY, Huang XX, Wong KC, Lo KL, Fu Yang, Xuan SH. Improvement on probabilistic small-signal stability of power system with large-scale wind farm integration. Electr Power Energy Syst 2014;61:482–488s.
[4] Nor Khalid Mohamed, Shaaban Mohamed, Rahman Hasimah Abdul. Feasibility assessment of wind energy resources in Malaysia based on NWP models. Renew Energy 2014;62:147–54.
[5] Cassola Federico, Burlando Massiamiliano. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. Appl Energy 2012;99:154–66.
[6] Erdem Ergin, Shi Jing. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405–14.
[7] Liua Hui, Tian Hong-qi, Li Yan-fei. An EMD-recursive ARIMA method to predict wind speed. J Wind Eng Ind Aerodyn 2015;141:27–38.
[8] Katris Christos, Daskalaki Sophia. Comparing forecasting approaches for Internet traffic. Expert Syst Appl 30 November 2015;42(21):8172–83.
[9] Yang Dazhi, Sharma Vishal, Ye Zhen, Lim Lihong Idris, Zhao Lu, Aryaputera Aloysius W. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy 2015;81:111–9.
[10] Li Yunhua, Ling Lina, Chen Jiantao. Combined grey prediction fuzzy control law with application to road tunnel ventilation system. J Appl Res Technol 2015;13:313–20.
[11] Wang Jujie, Zhang Wenyu, Li Yaning, Wang Jianzhou, Dang Zhangli. Forecasting wind speed using empirical mode decomposition and Elman neural network. Appl Soft Comput 2014;23:452–9.
[12] Ren Chao, An Ning, Wang Jianzhou, Li Lian, Hu Bin, Shang Duo. Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowledge-Based Syst 2014;56:226–39.
[13] Shi Jing, Guo Jinmei, Zheng Songtao. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. Renew Sustain Energy Rev 2012;16:3471–80.
[14] Cao Qing, Ewing Bradley T, Thompson Mark A. Forecasting wind speed with recurrent neural networks. Eur J Oper Res 2012;221:148–54.
[15] Sideratos George, Hatziargyriou Nikos D. Probabilistic wind power forecasting using radial basis function neural networks. IEEE Trans Power Syst 2012;27:1788–96.
[16] López Paz, Maseda Francisco. Wind speed estimation using multilayer perceptron Ramón Velo. Energy Convers Manag 2014;81:1–9.
[17] Liu Da, Niu Dongxiao, Wang Hui, Fan Leilei. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew Energy 2014;62:592–7.
[18] Salcedo-Sanz Sancho, Ortiz-García Emilio G, Pérez-Bellido Ángel M, Portilla-Figueras Antonio, Prieto Luis. Short term wind speed prediction based on evolutionary support vector regression algorithms. Expert Syst Appl 2011;38:4052–7.
[19] Hu Jianming, Wang Jianzhou, Zeng Guowei. A hybrid forecasting approach applied to wind speed time series. Renew Energy 2013;60:185–94.
[20] Guo Zhenhai, Zhao Weigang, Lu Haiyan, Wang Jianzhou. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renew Energy 2012;37:241–9.
[21] Amjady Nima, Keynia Farshid, Zareipour Hamidreza. Short-term wind power forecasting using ridgelet neural network. Electr Power Syst Res 2011;81:2099–107.
[22] Osorio GJ, Matias JCO, Catalao JPS. Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information. Renew Energy 2015;75:301–7.
[23] Alessandrini S, Sperati S, Pinson P. A comparison between the ECMWF and COSMO Ensemble Prediction Systems applied to short-term wind power forecasting on real data. Appl Energy 2013;107:271–80.
[24] Li Xue, Cao Jia, Du Dajun. Probabilistic optimal power flow for power systems considering wind uncertainty and load correlation. Neurocomputing 2015;148:240–7.
[25] Aien Morteza, Rashidinejad Masoud, Firuz-Abad Mahmud Fotuhi. Probabilistic optimal power flow in correlated hybrid wind-PV power systems: a review and a new approach. Renew Sustain Energy Rev 2015;41:1437–46.
[26] Pinson P. Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions. J R Stat Soc Ser C-Appl Stat 2012;61:555–76.
[27] Kou Peng, Gao Feng, Guan Xiaohong. Sparse online warped Gaussian process for wind power probabilistic forecasting. Appl Energy 2013;108:410–28.
[28] Jiang Yu, Song Zhe, Kusiak Andrew. Very short-term wind speed forecasting with Bayesian structural break model. Renew Energy 2013;50:637–47.
[29] Bessa R, Miranda V, Botterud A, Zhou Z, Wang J. Time-adaptive quantile-copula for wind power probabilistic forecasting. Renew Energy 2012;40:29–39.
[30] Baran Sándor. Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components. Comput Stat Data Anal 2014;75:227–38.
[31] Gilles Jérôme. Empirical wavelet transform. IEEE Trans Signal Process 2013;61:3999–4010.
[32] Shayeghi H, Ghasemi A. Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. Energy Convers Manag 2013;74:482–91.
[33] Catalão JPS, Pousinho HMI, Mendes VMF. Short-term wind power forecasting in Portugal by neural networks and wavelet Transform. Renew Energy 2011;36:1245–51.
[34] Hu Jianming, Wang Jianzhou, Ma Kailiang. A hybrid technique for short-term wind speed prediction. Energy 2015;81:563–74.
[35] Vanhatalo Jarno, Riihimäki Jaakko, Hartikainen Jouni, Jylänki Pasi, Tolvanen Ville, Vehtari Aki. GPstuff: Bayesian modeling with Gaussian processes. J Mach Learn Res 2013;14:1175–9.
[36] Fan Jianqing, Yao Qiwei. Nonlinear time series: nonparametric and parametric methods. 2nd ed. New York: Springer-Verlag; 2005.