Research Article

Multistep Wind Speed Forecasting Based on Wavelet and Gaussian Processes

Niya Chen,[1] Zheng Qian,[2] and Xiaofeng Meng[2]

[1] *Key Laboratory of Precision Opto-Mechatronics Technology, Ministry of Education, Beihang University, Beijing 100191, China*
[2] *School of Instrument Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China*

Correspondence should be addressed to Niya Chen; pekiny@aspe.buaa.edu.cn

Accurate wind speed forecasts are necessary for the safety and economy of the renewable energy utilization. The wind speed forecasts can be obtained by statistical model based on historical data. In this paper, a novel W-GP model (wavelet decomposition based Gaussian process learning paradigm) is proposed for short-term wind speed forecasting. The nonstationary and nonlinear original wind speed series is first decomposed into a set of better-behaved constitutive subseries by wavelet decomposition. Then these subseries are forecasted respectively by GP method, and the forecast results are summed to formulate an ensemble forecast for original wind speed series. Therefore, the previous process which obtains wind speed forecast result is named W-GP model. Finally, the proposed model is applied to short-term forecasting of the mean hourly and daily wind speed for a wind farm located in southern China. The prediction results indicate that the proposed W-GP model, which achieves a mean 13.34% improvement in RMSE (Root Mean Square Error) compared to persistence method for mean hourly data and a mean 7.71% improvement for mean daily wind speed data, shows the best forecasting accuracy among several forecasting models.

1. Introduction

Wind power is one of the fastest-developing renewable energy sources of which the current total capacity around the world is approximately 282 gig watts (GW) till the end of 2012, with a growing rate around 20% [1]. However, the variable and uncontrollable characteristics of wind pose several operational challenges. Thus, the wind power forecasting is an essential process for the wind farm units' maintenance and energy reserves scheduling [2, 3].

Since wind power is a function of wind speed, the wind power forecasts basically depend on wind speed forecasts. The short-term wind speed forecasting, of which the prediction horizon is from 1 hour to 3 days, is critical to minimize scheduling errors which impact grid reliability and market based ancillary service costs. Broadly speaking, there are two statistical approaches for short-term wind speed prediction: regression models contain Numerical Weather Prediction (NWP) data as inputs and time series based methods which only use historical data to obtain prediction results. The former downscales information from global

meteorological model to wind farm location therefore got inputs to regression model which estimate the future wind and have advantages in multihour (from several hours to dozens of hours) ahead prediction [4–6], while the latter use only historical wind speed data to build models and perform better in multistep ahead (usually 1–4 hours ahead for mean hourly wind data or several days ahead for mean daily data) prediction. Since the combination of NWP data requires long computational time (usually several hours) and does not show advantage in multistep ahead forecasts, models built only based on historical data are preferred for this prediction horizon. Time series based models for multistep wind forecasting have been investigated deeply and developed based on different methodologies, such as autoregressive integrated moving average [7, 8], Kalman filter [9, 10], artificial neural network [11, 12], and support vector machine methods [13].

Recently, wavelet decomposition method has been applied to establish different hybrid wind speed forecasting models. The main contribution of wavelet transform is to decompose and reconstruct a wind speed series into a set

of better-behaved constitutive series. Then each sub-series can be separately predicted by a suitable model according to its feature; hence, the new hybrid model improves the forecasting accuracy. Wavelet combined methods can be found in [14–16]. Liu et al. proposed a hybrid method based on the methods of wavelet and classical time series analyses to predict short-term wind speed and wind power, which gained more accurate simulation results than classical time series method and BP network method, especially with multistep forecasting and jumping data [14]. Catalão et al. presented a novel forecasting method by combining artificial neural networks with wavelet transform. Results from a real-world case study, which used wind speed data in Portugal, show the efficiency of this model [15]. An et al. proposed a prediction model for wind farm power forecasting by combining the wavelet transform, chaotic time series, and GM(1,1) method [16].

As an effective statistical method, Gaussian Process (GP) has been applied broadly in many domains, including wind energy prediction. Jiang et al. focused on very short-term (<30 min) wind speed prediction using GPs [17]. They evaluated their model on real-world datasets and found that the GP performs better than ARMA (a simpler variant of ARIMA) and Mycielski algorithms [18].

In this paper, a novel hybrid forecasting approach is proposed based on wavelet method and Gaussian Processes (GPs) for multiple steps ahead wind speed prediction. Compared with earlier work, this paper has the following contributions. First, the novel combination of wavelet method and GP (W-GP model) managed to improve forecasting accuracy, especially when forecast step grows. Second, lots of simulation work was done to determine the best W-GP model by comparing the forecasting error of several W-GP models decomposed by different levels. Third, not only mean hourly wind speed but also mean daily wind speed is forecasted by proposed model in this paper.

Actual wind speed data of 1 year from a wind farm in southern China is used to examine the proposed W-GP model. The proposed model is compared with persistence, MLP (Multilayer Perceptron) neural network, and the original GP approaches to demonstrate its effectiveness regarding forecast accuracy. The forecasting results are given and discussed hereinafter.

This paper is organized as follows. Section 2 presents the proposed W-GP approach to forecast wind speed. Section 3 first presents a case study of detailed forecast process based on W-GP model and then introduces three different criterions used to evaluate the forecasting accuracy, based on which the most suitable model for our database is finally chosen. Section 4 presents the large-scale simulation results for a real wind speed data set from a wind farm in southern China. Finally, Section 5 outlines the conclusions.

2. The Proposed W-GP Model

The proposed W-GP approach to forecast short-term wind speed is based on the hybrid of GP with wavelet method. The wavelet method is used to decompose the original wind

speed series into a set of sub-series which can be analyzed easier. Then, GP method is used to forecast the future values for all those sub-series. In turn, through the inverse wavelet decomposition, finally the wind speed forecast value can be obtained by aggregating the forecast value of sub-series.

2.1. Wavelet Method. The wavelet method used here is to decompose a wind speed series into a set of sub-series. With the filtering effect of the wavelet decomposition, these sub-series present a better behaviour than the original wind power series and therefore can be analyzed clearer and predicted more accurately.

Wavelet method can be divided in two categories: continuous wavelet transform (CWT) and discrete wavelet transform (DWT) [19]. In CWT, a wavelet can be defined as function $\psi(t)$ with a zero mean:

$$\int_{-\infty}^{+\infty} \psi(t) \, dt = 0. \tag{1}$$

A signal can be decomposed into many series of wavelets with different scales a and translation b:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right). \tag{2}$$

Thus, the wavelet transform of a signal $f(t)$ at translation b and scale a is defined by

$$wf(b,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \, \psi\left(\frac{t-b}{a}\right) dt. \tag{3}$$

The original signal $f(t)$ can be reconstructed by inverse wavelet transform:

$$f(t) = \int_{0}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{a^2} wf(b,a) \, \psi_{b,a}(t) \, db \, da. \tag{4}$$

Different from CWT, when the mother wavelet is scaled and translated using certain scales and positions, it is known as the DWT, which is more efficient and just as accurate as the CWT. The definition is as follows:

$$W(m,n) = 2^{-(m/2)} \sum_{t=0}^{T-1} f(t) \, \phi\left(\frac{t-n\cdot 2^m}{2^m}\right), \tag{5}$$

where T is the length of the signal $f(t)$. The functions of the integer variables m and n $(a = 2^m, b = n \cdot 2^m)$ are scaling and translation parameters, and t is the discrete time index.

2.2. Gaussian Process. Recently, there has been much activity concerning the application of Gaussian process to machine learning tasks. The systematic and detailed explanation of Gaussian process regression can be found in Rasmussen's book [20]; here we only provide a brief illustration on GP applied in regression.

A Gaussian process $f(x)$ can be completely specified by its mean function and covariance function, written as $f(x) \sim \text{GP}(m(x), k(x, x'))$, where the mean function $m(x)$ and covariance function $k(x, x')$ are defined as follows:

$$m(x) = E[f(x)]$$

$$k(x, x') = E[(f(x) - m(x))(f(x) - m(x'))]. \tag{6}$$

Usually, the mean function is assumed to be zero, and the target variables are normalised to have zero mean.

Consider GP in a classic regression problem, assume x_i represents a feature vector of input space with dimension N, y_i represents the output value to be estimated, based on the training vectors' set $D = \{(x_i, y_i), i = 1, \ldots, n\}$, and the key point of GP regression method is to model the relationship between inputs and targets, that is, to build a function to satisfy

$$y_i = f(x_i) + \varepsilon_i. \tag{7}$$

The observed value y is assumed to be different from the function value $f(x)$ by additive noise ε, which is assumed to be an independent and identically distributed Gaussian distribution with zero mean and variance σ_n^2, that is, $\varepsilon \sim N(0, \sigma_n^2)$.

Note that y is a linear combination of Gaussian variables and, hence, is itself Gaussian. The prior on y becomes as follows:

$$
\begin{aligned}
E[y] &= E[f + \varepsilon] = 0, \\
\text{cov}[y] &= K(X, X) + \sigma_n^2 I,
\end{aligned}
\tag{8}
$$

where K is a matrix with elements $K_{ij} = k(x_i, x_j)$, which is also known as the kernel function.

Given a training set $D = (X, y)$, our goal is to make predictions of the target variable f_* for a new input x_*. Since we already have $p(y \mid X, k) = N(0, K + \sigma_n^2 I)$, the distribution with new input can be written as follows:

$$
\begin{bmatrix} y \\ f_* \end{bmatrix} \sim \left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & k(X, x_*) \\ k(x_*, X) & k(x_*, x_*) \end{bmatrix} \right),
\tag{9}
$$

where $k(X, x_*) = [k(x_1, x_*), \ldots, k(x_n, x_*)]$, and it can be shortened to k_*. Then according to the principle of joint Gaussian distribution, the prediction result of target is given by

$$
\begin{aligned}
\overline{f}_* &= k_*^T (K + \sigma_n^2 I)^{-1} y \\
V[f_*] &= k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_*.
\end{aligned}
\tag{10}
$$

Now, the whole regression model based on Gaussian process is completed.

In the GP model for wind speed forecasting, we define the wind speed series as $\{d_n, n = 1, \ldots\}$, then the input vector x_i is constructed as $\{d_i, d_{i+1}, \ldots, d_{i+t-1}\}$; correspondingly the output value y_i is d_{i+t}. Therefore, the GP model can estimate the wind speed value at next time period based on historical data.

2.3. Process of W-GP Modelling. Most researches use symmetric WTs such as Symlet or Morlet for decomposition. However, this type of WT is not suitable for forecasting problem because in symmetric wavelet future information is also needed as well as previous information [19]. In this paper, a wavelet function of type Daubechies of order 4 (abbreviated
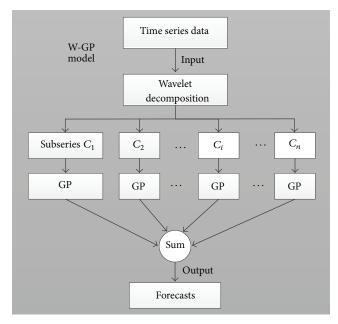

FIGURE 1: Structure of proposed W-GP model.

TABLE 1: Information of wind farm and its wind turbines.

Location	Installed capacity	Height	Number
Fujian	2000 kW	80 m	15

as Db4), which is an asymmetric WT, is used to solve this problem.

As shown in Figure 1, the modelling steps of the proposed W-GP method are described as follows.

(1) Use wavelet method to decompose original wind speed data series into a number of different sub-series (depending on level of decomposition) which can be analyzed and separately predicted. Denote these sub-series as $C_1, \ldots, C_n$.

(2) Build the prediction models for each sub-series based on GP method and estimate the multistep forecasting results.

(3) Through the inverse wavelet decomposition, attain final forecasting results for original wind speed series by aggregating the forecasting value of sub-series.

(4) Calculate the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE) of forecasting results.

3. A Case Study of W-GP Model

A real-world dataset based on wind farm is used in this paper to evaluate our approach. The dataset is from a wind farm in Fujian province, a coastal area located in very southern China, where wind source is sufficient and variable, and the integration of wind farm is important.

As shown in Table 1, the tested wind farm contains 15 wind turbines, of which installed capacity is 2000 kW and
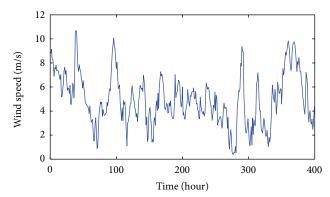
FIGURE 2: Original time series of wind speed.

turbine height is 80 m. With wind speed being measured consistently on different heights (70 m, 35 m, and 10 m) by a wind tower located in the centre of the wind farm, the data from the 70 m high sensor is chosen for calculation, since it is the nearest to turbine height. The original measured data has an interval of 10 mins, which we averaged into mean hourly and daily data for different forecast targets.

To illustrate the specific process of modelling and analyze the performance of W-GP model, in this section, we use the 1st-400th hours' data of each month as training set to build the model, the later ones as test set, and we obtain multihour ahead forecasting results.

3.1. Wavelet Decomposition.

According to Section 2, if Daubechies-4 wavelet is employed to do n level discrete wavelet decomposition in simulation, the original wind speed series would be decomposed to $n + 1$ sub-series: one low-frequency section and other n high-frequency series. To facilitate the latter modelling, the low-frequency section is recorded as $\{x_{1t}\}$ series; correspondingly, the other high-frequency sub-series are recorded as $\{x_{2t}\} \cdots \{x_{(n+1)t}\}$.

Here, we decompose the actual wind speed data in February 2012 to the 3rd level as an example. The original wind speed series is shown in Figure 2, while the sub-series ($\{x_{1t}\}, \ldots, \{x_{4t}\}$) are plotted in Figure 3.

As shown in the figures previously, the original wind speed series is decomposed into a set of better-behaved constitutive series. Therefore, it is easier for sub-series to obtain better performance in forecasting and eventually get results with higher accuracy.

3.2. Wind Speed Forecasting.

According to Section 2.3, after obtaining the sub-series by wavelet method, the prediction models for each sub-series should be built based on GP method. Then the final forecast value of original wind speed series can be attained by aggregating the forecasting results of sub-series.

Consequently, we calculate 1-hour-ahead forecasting based on GP model for $x_{1t}(1 \cdots 400)$, the result is recorded as $\hat{x}_{1t}(1)$. Similarly, and we obtain 1-hour-ahead forecasting results of, $\{x_{2t}\} \cdots \{x_{4t}\}$, recorded as $\hat{x}_{2t}(1), \ldots, \hat{x}_{4t}(1)$. Then we aggregate the forecast value of sub-series by

$$\hat{x}_t(1) = \hat{x}_{1t}(1) + \hat{x}_{2t}(1) + \cdots + \hat{x}_{4t}(1), \tag{11}$$

where $\hat{x}_t(1)$ represents the forecast result of the original wind speed series; thus, the final forecasting value is attained.

Correspondingly, calculate 1-hour-ahead forecasting for $x_t(2 \cdots 401), \ldots, x_t(60 \cdots 459)$; these 60 forecast values are shown in Figure 4(a), plotted by green star line.

Meantime, the basic GP model is applied on the same original wind speed data to obtain 60 forecast values. By subtracting from the actual wind speed value, the absolute value of forecasting error can be calculated. The comparison of the GP model and W-GP model's forecasting errors at each hour is shown in Figure 4(b).

As shown in Figure 4, it can be observed that W-GP model performs a better wind speed forecasting than the basic GP model at most of the time.

Since now we have one-hour-ahead forecasting value, it could be useful in the case of attaining several-hour-ahead forecasting value, based on $\{x_{1t}(2), \ldots, x_{1t}(400), \hat{x}_{1t}(1)\}$ series to rebuild the GP model and calculate 1-hour-ahead forecasting again, that is, the 2-hour-ahead forecasting for $\hat{x}_{1t}(1 \cdots 400)$, recorded as $\hat{x}_{1t}(2)$. Similarly obtain 3-hour-ahead forecasting results based on $\{x_{1t}(3), \ldots, x_{1t}(400), \hat{x}_{1t}(1), \hat{x}_{1t}(2)\}$ series and multi-hour-ahead forecasting likewise.

The performance of multihour forecasting by proposed W-GP model can be observed in Figure 5, of which the upper part displays 2-hour-ahead forecasting curve, while the lower part presents the corresponding forecast error (absolute value).

It is clearly seen that at most of the forecast points, the W-GP model decreases forecast error of GP model, even more obvious than 1-hour-ahead prediction displayed previously. Therefore, it is reasonable to say that the hybrid W-GP model is effective and applicable in wind speed forecasting problem and has greater advantage when forecast step grows.

3.3. Forecasting Accuracy Evaluation.

Clearly, accuracy is the most important criterion to compare the efficiency of alternative forecasting approaches. Therefore, different criteria are used here to evaluate the accuracy of the proposed approach. This accuracy is computed in function of the actual wind speed that occurred.

Three forecast error measures were employed for model evaluation and model comparison: the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE). The error is defined as follows:

$$\begin{aligned} e_t &= y_t - \hat{y}_t, \\[2mm] \text{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}, \\[2mm] \text{MAE} &= \frac{1}{n}\sum_{i=1}^{n} |e_i|, \\[2mm] \text{MAPE} &= \frac{1}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \times 100 \right|, \end{aligned} \tag{12}$$

(a) Sub-series x_{1t}

(b) Sub-series x_{2t}

(c) Sub-series x_{3t}

(d) Sub-series x_{4t}
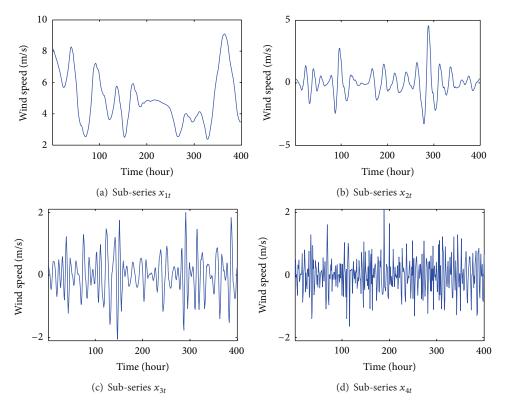
Figure 3: Decomposition results of the original series.
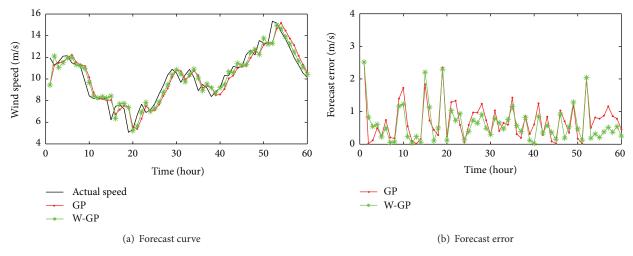

(a) Forecast curve

(b) Forecast error

Figure 4: One-hour-ahead forecast results.

where y_t represents the actual observation value at hour t and $\widehat{y}_t$ represents the forecast value for the same period. n is the number of forecasted hours.

3.4. Comparison of Different Decomposition Levels.

Currently, there is still no specific principle of determining wavelet decomposition level of W-GP model yet, though various decomposition levels of proposed forecast model

may lead to different forecasting performances. Therefore, we should recognise the most adequate decomposition level for our database by analyzing simulation results.

As shown in Figure 6, it is obvious that the forecasting error is highly relevant to decomposition level. Furthermore, the error is the minimum while the discrete wavelet decomposition level of Daubechies-4 wavelet is 3. Therefore, in this paper, the 3 level decomposed W-GP model is applied through the whole simulation process.
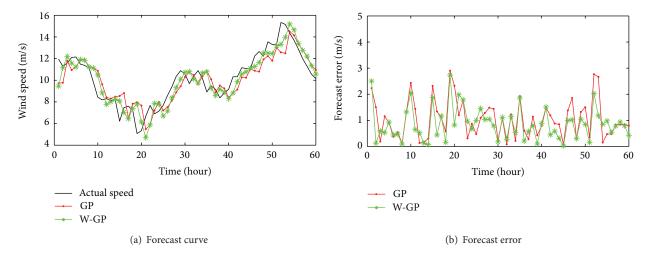
(a) Forecast curve

(b) Forecast error

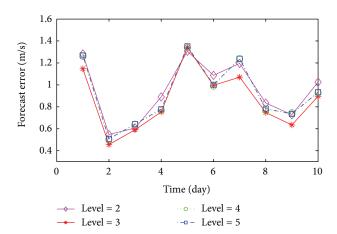Figure 5: Two-hour-ahead forecast results.

Figure 6: One-hour-ahead forecast error (averaged by 24 hour).

4. Wind Speed Forecasting Results

The actual measured wind speed data from a wind farm in southern China through a whole year has been applied in simulation. Since sometimes the wind speed data may be invalid because of anemometer fault, a preprocess was taken to modify invalid speed data by interpolation, after which the whole dataset contains 2 sub-sets: D1, 8700 points of mean hourly data and D2, 362 points of mean daily data. In order to reasonably testify the performance of proposed model, the sub-set D1 was divided into 12 parts; each part was separately modeled, using the first 400 data as training set and the rest as test set. Similarly, the sub-set D2 was divided into 2 parts, using the first 160 data of each part as training set, with the rest as test set. The short-term forecasting results, of both mean hourly and mean daily wind speed, are presented and analyzed in the following.

Besides the basic GP method, the other models represented as comparison in this paper are persistence and MLP methods. Persistence method is a quite simple method which only uses the current value as forecast and is impressively

Table 2: Comparison of forecast accuracy for mean hourly wind speed.

	Number of forecasting steps (hour)				
	1	2	3	4	Average
RMSE (m/s)					
Persistence	1.0028	1.5272	1.9233	2.2231	1.6691
MLP	1.0075	1.5697	1.8911	2.1545	1.6557
GP	1.0305	1.4741	1.8219	2.1572	1.6209
W-GP	0.9605	1.2672	1.6041	1.9539	1.4464
MAE (m/s)					
Persistence	0.7431	1.1399	1.5271	1.7293	1.2849
MLP	0.7647	1.2168	1.4515	1.7095	1.2856
GP	0.7675	1.1280	1.4200	1.7035	1.2548
W-GP	0.7204	0.9565	1.2093	1.5320	1.1046
MAPE (%)					
Persistence	11.0534	16.9723	26.3192	26.7187	20.2659
MLP	11.9014	19.0887	22.3889	27.7471	20.2815
GP	11.2190	17.4187	23.1315	28.4284	20.0494
W-GP	11.2430	14.3263	18.3666	23.9891	16.9812

effective for short-term prediction and therefore is considered as the most classical benchmark in wind forecasting area. The MLP network is a very popular machine learning method and has been applied in wind power forecasting widely. We established MLP model based on the same wind speed dataset and obtained simulation results to compare with the proposed W-GP model.

According to Section 3.3, the evaluation index applied in this paper is the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE). The comparison among the forecasting accuracy of persistence method, MLP, basic GP, and W-GP models is shown in Tables 2 and 3.

Table 2 shows the forecasting errors of persistence method, MLP, basic GP, and proposed W-GP models for mean hourly wind speed data. It can be observed that though

TABLE 3: Comparison of forecast accuracy for mean daily wind speed.

	Number of forecasting steps (day)			
	1	2	3	Average
RMSE (m/s)				
Persistence	1.6069	2.2461	2.2985	2.0505
MLP	2.2651	2.5642	2.4331	2.4208
GP	1.9024	2.1317	2.0892	2.0411
W-GP	1.6110	2.0241	2.0421	1.8924
MAE (m/s)				
Persistence	1.2979	1.8150	1.8368	1.6499
MLP	1.6953	1.9832	2.0775	1.9187
GP	1.4838	1.7028	1.7227	1.6364
W-GP	1.2775	1.6355	1.6357	1.5162
MAPE (%)				
Persistence	29.9989	49.0180	42.9931	40.6700
MLP	47.7015	52.4184	56.1476	52.0892
GP	40.8972	38.6730	47.5696	42.3799
W-GP	30.2413	39.4201	44.0601	37.9072

forecast performance gets worse when prediction time grows, normally basic GP method shows a better forecast accuracy than MLP model, and comparing with GP method, the proposed W-GP model presents positive error improvements over the entire forecast horizon:

$$\text{Improvement} = \frac{\left(e_{\text{Persistence}} - e_{\text{proposed}}\right)}{e_{\text{Persistence}}} \times 100\%. \quad (13)$$

With respect to the basic persistence method, the improvement of W-GP model in RMSE, computed by formula (13), ranges from a maximum of 17.02% to a minimum of 4.22% with a mean value of 13.34%.

Table 3 shows the forecasting errors of the same models for mean daily wind speed data. Here, the MLP model works the worst, which is understandable because of the small scale of training set. Fortunately, Gaussian process based models maintain a stable performance without needing large-scale dataset to train appropriate parameters that is why the proposed model still shows an obvious advantage than the other forecasts. Comparing with the basic persistence method, the improvement of W-GP model in RMSE, computed by formula (13), ranges from a minimum of -0.26% to a maximum of 11.16% with a mean value of 7.71%. Although for 1-day-ahead forecast, the W-GP model is slightly worse than persistence method, the other results with larger forecast step strongly reveal the efficiency of proposed W-GP model.

Since the wavelet method decomposes the original wind speed series into a set of better-behaved constitutive series, the proposed W-GP model achieves a higher level of accuracy at short-term wind speed forecast. What is more, the improvement approached by wavelet decomposition is getting more obvious as forecast step grows. However, as the prediction time grows, the forecast accuracy of each model decreases severely. Anyway, it is completely natural that models based only on historical data would behave this way considering the uncontrollable and unstable inhesion of wind.

Though due to the variable nature of wind, the forecast performance of all the models listed in the tables decade as the predict time grows, the proposed W-GP model continuously shows a better forecasting accuracy and, eventually, represents an obvious advantage.

5. Conclusion

In this paper, first, a novel historical forecast model is proposed based on the wavelet method and Gaussian Process (GP) method in order to predict multistep ahead wind speed. Second, based on analysis of W-GP model's forecasting error series, an appropriate level of wavelet decomposition is chosen to get the most accurate model. Finally, real-world dataset is applied with the proposed model to validate its efficiency.

The simulation results convincingly reveal the effectiveness and accuracy of proposed model for short-term wind speed forecasting, which achieves a mean 13.34% improvement in RMSE comparing to persistence method for mean hourly data and a mean 7.71% improvement for mean daily wind speed data.

Considering the unstability of wind, there must be a limit on single historical methodology. A tendency of future method is combinatorial models, of which the effective way of combination between methods is worth more research.
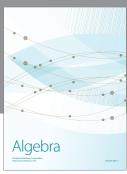
Acknowledgment

References

[1] Global Wind Energy Council, Global wind statistics 2012, http://www.gwec.net/wp-content/uploads/2013/02/GWEC-PRstats-2012_english.pdf.

[2] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 6, pp. 1725–1744, 2008.

[3] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.

[4] M. G. De Giorgi, A. Ficarella, and M. Tarantino, "Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods," *Energy*, vol. 36, no. 7, pp. 3968–3978, 2011.

[5] J. Ignacio, L. Alfredo, M. Claudio et al., "Comparison of two new short-term wind-power forecasting systems," *Renewable Energy*, vol. 34, no. 7, pp. 1848–1854, 2009.

[6] S. Salcedo-Sanz, A. M. P. Ángel M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and D. Paredes, "Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction," *Renewable Energy*, vol. 34, no. 6, pp. 1451–1457, 2009.
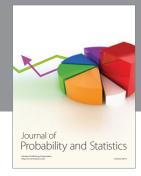
[7] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Applied Energy*, vol. 88, no. 4, pp. 1405–1414, 2011.

[8] H. Liu, E. Erdem, and J. Shi, "Comprehensive evaluation of ARMA-GARCH approaches for modeling the mean and volatility of wind speed," *Applied Energy*, vol. 88, no. 3, pp. 724–732, 2011.

[9] D. Pan, H. Liu, and Y. Li, "Optimization algorithm of short-term multi-step wind speed forecast," *Proceedings of the Chinese Society of Electrical Engineering*, vol. 28, no. 26, pp. 87–91, 2008 (Chinese).

[10] P. Louka, G. Galanis, N. Siebert et al., "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348–2362, 2008.

[11] E. Cadenas and W. Rivera, "Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks," *Renewable Energy*, vol. 34, no. 1, pp. 274–278, 2009.

[12] S. A. P. Kani and M. M. Ardehali, "Very short-term wind speed prediction: a new artificial neural network-Markov chain model," *Energy Conversion and Management*, vol. 52, no. 1, pp. 738–745, 2011.

[13] S. Sancho, G. Emilio, M. Angel et al., "Short term wind speed prediction based on evolutionary support vector regression algorithms," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4052–4057, 2011.

[14] H. Liu, H. Tian, C. Chen, and Y. Li, "A hybrid statistical method to predict wind speed and wind power," *Renewable Energy*, vol. 35, no. 8, pp. 1857–1861, 2010.

[15] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-term wind power forecasting in Portugal by neural networks and wavelet transform," *Renewable Energy*, vol. 36, no. 4, pp. 1245–1251, 2011.

[16] X. An, D. Jiang, C. Liu, and M. Zhao, "Wind farm power prediction based on wavelet decomposition and chaotic time series," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11280–11285, 2011.

[17] X. Jiang, B. Dong, L. Xie, and L. Sweeney, "Adaptive Gaussian process for short-term wind speed forecasting," in *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI '10)*, H. Coelho, R. Studer, and M. Wooldridge, Eds., Lisbon, Portugal, August 2010.

[18] F. O. Hocaoğlu, M. Fidan, and O. N. Gerek, "Mycielski approach for wind speed prediction," *Energy Conversion and Management*, vol. 50, no. 6, pp. 1436–1443, 2009.

[19] H. T. Nguyen and I. T. Nabney, "Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models," *Energy*, vol. 35, no. 9, pp. 3674–3685, 2010.

[20] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Mass, USA, 2006.
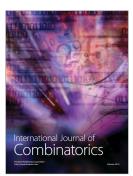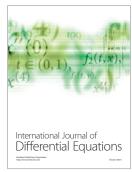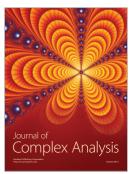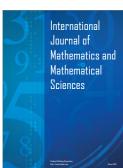
Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Hindawi

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

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization