

ORIGINAL RESEARCH

Open Access



Time series decomposition model for accurate wind speed forecast

V. Prema* and K. Uma Rao

Abstract

Climate change can be considered to be the greatest environmental challenge our world is facing today. Along with the need to ensure long-term assurance of energy supply, it imposes an obligation on all of us to consider ways of reducing our carbon footprint and sourcing more of our energy from renewable sources. Wind energy is one such source and forecasting methods for the prediction of wind speed are becoming increasingly significant due to the penetration of wind power as an alternative to conventional energy sources. This paper proposes time series models for short-term prediction of wind speed. The predictions are done for 1 day ahead using different time series models. For each model, these predicted values are compared with the actual values for the next day. Basic exponential smoothing for different duration of data was tested. A hybrid model with decomposition and exponential smoothing is proposed. A multiplicative decomposition is carried out for the measured data. Separate models were developed for seasonal and trend series and then combined to carry out the forecast. The models were tested for different durations of samples and different weather conditions. It is observed from the results that the prediction with decomposition model for 4 months data gave the least error.

Keywords: Wind speed, Forecast, Decomposition, Time series

Background

Intermittency of wind is the biggest challenge in employing wind energy as a reliable autonomous source of electric power. Energy crisis, global warming, depletion of fossil fuels are the major threats looming the world today. Adequate utilization of renewable energy sources such as wind, solar and biomass proves to be the only alternative to overcome these problems. Because of the indeterminate nature and complementary behavior of wind and solar energy, many researchers are focusing their efforts towards hybrid energy systems. A hybrid energy system uses a combination of sources such as wind and solar along with a battery and Diesel Generator set. Optimal allocation of the available resources is one of the challenges faced in the design of hybrid energy systems. For a well-planned system, it is essential to have accurate predictive models of the wind and solar sources. Wind energy is directly dependent on the wind speed available

at that location and this speed is highly erratic. Hence, a model is needed for accurate prediction of wind speed.

Since the generation of power from wind energy is very erratic due to its heavy dependence on weather, seasonal changes, geographical location, time of the day, direction of wind speed, etc., the forecasting methods may not give uniformly efficient results for all regions. Hence, there is a need to critically examine the seasonality and the nature of the data to determine the models that can be used for prediction of wind speed. There are many methods available in literature for prediction such as time series models, regression models, models based on Artificial Neural Network and Fuzzy logic.

Zhao et al. (2011) and Wang et al. (2011) give thorough review on various methods of wind power forecasting. The authors also propose the development direction of forecast. It is mentioned that combining physical and statistical forecasting models can improve prediction. Erdem et al. (2014), a mixed ARMA model is proposed which incorporates wind direction also with the model. K-Mean clustering is used to examine the association between wind speed and wind direction. Lv and

*Correspondence: premav@rvce.edu.in
R.V. College of Engineering, Mysore Road, Bangalore, India

Yue (2011), the authors propose a forecast model based on wavelet ARIMA-GARCH model. The input data are decomposed using wavelets to remove outliers and then modeled with ARIMA-GARCH model. Liu et al. (2011), the time series mean and volatility of wind speed is predicted with ARMA-GARCH model. Further, the wind power expected output equation is obtained using operational probability. Li et al. (2010) proposes a 2-step forecast technique based on Bayesian algorithm and neural network. Results were obtained for 2 different sites and the MAPE is in the range of 14–18 %.

In Wang et al. (2012), a new multi-variate Least Square-Support Vector Machine (LS-SVM) model is proposed. The proposed method is compared with the existing models such as ARIMA and neural network models. SVM is a regression model used to find out the mapping between the predicted values and actual values of wind speed. The data are collected from various wind farms for a duration of 1 year. MAPE is calculated and compared. The lowest MAPE is found to be 10.06 %. Sideratos and Hatziargyriou (2007), George Sideratos et al. combine Artificial Intelligence and fuzzy logic techniques and introduce a new forecasting technique. A preliminary forecast is done using NWP and quality is estimated using fuzzy rules. Three models are developed. The preliminary model predicts wind power for the next hour using neural network whose inputs are wind speed measurements from a wind farm, wind power prediction from NWP and hour of the day for which prediction is made. The maximum error obtained is 40–50 %. This output is fed to the second model which estimates quality of NWP. Wind power from NWP is compared with the theoretical wind power obtained from power curve of wind turbine. This decides the inputs of the fuzzy model. The output of the fuzzy model provides a quality index for the NWP. In the third model, each class of wind speed is forecasted separately. The error ranges between 5 and 14 %. Catalao et al. (2011), a hybrid approach is adopted for wind power prediction combining wavelet, Particle Swarm Optimization (PSO) and adaptive Network-based Fuzzy Inference System (ANFIS). The measured values of wind power are decomposed using wavelet transform. The prediction of this data is done using ANFIS. The performance is improved using PSO technique. It is used to train the parameters of membership function of ANFIS. This is to increase the accuracy of the model. The proposed model is compared with many existing models. The least MAPE obtained is 4.98 %.

In Yan et al. (2013), a new prediction model based on Gaussian Process is proposed. The computational complexity of the standard Gaussian process is reduced and is used to train the system and to predict the wind speed. This model is compared with the standard Gaussian model. It can be seen that the error has reduced tremendously. RMSE

for each model is calculated and compared. 15 min ahead to 12 h ahead prediction has been carried out. Ghadi et al. (2012), a new neural network model named ICA-NN is proposed to predict wind speed. The inputs given are NWP data and data measured through SCADA. Initially, a neural network model is developed with temperature, humidity and wind speed. The weights are adjusted using Imperialist Competitive algorithm. The Mean Square Error is found to be less than 20 %. Wu et al. (2011), the wind power data obtained from a wind farm are preprocessed using Grubb's test. This is used as input to ANN model. Grubb's test eliminated the faulty data corresponding to outliers in the wind power or the errors in the measurement. The wind powers which are outside a predefined limit are eliminated. This is normalized and fed to RBF neural network. MAPE obtained is in the range of 7–23 % depending on the out samples of prediction.

In this paper, a novel decomposition model is proposed to separately predict seasonal and trend components of the wind speed so as to improve the forecast as compared to the existing time series models such as ARIMA, persistence, BPNN (Back Propagation Neural Network) and TES (Triple Exponential Smoothing). The paper is organized as follows. Section "Data used" explains the data used for the proposed work. Section "Time series models for prediction" gives a detailed explanation of the various time series models. Section "Error measures" gives the basic definitions of performance parameters which are used to validate the models used in this paper. Section "Model development" gives the insight of the models developed. In section "Results and discussions", the results of the developed models are shown. The predicted wind speed is compared with the actual data and the errors are calculated for the developed models.

Data used

Two models are developed based on different time series prediction concepts and the results are compared. The data used for the proposed work were collected from a wind farm in Bagalkot, Karnataka. The recorded data consist of values of Air Temperature, Relative Humidity, pressure, wind speed and the wind direction measured for different altitudes for every minute over a period of five months. For analysis, the values at one particular altitude were considered. The choice is made based on the fact that wind velocity is higher in upper strata of atmosphere thereby having more potential to generate wind energy Holland (1977). The data as such consisted of more than 150 thousand points. Each of them had been measured with an interval of approximately 1 min. It was observed that some points were not exactly spaced at that interval. Thus, the data were averaged for an interval of 10 min to reduce the number of data points as well as

to have a uniform time interval. It has been found from the study made on the data that the accuracy of the forecast depends on the type of the model and the selection of forecast interval. Thus, every model developed has been trained with different sets of data. The available data included wind speed for a duration of 5 months from February 2013 through June 2013. Every model was trained with three different sets of data.

1. 1 Month data for the month of May was used to predict the wind speed for the first day of June.
2. 2 Months data for the months of April and May was used to predict the wind speed for the first day of June.
3. 4 Months data for the months of February through May was used to predict the wind speed for the first day of June.

Time series models for prediction

Time series analysis comprises of methods for analyzing time series data to extract meaningful statistics and other characteristics of the data. Sobu and Wu (2012), Kang et al. (2011), Mori and Takahashi (2012). The main types of time series models include—parametric and non-parametric time series models. The parametric approaches assume that the underlying stationary stochastic process has a certain structure which can be described using a small number of parameters Huang and Lu (2010). All the time series models used in this paper are parametric models, described as below Douglas and Kulahci (2008).

Single, double, centered and weighted moving average models were tested for the available data with different orders and intervals. It was observed from the analysis that the value of interval, when chosen to be small (that is, less than 11), the predicted values and the actual values in the test data were closely similar. Though the predicted values are good for higher values of wind speed, when predicting lower values, the moving average model of very small interval tends to saturate towards a constant value. Hence, exponential smoothing methods were developed. The models are developed using R-software and Microsoft Excel Zhao (2012).

Single exponential smoothing (SES)

Exponential Smoothing is a class of methods that imply exponentially decreasing weights as the observations get older. They have a common property that recent values are given relatively more weight in forecasting than the older observations. The model assumes that the data fluctuate Lim (2011) around a reasonably smean (no trend or consistent pattern of growth).

SES model is formulated as indicated in Eq. 1

$$S_{t+1} = \alpha \times x_t + (1 - \alpha) \times S_t \tag{1}$$

where S_{t+1} = next forecast value of the variable, S_t is the present forecast value of the variable, X_t is the present actual value of the variable, α is a constant between 0 and 1.

Double exponential smoothening (DES)

This method is used when the data show a trend. Exponential smoothing with a trend works much like simple smoothing except that two components must be updated each period—level and trend. The level is a smoothed estimate of the value of the data at the end of each period. The trend is a smoothed estimate of average growth at the end of each period. DES can be modeled as given in Eqs. 2 and 3

$$\text{Level: } S_{t+1} = \alpha X_t + (1 - \alpha)(S_{t-1} + b_{t-1}) \tag{2}$$

$$\text{Trend: } b_t = \beta(S_t - S_{t-1}) + (1 - \beta)b_{t-1} \tag{3}$$

Triple exponential smoothening (TES)

This method is used when the data show trend and seasonality. To handle seasonality, a third parameter is added. Now, a third equation is introduced to take care of seasonality. The resulting set of equations is called the “Holt-Winters” (HW) method after the names of the inventors. The model is given by Eqs. 4, 5 and 6

$$\text{Level: } L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}) \tag{4}$$

$$\text{Trend: } b_t = \beta(L_t - L_{t-1}) + (1 - \beta)(b_{t-1}) \tag{5}$$

$$\text{Seasonal: } S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s} \tag{6}$$

α is the level smoothing constant (lies between 0 and 1).
 β is the trend smoothing constant (lies between 0 and 1).

γ is the seasonal smoothing constant (lies between 0 and 1).

L_t is the estimate of the level of the series at time t .

Y_t is the actual value of the series at time t .

b_t is the estimate of the slope of the series at time t .

S_t is the seasonal component.

s is the length of seasonality.

m is the number of periods ahead to be forecast.

F_{t+m} is the forecast for m periods ahead.

The forecast is obtained from the level, trend and seasonal components as in Eq. 7

$$\text{Forecast: } F_{t+m} = L_t + b_{tm} + S_{t-s} + m \tag{7}$$

Decomposition model

It is possible to decompose time series data into various sub-components to see how these components affect the data in the series. Time series data are mainly composed of seasonal pattern and trend pattern. There are two different decomposition models possible.

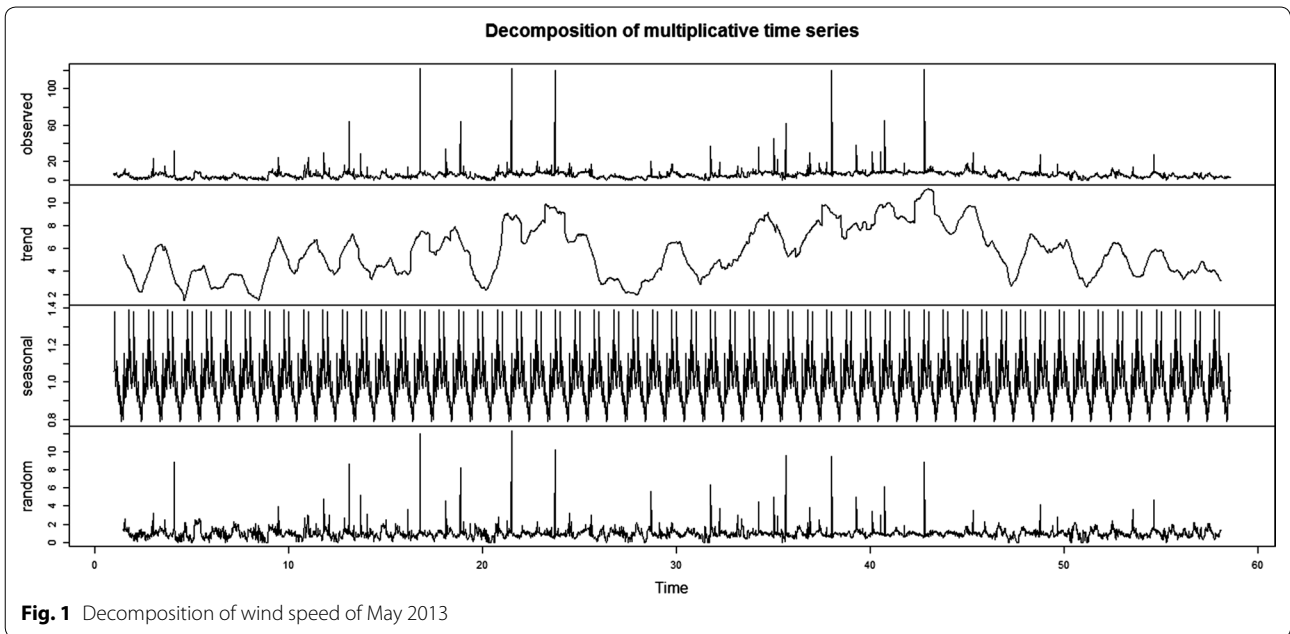


Fig. 1 Decomposition of wind speed of May 2013

1. Additive Decomposition: Here, the total data are taken as the sum of the decomposed patterns

$$X_t = \text{seasonal } (S_t) + \text{trend } (T_t) + \text{random.}$$

2. Multiplicative Decomposition: Here, the given time series data are treated as the product of the decomposed patterns

$$X_t = \text{seasonal } (S_t) \times \text{trend } (T_t) \times \text{random.}$$

Additive decomposition is effective when the peak values of the seasonal data do not vary much. Multiplicative models are effective when the seasonal value changes over time. Both additive and multiplicative models were tested for the available data. It was found that the multiplicative model works well. The reason could be that the wind speed changes with the change in weather. Fig. 1 shows the decomposition of wind speed for the month of May 2013 using R.

Forecasting the decomposed seasonal and trend patterns separately to predict the wind speed worked very well. This analysis was done in Microsoft Excel. The following procedure was used for the forecast. The data were smoothed by taking the centered moving average. The interval for centered moving average was taken as 144 samples which correspond to samples at every 10 min for each day. The de-trended series is calculated by dividing the smoothed series from the actual data. The seasonal component for each 10-min sample

is estimated by averaging the de-trended values for that particular time of each day. The total seasonal pattern is obtained by stringing together all the seasonal indices for each day of the data. The forecast for the seasonal component can easily be estimated by extending the seasonal pattern to the next day. The forecast for the de-seasonalised data is obtained by developing exponential smoothing model for the de-seasonalised data. This hybrid model was tested for various data sample durations, like 1 month, 2 months and 4 months. The steps involved in developing the multiplicative decomposition model are listed below.

1. Obtain the trend pattern by calculating the centered moving average for the entire data.
2. Find out the de-trended pattern by dividing the total data by the trend series.
3. Seasonal component is calculated by averaging the de-trended values for that particular time of all the days.
4. Calculate the de-seasonalised data by dividing the seasonal data from the actual data.
5. Bring out a single exponential model for the de-seasonalised data.
6. Forecast seasonal data by extending the data for the next day.
7. Forecast de-seasonalised data by forecasting the exponential smoothing model.
8. Multiply both the forecasted data to get the forecast for the actual data.

Error measures

When there are more than one model developed, a comparative analysis can be made by calculating the accuracy of each model and comparing them. The accuracy of any predictive model can be determined only by choosing appropriate error measures. In this paper, different models are compared using two error measuring parameters, Mean Absolute Percentage Error (MAPE) and Symmetric Mean Absolute Percentage Error (SMAPE) which can be defined as shown below. MAPE is a relative measure which reflects error as the percentage of actual data. Hence, the accuracy of the model can be easily judged.

Mean absolute percentage error (MAPE)

MAPE can be defined as,

$$MAPE = \frac{1}{N} \times \sum \frac{|X_t - F_t|}{X_t} \times 100$$

Symmetric mean average percentage error (SMAPE)

Symmetric MAPE can be defined as,

$$SMAPE = \frac{1}{N} \times \sum \frac{|X_t - F_t|}{(X_t + F_t)/2} \times 100$$

Model development

Decomposition model and TES models were developed. The predicted values of wind speed obtained for the first day of June were compared with the actual values of wind speed and errors were calculated for each model. Amongst the developed models, decomposition model is found superior and further analysis was carried out with decomposition model. The model was tested on different durations of past data for a period of 1 year. The prediction was done for different climatic conditions such as monsoon, spring and winter.

These two models are compared with classic methods of wind forecasting such as Persistence, ARIMA and BPNN. All these models were tested with 3 different sets of data, 1-month duration, 2-month duration (April and May 2013) and 4-month (February, March, April and May) duration to predict the wind speed of 1st of June 2013. Persistence and Decomposition models were developed in Microsoft Excel. TES model was developed using R software. Table 1 gives the best smoothing parameters for the developed models. ARIMA and BPNN models were developed in MATLAB. Different order of ARIMA models was tested. The best model obtained has the order ($p = 4, q = 1, d = 2$). BPNN model was developed with Non-Linear Auto-regressive (NAR) network with Gradient Descent as the training algorithm. A single hidden layer Neural network with

Table 1 Smoothing parameters of the TES models

Sl no	Model	α	β	γ
1	TES 1 month	0.59	0.09	0.17
2	TES 2 months	0.71	0.11	0.21
3	TES 4 months	0.68	0.099	0.2

Table 2 Comparison of errors

Data duration	Model	MAPE (%)	RMSE (m/s)
1 Month	Persistence	31.63	1.43
	ARIMA	25.21	1.07
	BPNN	37.6	1.98
	TES	23.94	1.47
	Decomposition	20.15	0.965
2 Months	Persistence	31.63	1.43
	ARIMA	22.15	0.979
	BPNN	40.94	1.89
	TES	26.03	1.772
	Decomposition	22.03	1.105
4 Months	Persistence	31.63	1.43
	ARIMA	23.26	1.022
	BPNN	44.37	1.91
	TES	28.63	1.35
	Decomposition	18.24	1.05

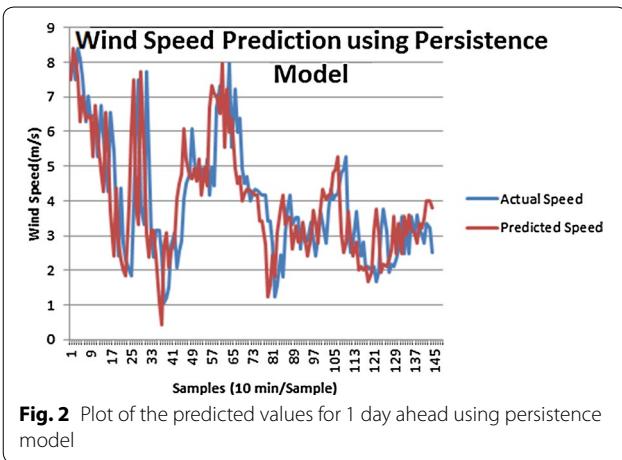
7 neurons was found to give optimal result. 80 % of the input data were used for training and 20 % for validation. Table 2 gives the comparison of MAPE and RMSE of the developed models.

Figures 2, 3, 4, 5 and 6 give the comparison of actual wind speed with predicted values for Persistence model, ARIMA model, BPNN model, TES model and Decomposition model, respectively.

Results and discussions

Table 2 shows the comparison of the errors calculated for the prediction. As per Government of India, a deviation of 30percent in the forecast is acceptable CERC (2012).

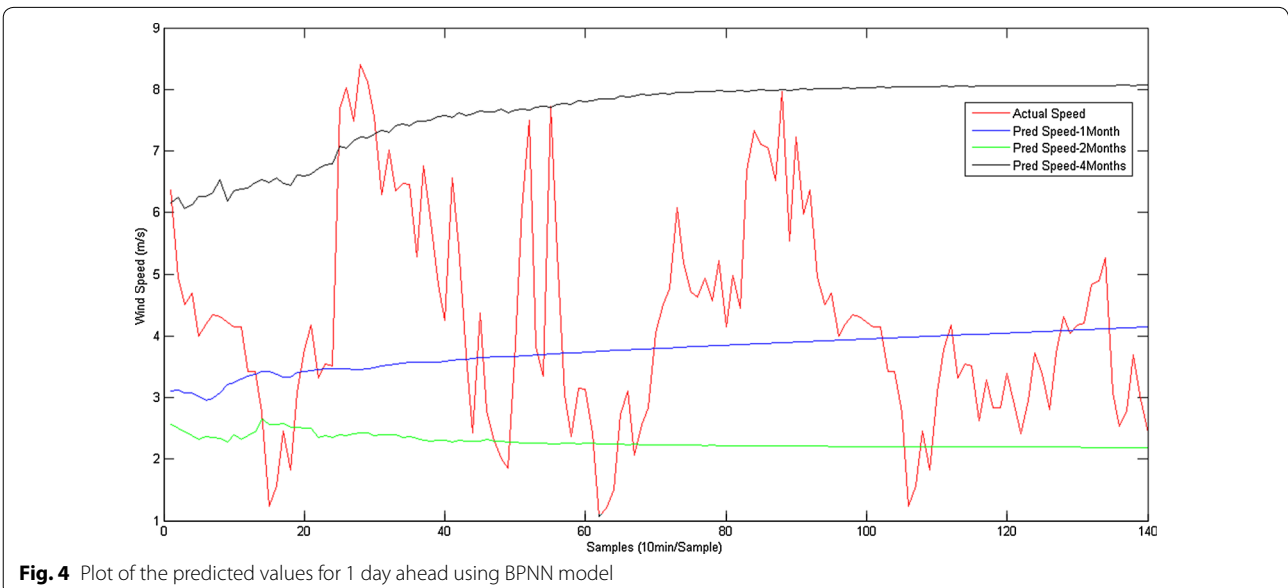
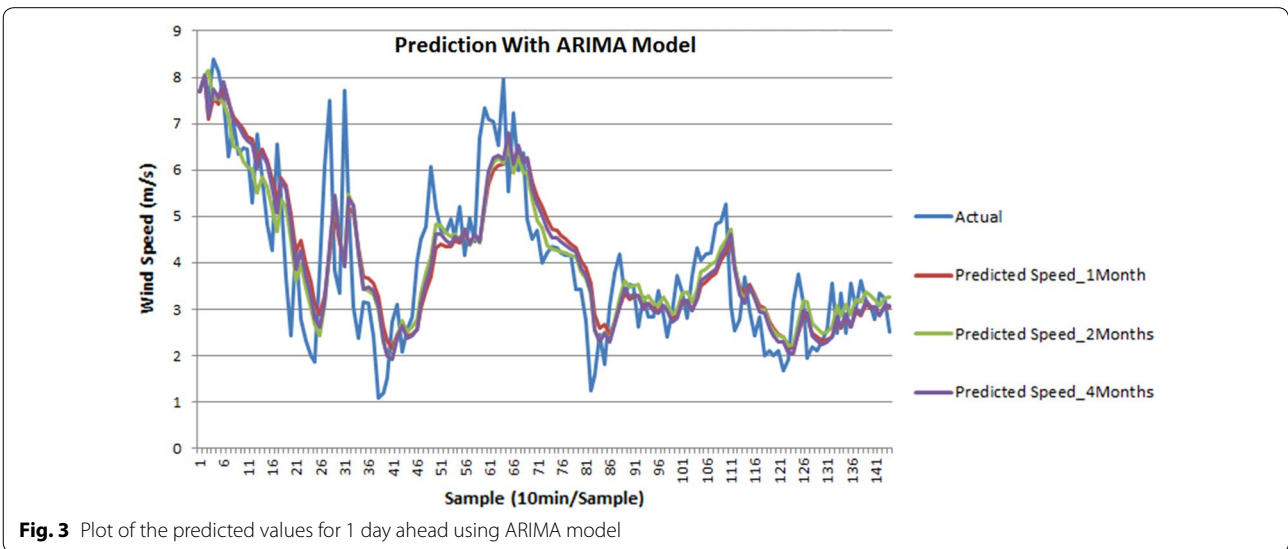
It can be observed from Table 2 that the proposed decomposition models outperform all other benchmark models such as Persistence, TES, ARIMA and BPNN. The benchmark models have a MAPE in the range of 40–23 % and RMSE of around 1 m/s. But decomposition models give an error as less as 18 %. Thus, decomposition models are considered for further analysis. The decomposition model proposed above is tested with various time durations of data and wind speed is predicted for different climatic conditions. For these studies, a data set



of wind speed for duration of 1 year from January 2013 to December 2013 is considered. The results are presented as various case studies explained below.

Case study 1: to determine ideal time period for model building

The main issue in developing a forecasting model is to determine the duration of past data samples to be considered for the model development. In this section, the wind speed for December 31 is predicted using different durations of data (1 year, 6 , 4 , 2 and 1 month) for model building. Table 3 shows the errors for the different cases. Figure 7 plots the actual and predicted wind speeds of December 31 for different data durations for



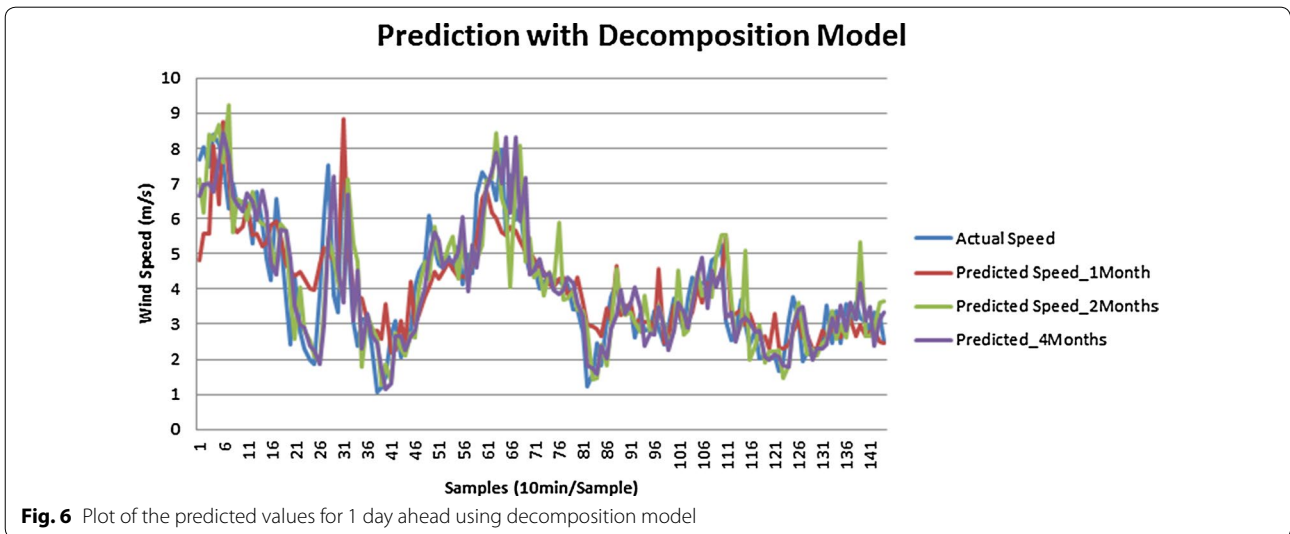
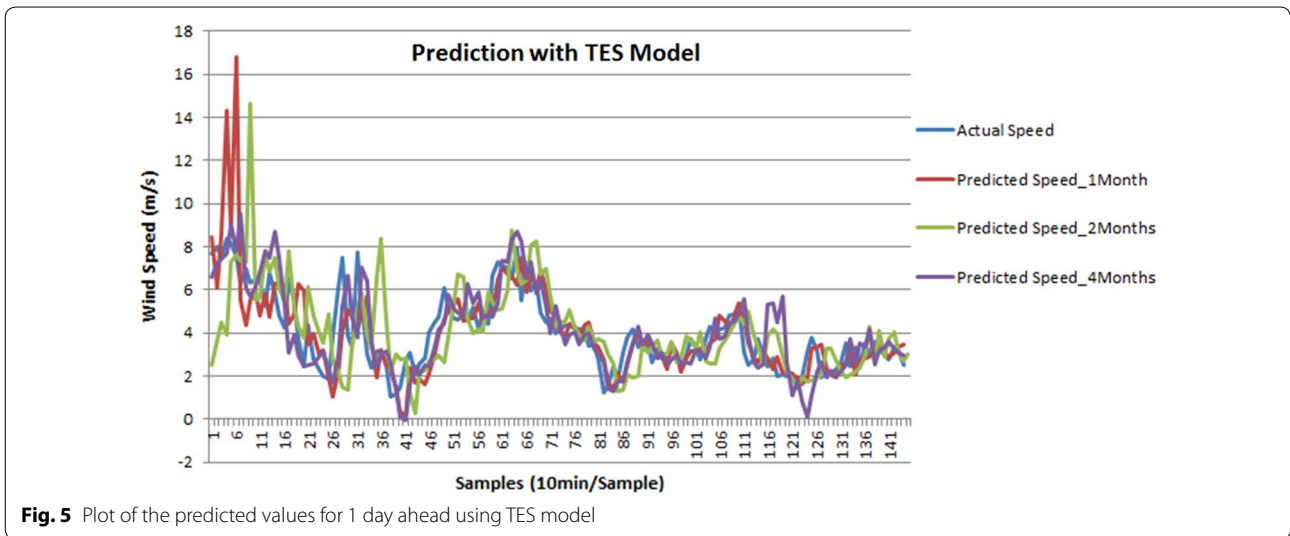


Table 3 Forecast of 31 Dec 2013 using different time periods for model building

Sl no	Time period	MAPE	SMAPE	Mean MAPE	RMSE
1	1 year	34.89	14.61	26.84	0.90
2	6 months	28.30	26.21	24.53	0.96
3	4 months	22.55	21.49	18.72	0.68
4	2 months	25.89	26.02	23.64	0.87
5	1 month	31.12	26.09	25.63	0.97

model building. It can be seen from the plot that the forecast using 4-month data exactly follows the actual wind speed and the error is least for the model using 4-month

data. As the duration of data is increased, the error is also increased. The model with 1-year data gives a maximum MAPE of 34 %. Thus, it can be concluded that the decomposition model with a historical data of 4 months is ideal for model building. The reason for this could be the drastic changes in the wind patterns every four months at the site location. It is possible that in a different country, a different duration of historical data set may yield better results.

Case study 2: prediction during different climatic conditions

To understand the performance of model, it is tested for different climatic conditions. Wind speeds for 5

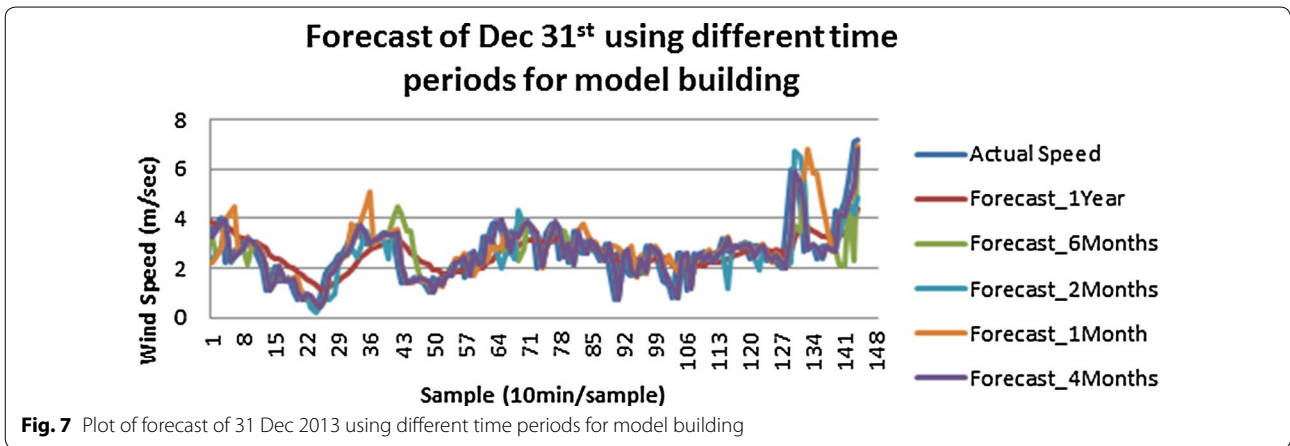


Fig. 7 Plot of forecast of 31 Dec 2013 using different time periods for model building

days each in monsoon, spring and winter seasons are predicted.

Monsoon

The location from which the data is collected is a tropical region where the monsoon hits during the second half of May every year. Because of the change in the weather from summer to monsoon, the wind is erratic

during this condition. Thus, prediction becomes difficult. Wind speed for 5 consecutive days in the end of May 2013 is predicted. The predicted speed is compared with actual speed and the calculated errors are tabulated in Table 4. It can be observed from the table that MAPE varies from 26.92 to 16.82 % between two consecutive days, 30 May and 31 May. This is due to the erratic nature of wind during the onset of monsoon. Figure 8 shows the comparison of actual speed and predicted speed for 30th May. It can be observed for the plots that wind speed varies for 2–16 m/s, which makes the prediction less accurate.

Table 4 Prediction of wind speed during monsoon

Sl no	Time period	MAPE	SMAPE	Mean MAPE	RMSE
1	01 June 2013	21.16	20.22	19.29	1.05
2	31 May 2013	26.92	16.10	14.05	0.92
3	30 May 2013	16.82	15.81	15.23	1.195
4	29 May 2013	19.13	19.37	16.36	0.96
5	28 May 2013	20.19	18.35	16.53	0.98

Spring

The spring season in India owns a pleasant weather and the prediction is easier during this time of a year. This can be seen from Table 5 where the prediction for 5 consecutive days in the month of September is shown. It can be observed from the table that spring season gives

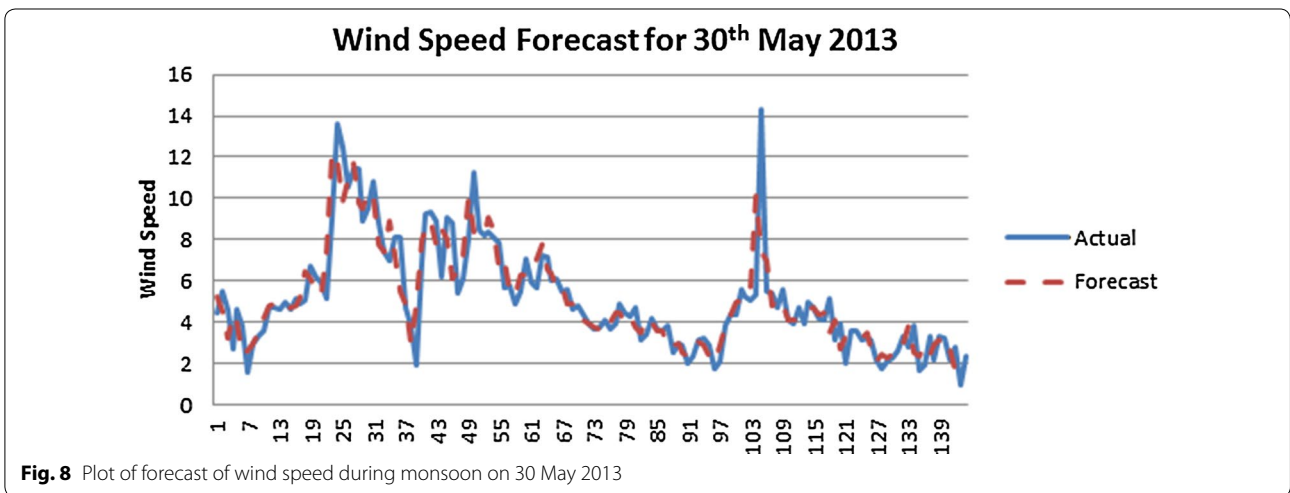


Fig. 8 Plot of forecast of wind speed during monsoon on 30 May 2013

lower MAPE compared to other weather conditions. The lowest MAPE observed is 6.72 % during 4 September. Figure 9 shows the comparison of actual speed and predicted speed for 5 September.

Winter

Winter season is during November to January in India. Wind speed is predicted for 5 days in December to study the performance of model. Wind is calm during winter and the lowest MAPE observed is 9.5 % as shown in Table 6.

Case study 3: extension of models to predict out sample

The objective of this case study is to see the extent of out-of sample prediction which can be performed with a developed model. Wind speed is predicted for 1 day ahead to 5 days ahead and errors are tabulated in Table 7. It can be observed that MAPE increases with increase in forecast horizon. The plots for 3 days ahead prediction are shown in Fig. 10. It can be concluded from this study that wind speed can be best predicted for 1 day ahead with the developed models.

Table 5 Prediction of wind speed during spring

Sl no	Time period	MAPE	SMAPE	Mean MAPE	RMSE
1	5 Sept 2013	8.3	8.26	8.14	0.66
2	4 Sept 2013	6.72	6.67	6.92	0.66
3	3 Sept 2013	8.14	8.16	8.51	0.79
4	2 Sept 2013	9.2	9.04	9.49	0.81
5	1 Sept 2013	7.44	7.41	7.31	0.69

Comparison with existing models

The classic models for prediction such as persistence, BPNN, ARIMA and TES are developed with the same set of data. MAPE and RMSE of these models are compared with the proposed model in Table 2. MAPE is above 30 % for persistence and BPNN and above 20 % for ARIMA and TES. RMSE is also on the higher side. This proves that these models cannot predict erratic data like wind speed efficiently. The lowest MAPE obtained for the proposed model is 6.72 %, which is very less compared to the classic models.

The total run time of the simulations is also a matter of concern in prediction. As model building is done offline, the time taken for model building does not significantly affect the system performance. Once the model is built, running the model to gather a day-ahead prediction takes only less than 2 s in the proposed model, whereas BPNN model took around 11 s to get the predicted wind speed.

A comparison of the proposed model with the models available in the literature becomes very difficult as the performance of a model is heavily dependent on the data and it is influenced by various parameters such as geographical locations and forecast out samples. The error measures considered in various models are also different. Still an effort is done to compare the proposed model with some of the existing models which use time series predictive models. Li et al. (2010), a 2-step forecasting technique based on Bayesian algorithm and neural network is developed. The results were obtained for two different sites and the MAPE is in the range of 14–18 %. Lv and Yue (2011) propose a forecast model based on wavelet,

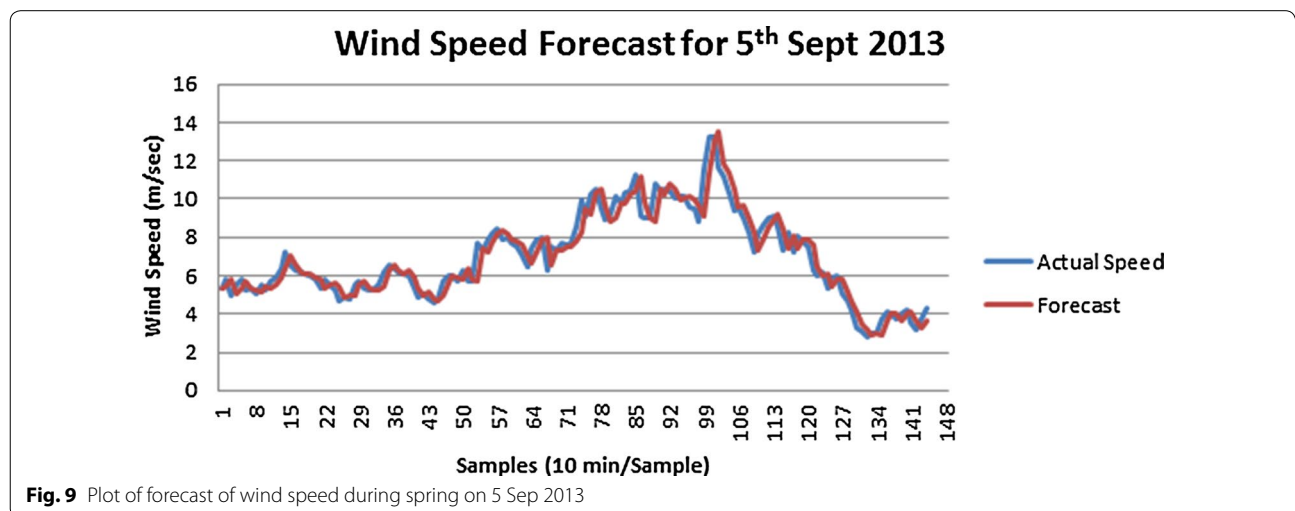


Fig. 9 Plot of forecast of wind speed during spring on 5 Sep 2013

Table 6 Prediction of wind speed during winter

Sl no	Time period	MAPE	SMAPE	Mean MAPE	RMSE
1	31 Dec 2013	22.54	21.46	18.70	0.68
2	30 Dec 2013	11.64	11.49	10.63	0.68
3	29 Dec 2013	9.50	9.31	8.68	0.51
4	28 Dec 2013	7.83	7.74	7.16	0.52
5	27 Dec 2013	10.85	10.77	9.48	0.52

Table 7 Comparison of errors for different out samples of prediction

Sl no	Forecast horizon	MAPE	SMAPE	Mean MAPE	RMSE
1	1 day (Dec 27)	10.85	10.77	9.48	0.52
2	2 days (Dec 27, 28)	15.42	16.78	14.36	0.88
3	3 days (Dec 27–29)	16.02	16.84	15.39	0.92
4	4 days (Dec 27–30)	20.51	21.87	19.63	1.18
5	5 days (Dec 27–31)	16.40	17.49	15.70	1.05

ARIMA and ARCH. The input is wavelet decomposed to remove outliers and the MAPE obtained is in the range of 17.85–8.72 %. Wen et al. (2012), Support Vector Machine (SVM) model and Lifting Wavelet Transform SVM (LWT-SVM) are compared and least MAPE obtained is 14.96 %. The decomposition model proposed in this paper gives a minimum MAPE of 6.7 %, which is less than the models discussed above. Johnson et al. (2007), wind forecast model is developed using adaptive neuro-fuzzy inference systems. The proposed model is compared with persistence

model and percentage improvement obtained is found to be 8 %. In this paper, the percentage improvement of the proposed model compared to persistence model is 23 %.

Conclusion

In this paper, a new method for forecast of wind speed has been proposed using decomposition of raw time series data. The forecast result has been compared with classical methods such as Persistence, ARIMA, BPNN and TES models. It is observed from the results and calculated errors that decomposition model gives better results compared to conventional forecasting methods. Models were tested with different durations of historical data and it was found that decomposition models worked well with data duration of 4 months. The model was tested for different weather conditions throughout a year. The prediction error was highest during onset of monsoon when wind speed is erratic and as less as 6.7 % during spring when wind speed is reasonably consistent. The models developed have better applications at State Load Regional Dispatch Center (SLRDC) for scheduling generation a day ahead thereby increasing penetration of wind power in the grid. If the data are viewed as a signal, the sudden variations can be treated as high-frequency components. The performance of the model can be further improved if these high-frequency components are captured. Work is being carried out using wavelets to capture the high-frequency components, which can be incorporated in the decomposition model to improve the prediction.

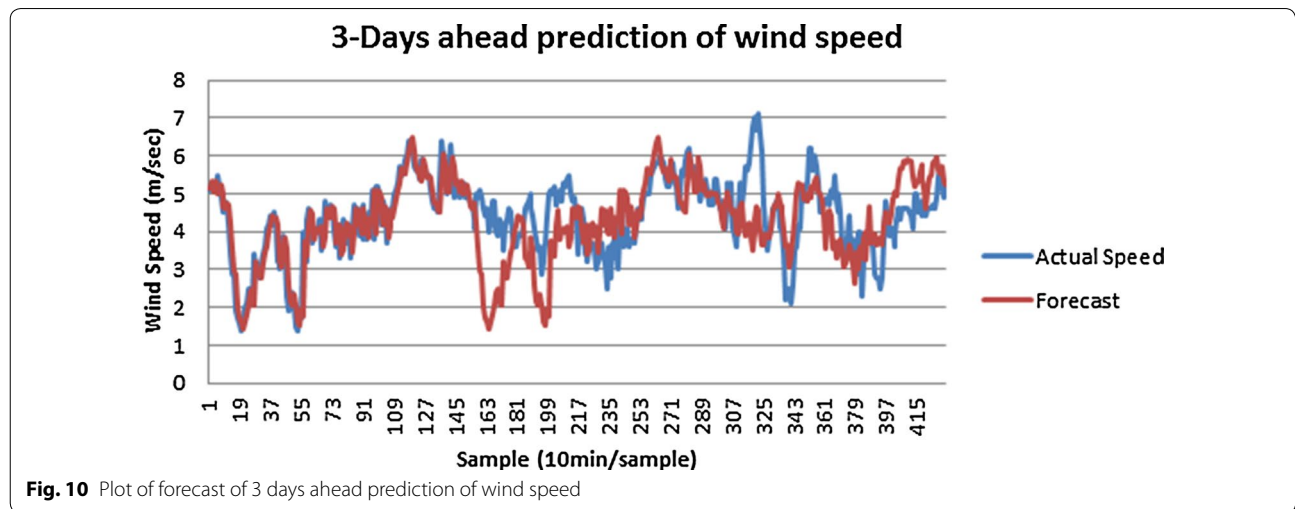


Fig. 10 Plot of forecast of 3 days ahead prediction of wind speed

Authors' contributions

Both the authors have equal contribution on the entire work done for this paper. Both authors read and approved the final manuscript.

Competing interests

The authors declare that there are no financial or non-financial competing interests with any other person or organization.

Received: 15 June 2015 Accepted: 23 October 2015

Published online: 12 November 2015

References

- Catalao, J. P. S., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Hybrid wavelet-soft computing approach for short-term wind power forecasting in Portugal. *IEEE Transactions on Sustainable Energy*, 2
- Central Electricity Regulatory Commission (CERC). (2012). New Delhi, Statement of Objects and Reasons (Terms and Conditions for Tariff Determination of Renewable Energy Sources). URL: <http://www.cercind.gov.in/2012/regulation/RE-Tariff-Regulations-2012-SOR%206-2-2012.pdf>
- Erdem, E., Peng, Y., & Shi, J. (2014). Short-term forecasting of wind speed and power—a clustering approach. In *Proceedings of the 2014 Industrial and Systems Engineering Research Conference*
- Ghadi, M. J., Sharifiyan, A., Gilani, S. H., & Afrakhteh, H. (2012). A new method for short-term wind power forecasting. *International Journal of Electrical Power and Energy Systems*, 34, 1–176
- Holland, R. E. W. P. W. (1977). *Communications in statistics: theory and methods*, 6th edn. New York: Wiley Student Edition
- Huang, Y., & Lu, J. (2010). Comparative study of power forecasting methods for PV stations. In *International Conference on Power System Technology (POWERCON)* (pp. 1–6)
- Johnson, P., Muttaqi, K. M., & Negnevitsky, M. (2007). Short term wind power forecasting using adaptive neuro-fuzzy inference systems. In *Power Engineering Conference (AUPEC)* (pp. 1–6)
- Kang, M. C., Sohn, J. M., Park, J. Y., Lee, S. K., & Yoon, Y. T. (2011). Development of algorithm for day ahead PV generation forecasting using data mining method. In *IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)* (pp. 1–4)
- Li, G., Zhou, J., & Shi, J. (2010). Bayesian adaptive combination of short-term wind speed forecasts from neural network models, 36th edn. *Renewable Energy* Elsevier
- Lim, Y. P. (2011). *Power management strategies for off-grid hybrid power systems*. PhD thesis, Curtin University, School of Electrical and Computer Engineering
- Liu, H., Erdem, E., & Shi, J. (2011). An integrated wind power forecasting methodology: interval estimation of wind speed, operation probability of wind turbine, and conditional expected wind power output of a wind farm, 12th edn. Elsevier
- Lv, P., & Yue, L. (2011). Short-term wind speed forecasting based on non-stationary time series analysis and arch model, 12th edn. IEEE
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2008). *Introduction to time series analysis and forecasting*. New York: Wiley publications
- Mori, H., & Takahashi, A. (2012). A data mining method for selecting input variables for forecasting model of global solar radiation. In *IEEE PES Transmission and Distribution Conference and Exposition*
- Sideratos, G., & Hatzigiorgiou, N. D. (2007). An advanced statistical method for wind power forecasting. *IEEE Transactions On Power Systems*, 22, 258–265.
- Sobu, A., & Wu, G. (2012). Dynamic optimal schedule management method for microgrid system considering forecast errors of renewable power generations. In *IEEE International Conference on Power System Technology (POWERCON)* (pp. 1–6)
- Wang, X., Huang, X., & Guo, P. (2011). *A review of wind power forecasting models*, 12th edn. Elsevier
- Wang, Q., Lai, K. K., Niu, D., & Zhang, Q. (2012). A multivariate wind power forecasting model based on LS-SVM. In *Fifth International Joint Conference on Computational Sciences and Optimization*
- Wen, J., Wang, X., Zheng, Y., Li, L., Zhou, L., Yao, G., & Chen, H. (2012). Short-term wind power forecasting based on lifting wavelet transform and SVM. In *IEEE Power Engineering and Automation Conference (PEAM)* (pp. 1–4)
- Wu, X., Wen, F., Hong, B., Peng, X., & Huang, J. (2011). Radial basis function neural network based short-term wind power forecasting with Grubbs test. In *4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*
- Yan, J., Li, K., & Bai, E. W. (2013). Prediction error adjusted Gaussian process for short-term wind power forecasting. In *IEEE International Workshop on Intelligent Energy Systems (IWIES)* (pp. 173–178)
- Zhao, X., Li, T., & Wang, S. (2011). Review of evaluation criteria and main methods of wind power forecasting, 12th edn. Elsevier
- Zhao, Y. (2012). *R and Data Mining: Examples and Case Studies*. New York: Elsevier

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com