



Performance monitoring of wind turbines using advanced statistical methods

ANIL KUMAR KUSHWAH* and RAJESH WADHVANI

Department of Computer Science and Engineering, Maulana Azad National Institute of Technology,
Bhopal 462003, India
e-mail: akushwah190@gmail.com; rajeshwadhvani@manit.ac.in

MS received 21 March 2018; revised 15 April 2019; accepted 22 April 2019; published online 17 June 2019

Abstract. Estimation of wind power generation for grid interface helps in calculation of the annual energy production, which maintains the balance between electricity production and its consumption. For this purpose, accurate wind speed forecasting plays an important role. In this paper, linear statistical predictive models such as autoregressive integrated moving average (ARIMA), generalized autoregressive score (GAS) model and a GAS model with exogenous variable x (GASX) have been applied for accurate wind speed forecasting. Along with this, a non-linear statistical predictive modelling technique called non-linear GASX (NLGASX) has been proposed and applied to model non-linear time-series data. Furthermore, the proposed NLGASX model is optimized using modelling techniques based on neural networks, namely Sigmoid, TANH, Softmax and RELU. The proposed optimized NLGASX model performs far better as compared with other models. Wind speed is also used as an input to wind power curve model for predicting the wind power. According to the predicted wind power the annual energy has been calculated.

Keywords. Statistical models; wind power curve models; TANH; softmax; RELU; sigmoid; annual energy production.

1. Introduction

Wind power utilization plays an important role in electricity generation industries. The balance between electricity production and its consumption is maintained by transmission system operators (TSO). Wind energy is a non-stationary as well as fluctuating energy source. A specific measure for saving power is expected to compensate for these fluctuations [1]. The utilization of wind power expectations minimizes the cost of realizing wind power into the power supply framework. There are two methodologies proposed for wind power forecast. First, design a physical model of the wind farm to decide the connection between climate information and wind power. Second is a mathematical model in which statistical and artificial intelligence is used to define the relationship between weather prediction data and power output from the historical datasets [2]. In recent years, to improve the accuracy of prediction, various advanced methods have been suggested, which categorize the prediction models into three classes: physical models [3], statistical models [4–8] and models based on neural networks [9–11]. In physical models (also referred to as analytical models) approach, the objective is to develop a mathematical function covering all the

predictor factors. Since the predictor values are complex and computation intensive, accurate values of all predictors are difficult to derive. Difficulties in the analytical model motivate the use of statistical modelling techniques. Statistical models such as autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), autoregressive model with exogenous variable (ARX) and ARIMA model with exogenous variable (ARIMAX) are well-defined models for forecasting but suitable in the case when available data for modelling are linear in nature [4–8]. However, artificial intelligence models such as support vector machine (SVM) [9], support vector regression (SVR) [10] and artificial neural network (ANN) [11] are effectively used for non-linear data. Models of all three classes perform well when dealing with homogeneous data. However, if the data are heterogeneous, then these models do not perform well due to poor handling of heteroscedasticity. For improving wind speed and wind power prediction accuracy further, hybrid models (usually referred to as heterogeneous estimators) have been developed that incorporate the advantages of individual forecasting methods [12, 13]. Heterogeneous estimators present a different nature of the variable in the same dataset. These estimators help discover an impressive inconsistency in the parameter estimates and produce unattainable evaluations

*For correspondence

while the homogeneous estimators show the same nature of the variable in the same dataset. They produce significantly comparable short-run estimates that differ only in the long-run effect. To handle the heteroscedasticity, special statistical models such as generalized autoregressive score (GAS) model, generalized autoregressive score model with exogenous variable (GASX) and non-linear GASX model (NLGASX) are used and discussed in this paper. The GAS model is applied to use the complete density structure along with mean and variance. It is used to represent all forms of time-series data, which can be either real-valued, integer-valued or bounded. The values should ensure conditional density when the score function is well defined [14]. The main challenge of using the GAS model on non-linear data is to evaluate the score and the maximum likelihood estimation. The current limitation of the GAS model is that the multivariate GAS model for more than four variables does not report the exact update for the correlation parameters. As such, to overcome this limitation, a hybrid model is being proposed in this work. The hybrid model is developed by the combination of the statistical and ANN models for improving the performance of wind speed forecasting and also overcoming the limitations of a single forecasting mechanism.

In this paper, statistical models and some modelling techniques based on neural networks are presented that play a vital role in modelling and forecasting of wind energy. Using these techniques, this paper presents a hybrid model to forecast wind speed by combining NLGASX model and modelling techniques based on neural networks activation function (TANH, Sigmoid, Softmax and rectified linear unit (RELU)) to achieve better accuracy. The accuracy of the proposed model is verified by simulation of the measured wind speed. For learning the models, Maximum Likelihood Estimation (MLE) is used as shown in Eq. (1):

$$\hat{\theta} = \operatorname{argmax} l(\theta; x) \quad (1)$$

where θ is the vector of model coefficients, which has to be estimated by MLE, and $\hat{\theta}$ is the estimated value of θ .

The MLE is defined as a method of estimating the value of θ to maximize $l(\theta; x)$. MLE is used for the learning of the statistical models.

Another parameter to evaluate the performance of proposed approach is to determine annual energy production (AEP) as shown in Eq. (2):

$$AEP = \sum_{i=1}^N \bar{p}(v_i) \quad (2)$$

where N is the total number of hours in a year, v is the wind speed and \bar{p} is the average hourly power output, which is calculated from wind turbine power curve model. The contributions and novelties of this study are illustrated here.

(a) The developed model NLGASX, using modelling techniques, enhances the prediction accuracy of wind

speed prediction. First, the wind speed is predicted using NLGASX model; later this predicted wind speed is given as an input to modelling techniques, which will improve the prediction accuracy.

- (b) Wind power is predicted also using the wind turbine power curve models, which can further be applied to wind speed that is predicted by the hybrid model.
- (c) Using wind power generated by the wind turbine power curve model, the production of annual energy is calculated.

The rest of the paper is organized as follows. In section 2, various time-series models for wind speed prediction are presented. Section 3 presents the wind turbine power curve models used for the prediction of wind power. To optimize the predictions, several modelling techniques for wind speed forecasting are shown in section 4. Section 5 presents the proposed non-linear GASX model along with the technique for developing hybrid models using optimizations based on neural networks. Several developed hybrid models are evaluated, and corresponding results are presented in section 6. Section 7 concludes the work.

2. Time-series model for wind speed forecasting

The time-series model has been defined in terms of forecasting the data, horizon and accuracy. The forecasting is done using factors such as wind speed, temperature, density, etc. To make accurate estimations about energy production at a geographical location, wind speed at periodic intervals is recorded. The collected wind speed data at periodic or timed intervals can be considered as a time-series data, which can be used for time-series forecasting. Several statistical models are applied for predicting the wind speed such as ARIMA model, GAS model and GAS model with exogenous variable x (GASX), which are explained here.

2.1 ARIMA model

Maatallah *et al* [15] have proposed the ARIMA model, which can represent different types of time series such as purely AR, purely MA and combination of both AR and MA (ARMA) series. The real constraint is the pre-assumed linear form of the model, i.e., a direct link structure is expected among the time-series values, and hence no non-linear pattern can be caught by the ARIMA model. The notation ARIMA (p, d, q) refers to the model with p AR terms, d degree of a differencing term and q MA terms. This model contains the AR(p) and MA(q) models defined as shown in Eq. (3):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \cdots + \alpha_q \epsilon_{t-q} \quad (3)$$

where $\beta_0, \beta_1, \dots, \beta_p$ are the AR coefficients, $\alpha_1, \alpha_2, \dots, \alpha_q$ are the MA coefficients and ϵ_t is assume to be a white noise. The ARIMA model performs better for linear time-series data and can be applied only on stationary time series. If a series is non-stationary, then the series is first converted into stationary and then the ARIMA model is applied.

2.2 GAS model

Creal *et al* [16] and Harvey [17] have proposed the GAS model, which can handle the varying density time series. For a given time-series observation y_t , the GAS model is represented by a conditional observation density $p(y_t/\theta_t)$, where θ_t is a latent time-varying parameter and follows the recursion as shown in Eq. (4):

$$\theta_t = \mu + \sum_{(i=1)}^p \phi_i \theta_{t-i} + \sum_{(j=1)}^q \alpha_j s(\theta_{t-j}) \frac{\partial \log p(y_{t-j}) / (\theta_{t-j})}{\partial \theta_{t-j}} \quad (4)$$

where μ is an intercept in AR process of order 1; ϕ is an AR coefficient; α is a scaling parameter and s is a strictly positive scaling factor multiplied with first derivative of the conditional density contribution for a single observation at time j . The GAS model considers all types of time-series data. It does not make a difference whether the data is real-valued, integer-valued or (0,1)-bounded as long as there is a conditional density for which the score function and the Hessian are well characterized. The main difficulty in using GAS models is inferring the score and executing the MLE on the non-linear data.

2.3 GASX model

The GASX model is the enhanced version of the GAS model improved by including exogenous factors X . Here, conditional observation density $p(y_t/\theta_t)$ is calculated using the observation y_t and a latent time-varying parameter θ_t , where parameter θ_t follows the recursion as shown in Eq. (5):

$$\theta_t = \mu + \sum_{(k=1)}^K \beta_k X_{t,k} + \sum_{(i=1)}^p \phi_i \theta_{t-i} + \sum_{(j=1)}^q \alpha_j s(x_{j-1}) \frac{\partial \log p(y_{t-j}) / (\theta_{t-j})}{\partial \theta_{t-j}} \quad (5)$$

where μ is an intercept in the AR process of order 1; X is an exogenous variable and β is a coefficient used in exogenous factor; ϕ is an AR coefficient; α is a scaling parameter and s is a strictly positive scaling function multiplied with first derivative of the conditional density distribution for a single observation at time j . The advantage of GASX model is to use the extra factor in the modelling to improving the

accuracy of the model. GASX model is restricted by the use of more than four factors, which does not report the exact update for the correlation parameters.

3. Models for wind turbine power curve

Wind turbine power curve is a tool that is provided by a turbine manufacturer to test its performance in optimal conditions. Usually, the power curve of the commercially available wind turbine is released either in the form of a power rating of the turbine or nominal wind power reading. Here, the turbine power curve includes cut-in speed, rated speed and cut-out speed. Similarly, nominal wind power reading is recorded under standard test conditions [18]. In order to calculate the AEP of a given wind turbine, an empirical power curve is needed, which can predict the output power of the wind turbine for all possible wind speeds between cut-in and cut-out.

Pallabazzer [19] and Jangamshetti and Guruprasada Rau [20] have proposed empirical power curve models, which are based on the quadratic and cubic law, respectively. The quadratic power curve (QPC) model is given by Eq. (6):

$$P(v) = \frac{p_r(v^2 - v_{cin}^2)}{v_{rate}^2 - v_{cin}^2} \quad (6)$$

and the cubic power curve (CPC) model is given by Eq. (7):

$$P(v) = \frac{p_r(v^3 - v_{cin}^3)}{v_{rate}^3 - v_{cin}^3} \quad (7)$$

where $P(v)$ is the non-linear region between cut-in and rated speed; v is wind speed; p_r is rated power; v_{cout} is the cut-out wind speed; v_{cin} is the cut-in wind speed and v_{rate} represents the rated speed. Though these models are simple, they do not replicate the actual behaviour of wind turbines.

In order to increase the prediction accuracy, Kazemi and Goudarzi [21] have proposed the generalized QPC and CPC models, which can be developed using nominal wind power reading given by the manufacturer. The generalized QPC model is given by Eq. (8):

$$P(v) = a_1 v_i^2 + a_2 v_i + a_3. \quad (8)$$

Similarly, generalized CPC model is given by Eq. (9):

$$P(v) = b_1 v_i^3 + b_2 v_i^2 + b_3 v_i + b_4 \quad (9)$$

where $a_1, a_2, a_3, b_1, b_2, b_3$ and b_4 are the coefficients. Here $P(v)$ is the wind power that has been measured between cut-in and rated wind speed. Wind power is constant between rated and cut-out wind speed and also tends to zero before cut-in and after cut-out wind speed.

Further, Thapar *et al* [22] and Ai *et al* [23] have proposed the piecewise polynomial power curve (PPPC) model, which is used to calculate the wind power by dividing the

domain of wind speed into contiguous intervals that can be represented by a separate polynomial in each interval. In each region, a third-degree polynomial has been applied on three points using the least squares method, which achieved a hundred percent accuracy for certain cases. The generalized PPC model equation is shown as Eq. (10):

$$p(v) = \begin{cases} 0, & (v < v_c, v > v_f) \\ a_1 + b_1v + c_1v^2 + d_1v^3, & (v_c \leq v < v_1) \\ a_2 + b_2v + c_2v^2 + d_2v^3, & (v_1 \leq v < v_2) \\ a_3 + b_3v + c_3v^2 + d_3v^3, & (v_2 \leq v < v_f) \end{cases} \quad (10)$$

where a_1, b_1, c_1 and d_1 are the coefficients of cubic equations. Here, v_1, v_2, v_c and v_f are the breakpoints also known as knots. This model is used to predict wind power using the given wind speed since a generalized model can be developed by best fitting lower degree polynomial on an optimal number of data points.

4. Modelling techniques for wind speed forecasting

There are various modelling techniques used for predicting the wind speed, which include optimization techniques, data mining algorithms and neural network activation functions. In this paper, different neural network activation functions are used as a modelling technique for building the hybrid model to predict wind speed. Sigmoid, TANH, Softmax and RELU are some of the modelling techniques used to build the hybrid method and explained here.

4.1 Sigmoid

A sigmoid function is a scientific function having a trademark S-form curve or sigmoid curve. Frequently, sigmoid function alludes to the unique instance of the logistic function. The function is defined using Eq. (11):

$$S(x) = \frac{1}{1 + e^{-x}}. \quad (11)$$

The sigmoid function is a non-linear estimator that is used to store the non-linear property of NLGASX model. Sigmoid functions have a domain of all real numbers, with return value monotonically increasing frequently from 0 to 1 or in another case from -1 to 1, contingent upon tradition [24].

4.2 TANH

Hyperbolic tangent (TANH) is an alternative of sigmoid function [25]. The disadvantage of the sigmoid function is that it can be stuck during training. If the dataset has strongly negative values then its output is near zero. Although this phenomenon is lesser during the training period of TANH, it gives negative outputs for negative inputs and zero outputs for zero inputs. The TANH function

is defined as the ratio of the hyperbolic sine to the hyperbolic cosine function, which is given in Eq. (12):

$$T(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (12)$$

4.3 Softmax

Softmax function is a generalization of the logistic function [26]. In likelihood hypothesis, the output of the softmax function can be utilized for clear cut dissemination, that is a likelihood dispersion over K distinctive conceivable results. As a matter of fact, it is the angle log-normalizer of the clear cut likelihood dissemination. The function is given by Eq. (13):

$$S(x_j) = \frac{e^{x_j}}{\sum_{k=1}^K e^{x_k}}. \quad (13)$$

The softmax function is utilized as a part of different multi-class grouping strategies, for example multinomial logistic regression (otherwise called softmax regression), artificial neural systems, naive Bayes classifiers and multi-class direct discriminant examination.

4.4 RELU

The positive part of an ANN argument is due to the rectifier, which is an activation function that can be represented by numerous definitions [27]. It is defined as shown in Eq. (14):

$$R(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (14)$$

where x is the input value. This activation function has been utilized as a part of a convolutional network more viable than the broadly utilized logistic sigmoid function. Due to the effective training, it is the most popular activation function for the deep neural network. The unit involved in the rectifier is known as the RELU. A simple approximation to the rectifier is the analytical function defined by Eq. (15):

$$F(x) = \log(1 + e^x). \quad (15)$$

It is also called a soft plus function. The first derivative of the soft plus function equals the sigmoid function. RELU function is faster than the sigmoid or similar functions for large and complex datasets.

5. The proposed models

Proposed models of NLGASX and hybrid models are discussed in this section, which are used for wind speed prediction. The NLGASX model is ideal for non-linear

datasets, and hybrid models are suitable for linear as well as non-linear datasets. The hybrid models are formed by combining NLGASX model with modelling techniques.

5.1 NLGASX model

The non-linear generalized autoregressive score model with exogenous factor x (NLGASX model) specifies non-linear structures using a parallel combination of linear and non-linear blocks. The NLGASX function is defined as shown in Eq. (16):

$$y_t = G(y_{t-1}, y_{t-2}, \dots, x_{t-1}, x_{t-2}, \dots) + \epsilon_t \quad (16)$$

where y is the variable of interest; x is the exogenous variable and ϵ_t is an error term. The function G is the non-linear function expressed using the Poisson distribution. When a number of events are given that occur in a fixed time period or space at a steady rate that is independent of the time of the last occurrence, then the Poisson distribution is a discrete likelihood distribution that yields the likelihood of these events. Instead of Poisson family, there are various families available for calculating the nonlinear block such as Student- t , normal, skewed Student- t , exponential, etc. It gives poor performance for linear time-series data.

5.2 Hybrid model

This paper proposes a hybrid model to forecast wind speed combining NLGASX model and modelling techniques based on neural networks activation function (TANH, Sigmoid, SOFTMAX and RELU). The hybrid model is defined as shown in Eq. (17):

$$y_t = F(G(y_t)) \quad (17)$$

where G is the NLGASX function defined earlier and F is a neural network function. There are four types of hybrid models that are formed, namely NLGASX + TANH, NLGASX + Sigmoid, NLGASX + Softmax and NLGASX + RELU. Figure 1 shows the detailed structure of our work.

6. Results and discussion

After developing the model that represents the behaviour of the actual data, the selection of suitable criteria to assess the ability of a model to generalize is significant. In general, mean absolute error (MAE), root mean squared error (RMSE) and R -square are used to measure the performance of wind speed and wind turbine power curve. The MAE, RMSE and R -square are defined in Eqs. (18), (19) and (20), respectively:

$$MAE = \frac{1}{N} \sum_{i=1}^N y(i) - x(i) \quad (18)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y(i) - x(i))^2} \quad (19)$$

$$R\text{-square} = 1 - \frac{\sum_{i=1}^N (y(i) - x(i))^2}{\sum_{i=1}^N (x(i) - \bar{x}(i))^2} \quad (20)$$

where N is the total number of input data, y is the estimated variable, x is the input variable and \bar{x} is the average of the input variable.

The wind power is tested using the variable wind speed, air temperature and air density. The model that gives smaller values of MAE and RMSE as well as the largest value for R -square performs better. The statistical models (ARIMA, GAS, GASX and NLGASX) and the optimization techniques (TANH, Softmax, RELU and Sigmoid) are implemented in python (3.6). There are some other criteria like Akaike information criterion (AIC) and Bayesian information criterion (BIC) that are used for finding the best model fit for a given dataset. According to the AIC and BIC value, it can be established that a particular model performs better than the others.

AIC and BIC: In statistics, the AIC and the BIC are criteria for model selection from a finite set of models [20, 21]. They are based on likelihood estimation. The addition of parameters would improve the likelihood during model fitting but may also result in overfitting. The overfitting problem is solved by the AIC and BIC, which use the penalty term for the number of parameters. The penalty term in BIC is larger than the penalty term of AIC. Equations (21) and (22) present AIC and BIC, respectively:

$$AIC = \ln\left(\frac{\sum_{t=1}^T e_t^2}{T}\right) + \left(\frac{2p}{T}\right) \quad (21)$$

$$BIC = \ln\left(\frac{\sum_{t=1}^T e_t^2}{T}\right) + \left(\frac{p \ln(T)}{T}\right) \quad (22)$$

where T periods of data have been used to fit a model with p parameters and e_t is the residual from the model-fitting process in period t . Additional parameters can be included in the model when the AIC and the BIC penalize the sum of squared residuals. Models that have small values of the AIC or BIC are considered good models.

6.1 Wind speed forecasting models

In this section, the practical implementation of statistical models is discussed. Here, we use the datasets obtained from the resource file of NREL (National Renewable Energy Laboratory) with site id 72509. The geographical

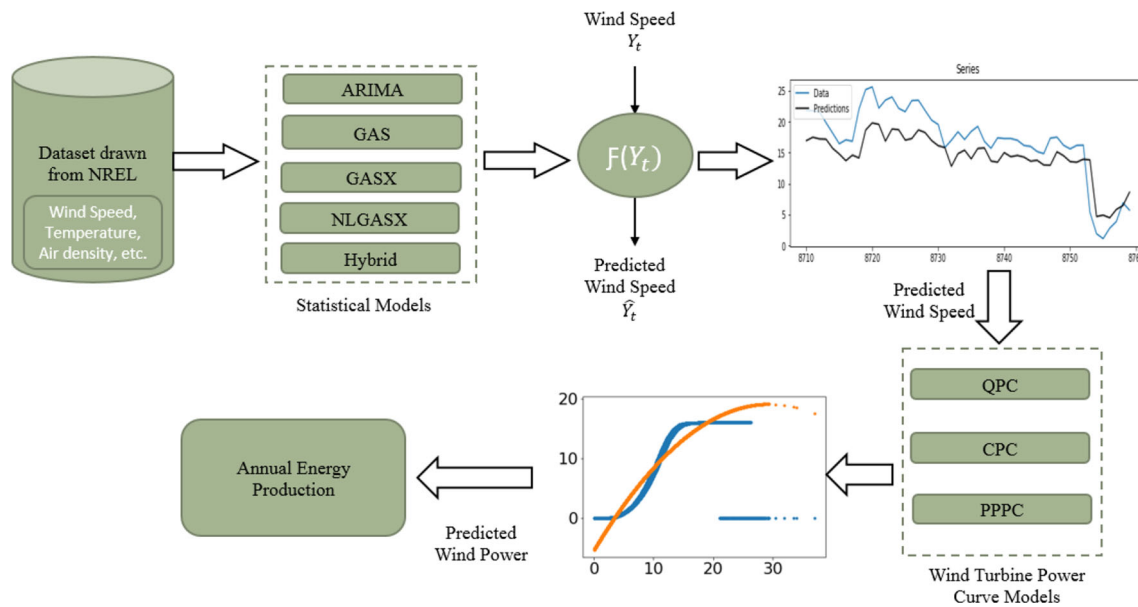


Figure 1. Proposed methodology for wind turbine performance monitoring system.

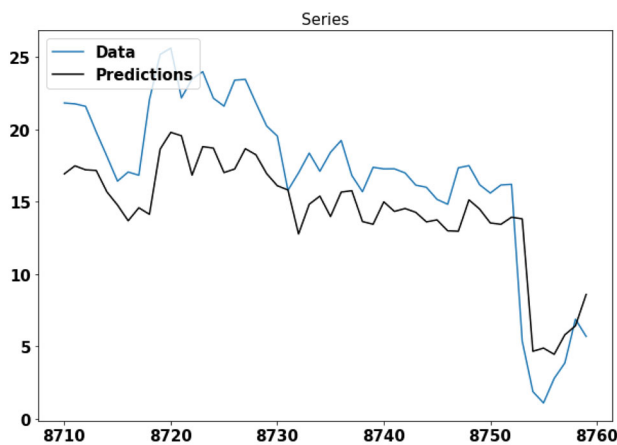


Figure 2. Wind speed prediction using ARIMA (MLE-202) on Dataset-2007.

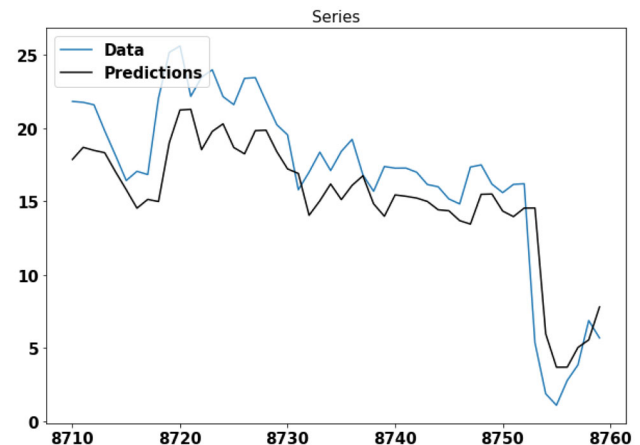


Figure 3. Wind speed prediction using ARIMA (MLE-404) on Dataset-2007.

location of the site is at longitude -106.259064° and latitude 41.775928° , which has an average wind speed of 9.24 and 9.83 m/s for Dataset-2007 and -2008, respectively [28]. NREL includes 5-min average wind speed and corresponding 5-min average power output values obtained from SCADA (supervisory control and data acquisition) system of wind plant recorded at 100-m height, and the capacity factor (CF) equals 0.359. The CF is the ratio of actual energy output to maximum possible energy output in that period. It is used to compare different types of electricity production. There are more than one lakh data observations recorded during the time period of January 2007–December 2007 and also dataset in the year of 2008.

In this work, ARIMA, GAS, GASX and NLGASX models have been applied with MLE method of two versions: MLE-202 and MLE-404. The predicted wind speed and actual wind speed are shown in figures 2 and 3 for the ARIMA model using MLE-202 and MLE-404, respectively, on Dataset-2007. The MLE-404 version shows better performance as compared with MLE-202 version. Figures 4 and 5 present the GAS model, which shows a better prediction as compared with the ARIMA model. Figures 6 and 7 show the actual and predicted wind speed for GASX model and figures 8 and 9 for the NLGASX model. The GASX model shows better prediction as compared with the GAS model. Furthermore, the forecasted wind speed is shown in figures 10–17 for ARIMA, GAS,

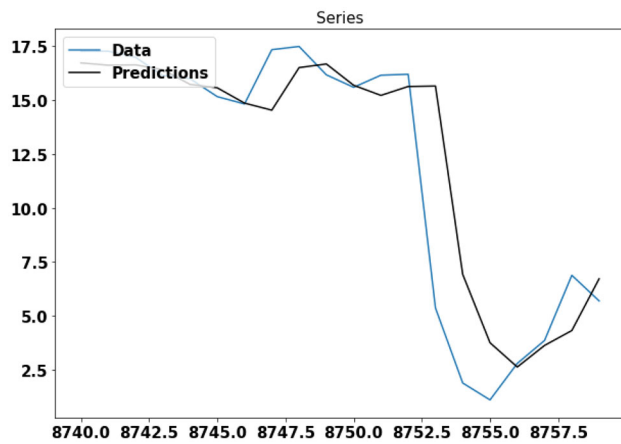


Figure 4. Wind speed prediction using GAS (MLE-202) on Dataset-2007.

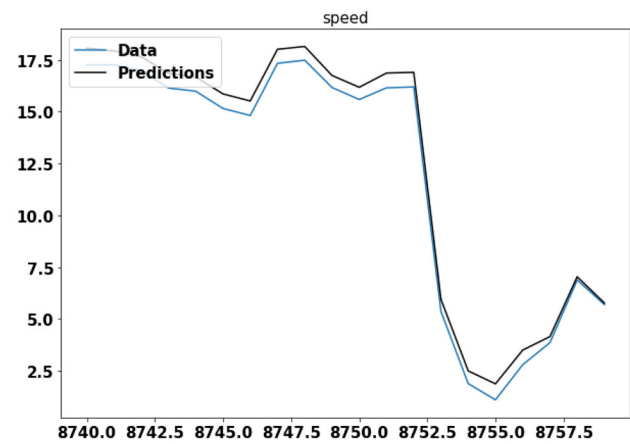


Figure 7. Wind speed prediction using GASX (MLE-404) on Dataset-2007.

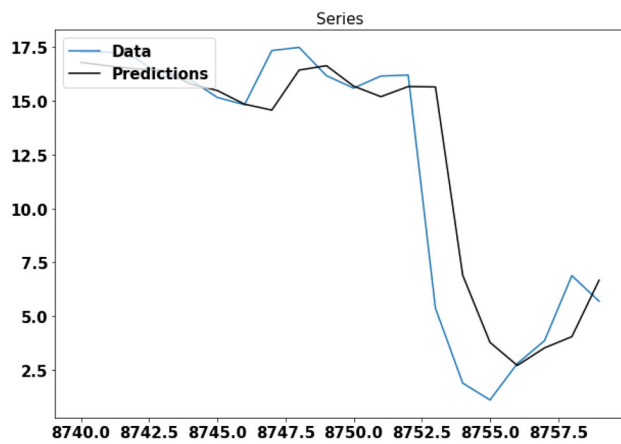


Figure 5. Wind speed prediction using GAS (MLE-404) on Dataset-2007.

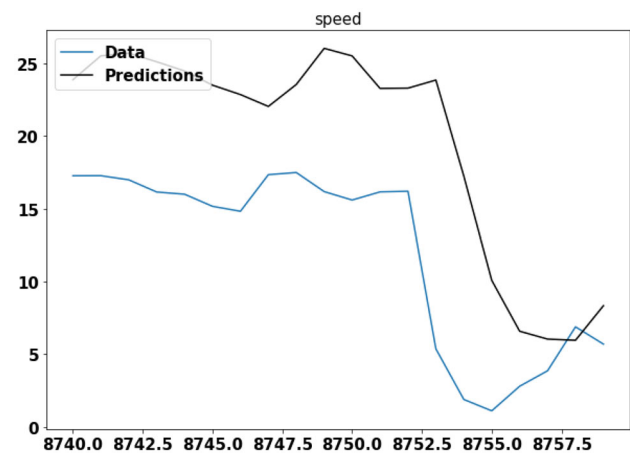


Figure 8. Wind speed prediction using NLGASX (MLE-202) on Dataset-2007.

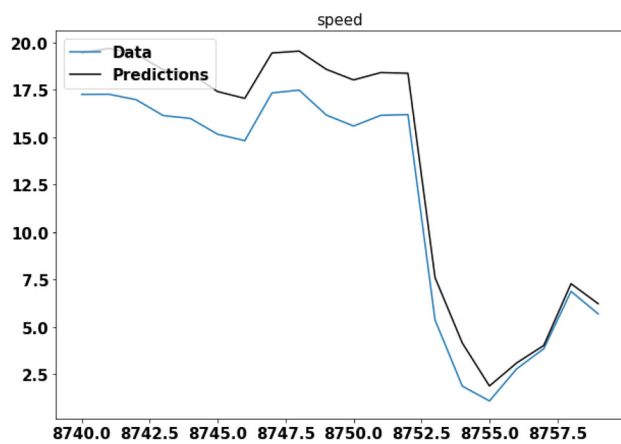


Figure 6. Wind speed prediction using GASX (MLE-202) on Dataset-2007.

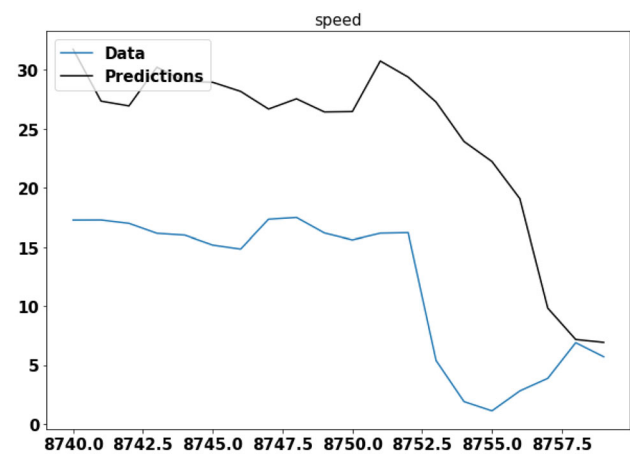


Figure 9. Wind speed prediction using NLGASX (MLE-404) on Dataset-2007.

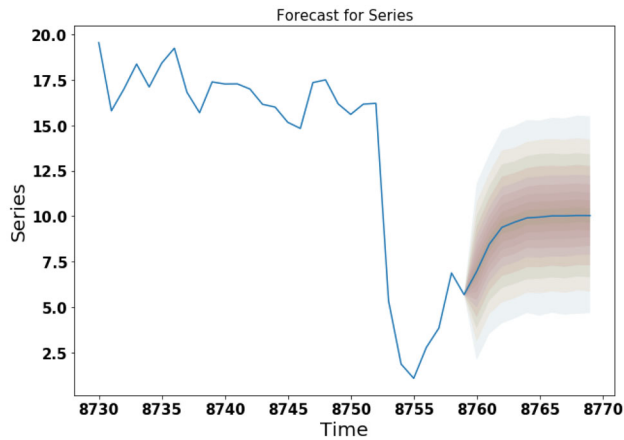


Figure 10. Next 10 values of wind speed forecasting using ARIMA (MLE-202) on Dataset-2007.

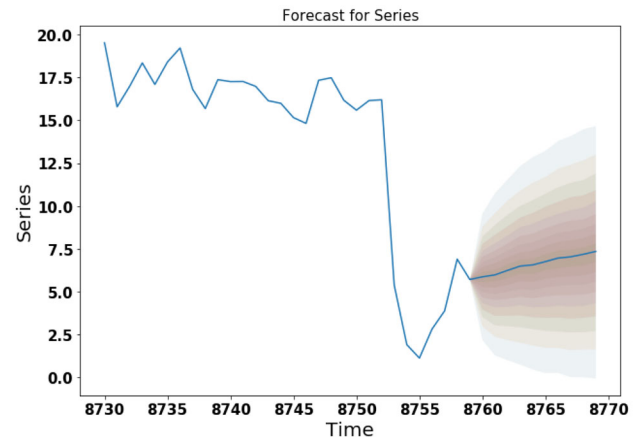


Figure 13. Next 10 values of wind speed forecasting using GAS (MLE-404) on Dataset-2007.

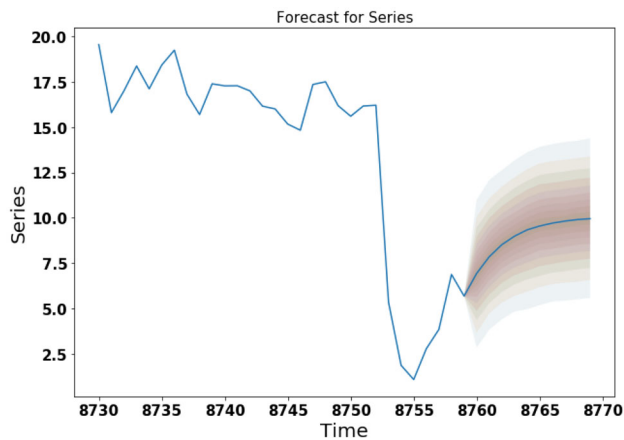


Figure 11. Next 10 values of wind speed forecasting using ARIMA (MLE-404) on Dataset-2007.

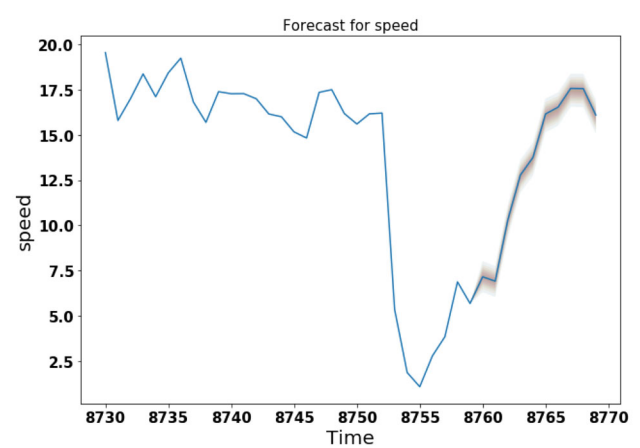


Figure 14. Next 10 values of wind speed forecasting using GASX (MLE-202) on Dataset-2007.

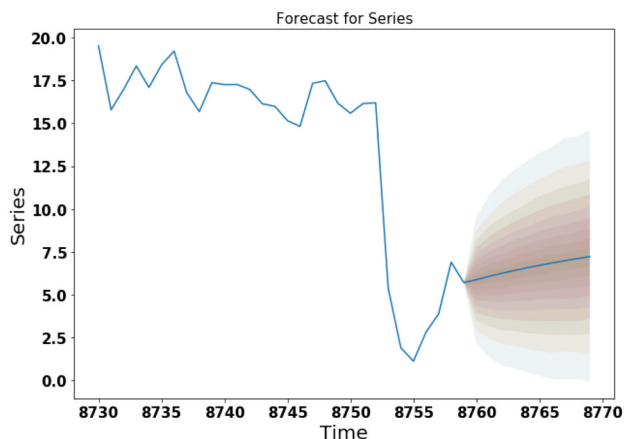


Figure 12. Next 10 values of wind speed forecasting using GAS (MLE-202) on Dataset-2007.

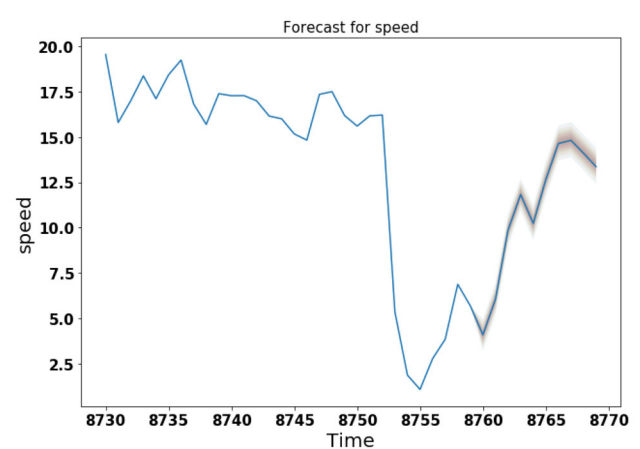


Figure 15. Next 10 values of wind speed forecasting using GASX (MLE-404) on Dataset-2007.

GASX and NLGASX models using MLE-202 and MLE-404, respectively. The performance metrics of the given model are listed in table 1 for Dataset-2007 and in table 2 for Dataset-2008.

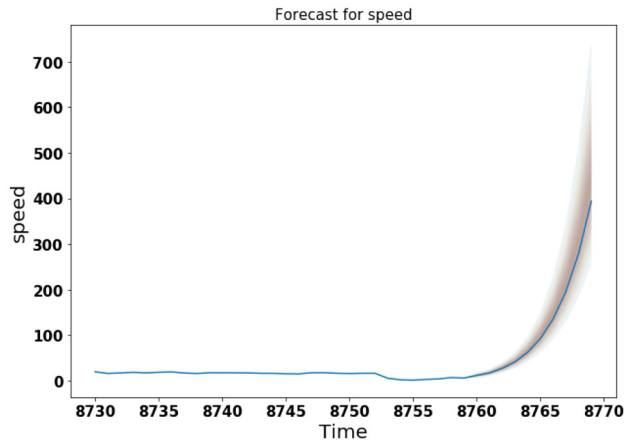


Figure 16. Next 10 values of wind speed forecasting using NLGASX (MLE-202) on Dataset-2007.

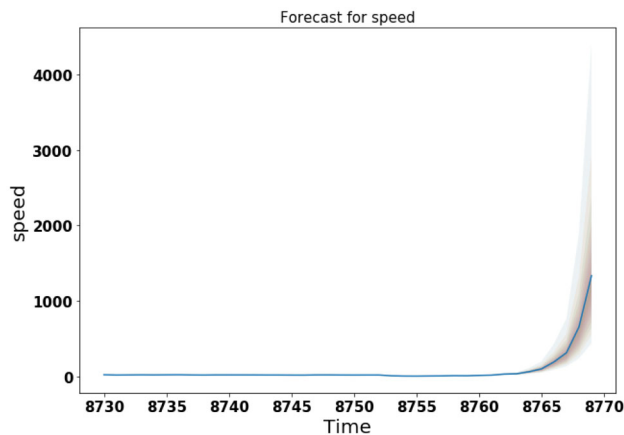


Figure 17. Next 10 Values of wind speed forecasting using NLGASX (MLE-404) on Dataset-2007.

The following factors are considered to build the linear and non-linear forecasting models: wind speed, air temperature and density of the hub. Temperature and density are considered as the exogenous variables. For the proposed model, only temperature is used as the exogenous variable for the experimentations. For Dataset-2007, the hybrid (NLGASX + Sigmoid) model has the lowest MAE, RMSE values and highest R -square value whereas the ARIMA model using MLE (202) records the highest MAE, RMSE values, and lowest R -square value. Hence the hybrid model using Sigmoid modelling technique performs better as compared with other models. In table 1 the GAS model has the lowest MAE/RMSE values as compared with the ARIMA model, so we can say the GAS model performs better than ARIMA model. On comparing GAS and GASX models, GASX performs better than the GAS model. For Dataset-2008, the Hybrid (NLGASX + Sigmoid) model records the lowest MAE, which makes it the best-performing model whereas the ARIMA model gives the highest MAE. In terms of RMSE and R -square value, GASX model records the lowest RMSE and highest R -square, which makes it also the best-performing model. Therefore, GASX model and hybrid model using Sigmoid modelling technique perform well in comparison with other models.

6.2 Wind turbine power curve model

The wind power is calculated using the QPC, CPC and PPPC models. Wind power is calculated for the QPC model using Eq. (8), the CPC model using Eq. (9) and the PPPC model using Eq. (10). The performance metrics of all the models are listed in tables 3 and 4 for Dataset-2007 and Dataset-2008, respectively. Here, the performance is measured using the criterion MAE, RMSE and R -square value for different datasets. The model that has the lowest MAE and RMSE values and highest R -square value is considered as the better-performing model. For both the datasets, results show that the PPPC model has the lowest MAE and

Table 1. Performance of wind speed forecasting models for Dataset-2007.

Model	Criteria	AIC	BIC	MAE	RMSE	R -square
ARIMA	MLE(202)	42771.448	42813.914	3.145	3.528	0.109
ARIMA	MLE(404)	40752.736	40823.511	2.472	2.895	0.430
GAS	MLE(202)	39177.664	39220.130	2.538	2.979	0.786
GAS	MLE(404)	39170.967	39241.742	1.941	2.584	0.795
GASX	MLE(202)	2317.283	2373.905	2.212	2.319	0.806
GASX	MLE(404)	2284.441	2299.511	1.654	2.302	0.807
NLGASX	MLE(202)	57813.262	57876.962	5.912	6.864	0.152
NLGASX	MLE(404)	52528.297	52620.305	5.345	6.637	0.198
NLGASX	TANH	17456.534	14500.079	1.577	2.290	0.808
NLGASX	Softmax	17456.552	14422.086	1.611	2.279	0.809
NLGASX	RELU	17456.566	14359.779	1.573	2.276	0.810
NLGASX	Sigmoid	17456.559	14391.982	1.568	2.271	0.810

Table 2. Performance of wind speed forecasting models for Dataset-2008.

Model	Criteria	AIC	BIC	MAE	RMSE	R-square
ARIMA	MLE(202)	42337.796	42380.262	3.338	4.093	0.121
ARIMA	MLE(404)	40179.142	40249.917	2.721	3.522	0.376
GAS	MLE(202)	38427.485	38469.951	2.089	3.116	0.650
GAS	MLE(404)	38416.291	38487.066	2.086	3.098	0.651
GASX	MLE(202)	3949.869	3893.248	1.778	1.946	0.795
GASX	MLE(404)	3736.894	3821.824	1.749	1.834	0.902
NLGASX	MLE(202)	62635.439	62699.138	4.457	5.503	0.175
NLGASX	MLE(404)	51701.804	51793.811	4.741	6.335	0.145
NLGASX	TANH	17456.749	13561.498	1.529	2.170	0.832
NLGASX	Softmax	17456.766	13486.932	1.497	2.161	0.833
NLGASX	RELU	17456.749	13563.068	1.499	2.170	0.832
NLGASX	Sigmoid	17456.764	13497.118	1.485	2.162	0.833

Table 3. Performance of wind power forecasting models for Dataset-2007.

Model	AIC	BIC	MAE	RMSE	R-square
QPC	210213.5737	173499.8325	1.7850	2.2822	0.8617
CPC	210216.4968	19864.0456	0.6891	1.0989	0.9679
PPPC	210216.9029	1481.5823	0.6277	0.9929	0.9738

Table 4. Performance of wind power forecasting models for Dataset-2008.

Model	AIC	BIC	MAE	RMSE	R-square
QPC	210213.5775	173301.2935	1.7785	2.2800	0.8647
CPC	210216.3815	25924.7641	0.7224	1.1311	0.9667
PPPC	210216.7131	8494.2321	0.6487	1.0411	0.9718

Table 5. Performance of annual energy production.

Model	Dataset-2007	Dataset-2008
QPC	59464.2598	65713.1751
CPC	59464.8145	65317.3461
PPPC	59465.0909	65317.6775

RMSE, and highest *R*-square value, whereas the QPC model has the highest MAE and RMSE, and lowest *R*-square value; therefore, the PPPC model performs better in comparison with QPC and CPC models.

Wind power is calculated using wind speed given in Dataset-2007 and -2008. Wind speed is given in the unit of m/s, and wind power is calculated in MW.

6.3 AEP

The AEP on two different datasets is estimated using Eq. (3) and is presented in table 5, which has the unit of MWh. The wind power is calculated using three different wind power curve models, namely QPC model, CPC model

and PPPC model as discussed in previous sections. The annual energy estimated with Dataset-2007 is the maximum using PPPC model and the least when QPC model is employed. AEP with Dataset-2008 is the maximum using PPPC model and the least using QPC model. Therefore, the performance of PPPC model is better when compared with QPC and CPC models.

The power calculated using PPPC model is better for all the datasets and therefore AEP using PPPC model is also better in comparison with QPC and CPC models. The AEP using QPC model produces the smallest value in comparison with the other models.

7. Conclusion

For the planning of an ideal framework for wind energy, the statistical model for wind speed prediction plays an important role. The proposed statistical and hybrid model is developed for wind speed forecasting, and the wind turbine power curve model is developed for wind power forecasting. In this work, the first non-linear GASX model is

presented. The proposed work focused on the selection of the most suitable model for wind speed prediction and wind power forecasting. For the selection, these proposed hybrid models were evaluated on two datasets (2007 and 2008). The most suitable model was selected to forecast the next 100 value of wind speed accurately. It was observed that the hybrid model based on our NLGASX model using Sigmoid technique performed well for wind speed prediction and piecewise polynomial power curve model performed better for wind power prediction. The proposed hybrid models are suggested to be used by the commercial wind farm to monitor the performance of their operational systems.

References

- [1] Jursa R and Rohrig K 2008 Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. *Elsevier International Journal of Forecasting* 24: 694–709
- [2] Meiqin M, Ling J, Chang L, Hatzigiargyriou N D, Zhang J and Ding Y 2016 A novel short-term wind speed prediction based on MFEC. *IEEE Journal of Emerging and Selected Topics in Power Electronics* 4: 1206–1216
- [3] Lei M, Shiyan L, Chuanwen J, Hongling L and Yan Z 2009 A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews* 13(4): 915–920
- [4] Maatallah O A, Achuthan A, Janoyan K and Marzocca P 2015 Recursive wind speed forecasting based on Hammerstein auto-regressive model. *Applied Energy* 145: 191–197
- [5] Chen D, Li P, Li X, Xu C, Xiao Y and Lei B 2011 Short-term wind speed forecasting considering heteroscedasticity. In: *Proceedings of the International Conference on Advanced Power System*, vol. 2, pp. 884–888
- [6] Kavasseri R G and Seetharaman K 2009 Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy* 34(5): 1388–1393
- [7] Lydia M, Suresh Kumar S, Immanuel Selvakumar A and Edwin Prem Kumar G 2015 Wind resource estimation using wind speed and power curve models. *Renewable Energy* 83: 425–434
- [8] Yunus K, Thiringer T and Chen P 2016 ARIMA-based frequency-decomposed modeling of wind speed time series. *IEEE Transactions on Power Systems* 31(4): 2546–2556
- [9] Pinto T, Ramos S, Sousa T M and Vale Z 2014 Short-term wind speed forecasting using support vector machines. In: *Proceedings of the IEEE Computational Intelligence Dynamic Uncertain Environments (CIDUE) Symposium*, pp. 40–46
- [10] Santamara-Bonfil G, Reyes-Ballesteros A and Gershenson C 2016 Wind speed forecasting for wind farms: a method based on support vector regression. *Renewable Energy* 85: 790–809
- [11] Guo Z, Zhao W, Lu H and Wang J 2012 Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renewable Energy* 37(1): 241–249
- [12] Guo Z, Zhao J, Zhang W and Wang J 2011 A corrected hybrid approach for wind speed prediction in Hexi Corridor of China. *Energy* 36(3): 1668–1679
- [13] Zhang W, Wang J, Wang J, Zhao Z and Tian M 2013 Short-term wind speed forecasting based on a hybrid model. *Applied Soft Computing* 13(7): 3225–3233
- [14] Ardia D, Boudt K and Catania L 2016 Generalized Autoregressive Score Models in R: The GAS Package. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v088.i06> Computation
- [15] Maatallah A, Achuthan A, Janoyan K and Marzocca P 2015 Recursive wind speed forecasting based on Hammerstein Auto-Regressive model. *Applied Energy* 145: 191–197
- [16] Creal D, Koopman S J and Lucas A 2013 Generalized autoregressive score (GAS) models with applications. *Applied Econometrics* 28: 777–795
- [17] Harvey A C 2013 *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*. New York, USA. Cambridge University Press
- [18] Shohoni V, Gupta S and Nema R 2016 A comparative analysis of wind speed probability distributions for wind power assessment of four sites. *Turkish Journal of Electrical Engineering and Computer Science* 24: 4724–4735
- [19] Pallabazzer R 1995 Evaluation of wind-generator potentiality. *Solar Energy* 55: 49–59
- [20] Jangamshetti S and Guruprasada Rau V 1999 Site matching of wind turbine generators: a case study. *IEEE Energy Conversion* 14(4): 1537–1543
- [21] Kazemi M and Goudarzi A 2012 A novel method for estimating wind turbines power output based on least square approximation. *International Journal of Engineering and Advanced Technology* 2(1): 97–101
- [22] Thapar V, Agnihotri G and Sethi V K 2011 Critical analysis of methods for mathematical modelling of wind turbines. *Renewable Energy* 36: 3166–3177
- [23] Ai B, Yang H, Shen H and Liao X 2003 Computer-aided design of PV/wind hybrid system. *Renewable Energy* 28: 1491–1512
- [24] Han J and Moraga C 1995 The influence of the sigmoid function parameters on the speed of backpropagation learning. In: Mira J and Sandoval F (Eds.) *From Natural to Artificial Neural Computation, Lecture Notes in Computer Science*, pp. 195–201
- [25] Krizhevsky A, Sutskever I and Hinton G E 2017 Imagenet classification with deep convolutional neural networks. *Communications of the ACM* 60(6): 84–90
- [26] Goodfellow I, Bengio Y and Courville A 2016 *Softmax units for multinoulli output distributions, deep learning*. Cambridge Massachusetts London, England, The MIT Press, pp. 180–184
- [27] Nair V and Hinton G 2010 Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the International Conference on Machine Learning (ICML)*
- [28] National Renewable Energy Laboratory (NREL) 2007–2008 Western dataset. Site-id 72509, <http://wind.nrel.gov/> Web nrel/