

Research Article

Development of Wind Power Prediction Models for Pawan Danavi Wind Farm in Sri Lanka

Piyal Ekanayake ¹, Amila T. Peiris,¹ J. M. Jeevani W. Jayasinghe ¹
and Upaka Rathnayake ²

¹Faculty of Applied Sciences, Wayamba University of Sri Lanka, Kuliyaipitiya, Sri Lanka

²Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

Correspondence should be addressed to J. M. Jeevani W. Jayasinghe; jeevani@wya.ac.lk

Received 15 April 2021; Revised 12 May 2021; Accepted 28 May 2021; Published 4 June 2021

Academic Editor: Juan C. Jauregui-Correa

Copyright © 2021 Piyal Ekanayake et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents the development of wind power prediction models for a wind farm in Sri Lanka using an artificial neural network (ANN), multiple linear regression (MLR), and power regression (PR) techniques. Power generation data over five years since 2015 were used as the dependent variable in modeling, while the corresponding wind speed and ambient temperature values were used as independent variables. Variation of these three variables over time was analyzed to identify monthly, seasonal, and annual patterns. The monthly patterns are coherent with the seasonal monsoon winds exhibiting little annual variation, in the absence of extreme meteorological changes during the period of 2015–2020. The correlation within each pair of variables was also examined by applying statistical techniques, which are presented in terms of Pearson's and Spearman's correlation coefficients. The impact of unit increase (or decrease) in the wind speed and ambient temperature around their mean values on the output power was also quantified. Finally, the accuracy of each model was evaluated by means of the correlation coefficient, root mean squared error (RMSE), bias, and the Nash number. All the models demonstrated acceptable accuracy with correlation coefficient and Nash number closer to 1, very low RMSE, and bias closer to 0. Although the ANN-based model is the most accurate due to advanced features in machine learning, it does not express the generated power output in terms of the independent variables. In contrast, the regression-based statistical models of MLR and PR are advantageous, providing an insight into modeling the power generated by the other wind farms in the same region, which are influenced by similar climate conditions.

1. Introduction

According to the Sustainable Development Goals adopted by the United Nations in 2015 and the Paris Agreement on Climate Change, the countries are bound to “ensure access to affordable, reliable, sustainable and modern energy for all” by the year 2030. In this connection, it is estimated that about two-thirds of global energy demand needs to be fulfilled by renewable energies by the year 2050 in order to contain the rise of global temperature under 2°C [1]. It is encouraging to note that more and more countries focus on clean energy, low carbon, and high energy efficiency with emphasis on countering harmful effects of climate changes [2]. Fossil fuel generated energy is also discouraged as it is

known to incite global warming, air pollution, and ozone depletion. In the context of adopting renewable energies like bioenergy, tidal, solar, and wind, wind energy has received wide attention in producing electricity due to its availability, cost effectiveness, efficiency, and environmental friendliness [3].

It is known that kinetic energy of the wind is converted into electricity using wind turbines installed in wind farms, which is both economical and innocuous to the environment compared to harmful effects of thermal plants maintained at unaffordable fuel prices, particularly in low- and middle-income countries. Using wind-powered electricity in developing countries is becoming more popular due to their capacity to bear the maintenance cost and the downward

trend in wind turbine costs, given an ideal location for a wind farm assured with strong winds throughout the year. Continuous progress in wind power generation would enable the countries to effectively reduce their dependence on power generated by other sources like coal, fossil fuel, and hydropower, the latter of which is interrupted during dry seasons.

Analyzing the wind power by country, the Global Wind Energy Council reported that the estimated cumulative worldwide wind power generation stands at 651 GW as at the end of 2019, which is an increase of 10% compared to 2018 [4]. The climate think-tank Ember revealed key findings on the status of global electricity production in which wind and solar generation is reported to have risen by 15% in 2019 accounting for 8% of the world's electricity [5]. USA, Russia, China, and India have made remarkable progress in contributing wind power to their national grids. From a regional perspective, the Asia Pacific region has an estimated wind power capacity of 290.6 GW, which is about 44% of total global installations. China and India have led the region with wind power-generated electricity of 26.2 GW and 2.4 GW, respectively [6]. As the wind and solar power is predicted to grow significantly in the years ahead, it is imperative that South East Asia too has to play a significant role in boosting its wind power capacity. Sri Lanka has an estimated 5,600 MW of exploitable wind power potential, of which only about 2.3% has been developed with further 2.3% in progress [7].

A number of research studies have been carried out to predict wind power in the short, medium, or long term using different types of models: physical, data-driven, hybrid, and time-series models. Wind speed is transformed into wind power in physical models that predict wind speed based on equations governing atmospheric motions and weather features, whereas data-driven models use past wind power, meteorological data, and numerical weather prediction measurements to fit a wind power forecasting mapping function [8]. Further, physical models are more appropriate for long-term predictions while statistical models are favored in short-term predictions as their error becomes cumulative in long-term predictions [8]. Moreover, different geographical and weather characteristics like terrains, obstacles, pressures, and temperatures are taken into account in physical models for predicting wind speed and the power generated [9]. Historical data are also used to create a mathematical model in time-series modeling, estimating parameters and verifying simulation characteristics [9].

Applications of artificial intelligence algorithms in wind farms are reported in literature [10]. Gomes and Castro proved that both autoregressive moving average and ANN performed well in forecasting wind power using statistical models [11]. Ouyang et al. trained support vector machine, multilayer perception, k -nearest neighbor, and random forest to develop a combined multivariate model [12]. Artificial Neural Network- (ANN-) based wind power prediction models have been developed in Denmark [13], Senegal [14], North India [15], USA [16], and Tamil Nadu, India [17]. Marginal Weibull distributions have been used to define synthetic wind speed variates in simulating correlated

wind speed and power variates in wind parks, and nonlinear turbine power curves have been applied to obtain power output variates [18]. Similar research studies have been conducted for short-term forecasting of wind power using time-series [19, 20] and hybrid [21] models. Rahmani et al. used swarm intelligence for short-term wind energy forecasting in Iran and showed that better results could be obtained by the hybridization of the two algorithms used in their study, namely, the ant colony optimization and particle swarm optimization [22]. Similarly, a data filtering technique based on wavelet transform and neural network were optimized by using particle swarm optimization algorithm [23]. The temporal-geographical correlations of wind speed in different geographical conditions were studied in China for both linear and nonlinear situations using the linear Pearson coefficient and the nonlinear Spearman rank coefficient and tail correlation coefficient [24]. The performance of wind power prediction models can be assessed by the statistical measures of normalized error, normalized mean biased error, normalized mean absolute error percentage, mean squared logarithmic error, R^2 , explained variance score, and the median absolute error [25].

Very few studies have been undertaken to explore the wind power generation and the effect of climatic variables on it in Sri Lanka. As pointed out by the Asian Development Bank too, Sri Lanka has been found to possess vast potential to exploit wind power as a viable option to generate electricity, to be integrated to the national grid [26]. However, it is not a smooth process owing to many reasons associated with integration of wind power generated electricity to a national grid, as described below. Accurate predictions on the wind power to be produced by the turbines are required for the grid operators from very short time scales to a few days ahead, in order to strike the balance between power supply and demand. Uncertainties and frequent fluctuations in wind speed and direction are detrimental to the increase of wind power within the national grid, because of their impact on the variation of power generated at a wind farm (turbines). Further, the temporal, geographical, and climatic factors like the particular time of the day, terrain, humidity, and temperature also affect the wind power generation that depends on generator hours of wind turbines, speed, and direction of the wind, air density, and position of the turbines. If the wind speed changes could be precisely forecast, it could bring about multiple benefits in adjusting control systems, maintenance scheduling, reducing economic and technical risks, and profit maximization of the power trade. A recent study has pointed out that Sri Lanka possesses about 20% of its land in coastal areas, to be utilized for wind power generation [27]. In [28], an adaptive filter was used for forecasting power fluctuations of a wind turbine installed in Kalpitiya area of the north-western region of Sri Lanka, which predicted its power generation with 5.07% root mean square error of its mean value. In this technique, filter weights are constantly modified with the fluctuation of wind pattern, which claimed better predictions over the time-series method. ANN-based models have also been developed to forecast the wind power generation of a wind farm installed in the same region resulting in $R > 0.91$, $MSE < 0.22$,

and bias < 1 [29]. However, none of the aforementioned research presents a formula for the predicted power in terms of the independent variable.

Due to the importance attached to the wind speed among all numerical weather prediction data, the objective of this research was focused on studying the variation of wind power generated by an onshore wind farm with the wind speed and temperature, based on five years of data since 2015. This is essential as the country needs an initiation of research related to wind power. Therefore, this paper presents a novel research study based on the wind farm Pawan Danavi to forecast its wind power generation based on the available and most important climatic factors. In this sense, this study provides an insight into the other existing wind farms in Sri Lanka as well as to the proposed wind power farms, to understand the future wind power generation.

The next section will describe both wind farm and meteorological data. Further, it presents the techniques used, namely, the Feed Forward Artificial Neural Network (ANN), Multiple Linear Regression (MLR), Power Regression (PR), and the evaluation criteria of the models. Section 3 presents the results obtained with a discussion and the conclusions are summarized in Section 4.

2. Materials and Methods

2.1. Wind Farm Data. This research presents the analysis of wind power generation for the “Pawan Danavi” wind farm, which is located in Kalpitiya (around $08^{\circ}02'56''N$ $79^{\circ}43'08''E$), north-western province in Sri Lanka. This geographical area was identified as one of the best locations in Sri Lanka to establish a wind farm [26]. Therefore, wind farm called Pawan Danavi was constructed and connected to national electricity grid for its operation in August 2012. The wind farm is constructed onshore and has 10.2 MW nameplate capacity with 12 wind turbines (model: Gamesa G58-850). The height of each tower is around 65 m, and the diameter of the blades is around 58 m.

Each wind turbine has three blades and its rated power of is 850 kW. The generation voltage of the plant is 690 V (AC) and it is stepped up using a transformer to 33 kV for connecting to the Sri Lanka’s national power grid, which has a transmission voltage of 33 kV. The rotational speed of the rotor varies between 19.44 and 30.8 rpm and the rated stator current at 690 V is 670 A. The standard power factor at the generator output terminals at low-voltage side before transformer input terminals is 0.95 at partial loads and 1 at nominal power.

Monthly average power generation data from January 2015 to December 2019 were obtained from the wind farm authorities (Lanka Transformers Private Limited, Sir Lanka). Figure 1(a) exhibits the variation of generated power over the 5 years. The peak power generation can be seen in the months of June followed by July and August, respectively (refer to Figure 1(b)). The power generation as well as the variation is less than 40 kW from November to April. During the past 5 years, the power generation varied between 113 kW and 3,064 kW with a mean of 1,588 kW. The highest

variation in power generation was observed in October, which varies between 120 kW and 1305 kW.

The statistics of the wind power generation can further be seen in Table 1. On average, May to September, where Sri Lanka is experiencing the southwestern monsoon winds, are having relatively higher wind power generation. Out of them, June has the highest wind power generation and the month’s minimum is also higher than most of the maximums in other months.

2.2. Meteorological Data. Climate parameters of mean wind speed (WS in m/s) and mean ambient temperature (T in $^{\circ}C$) were considered input variables in the modeling process. Their statistics were used with the average power output (MW) data in the period of January 2015 to December 2019. The monthly climatic data were collected from the wind farm, as there is an in-house meteorological station. The variations of mean wind speed and mean ambient temperature are illustrated in Figure 2(a). Both weather indices indicated the peak values in the mid of each year. The wind speed varied between 11 m/s and 2.4 m/s having a mean of 6 m/s. The highest variation in wind speed was observed in December (3.1–8.6 m/s) followed by October (2.7–6.9 m/s). More importantly, the wind speed in August was somewhat steady (8.2–8.9 m/s). In addition, the ambient temperature was varied between $40.7^{\circ}C$ and $33.4^{\circ}C$ with a mean of $37^{\circ}C$. These statistical particulars for the period of 2015–2020 can also be seen in Figure 2(b).

2.3. Wind Power Generation and Climatic Factors. The kinetic energy of the wind depends on the density of the air; thus, the wind power generation is also proportional to the air density. Therefore, the wind power generation can be formulated as shown in the following equation:

$$P = 0.5D_{\text{air}}AV^3, \quad (1)$$

where D_{air} , A , and V are the air density (in kg/m^3), the swept area of wind turbine (in m^2), and the upstream wind speed (in m/s), respectively [25]. The density of the air can be represented as function of d , density of the dry air at atmospheric temperature (at $25^{\circ}C$, $d = 1.168 kg/m^3$), T , the absolute temperature (in kelvin), B , the barometric pressure (in torr; 1 atm = 760 torr) and e , the vapor pressure of moist air (in torr) and given in the following equation [30]:

$$D_{\text{air}} = d \left(\frac{273.15}{T} \right) \times \left(\frac{B - 0.3783e}{760} \right). \quad (2)$$

The power output of the turbine depends on the wind capacity because the garnered power is cubic proportional to the wind speed. However, the air density is inversely proportional to the absolute temperature and directly proportional to the barometric pressure. Thus, the air density is less influential in comparison to wind variations.

In addition, the generated wind power depends on the average wind direction too. However, the rotor is rotated based on the wind direction. Furthermore, the blades are angled, so that the energy generation is optimum. The wind

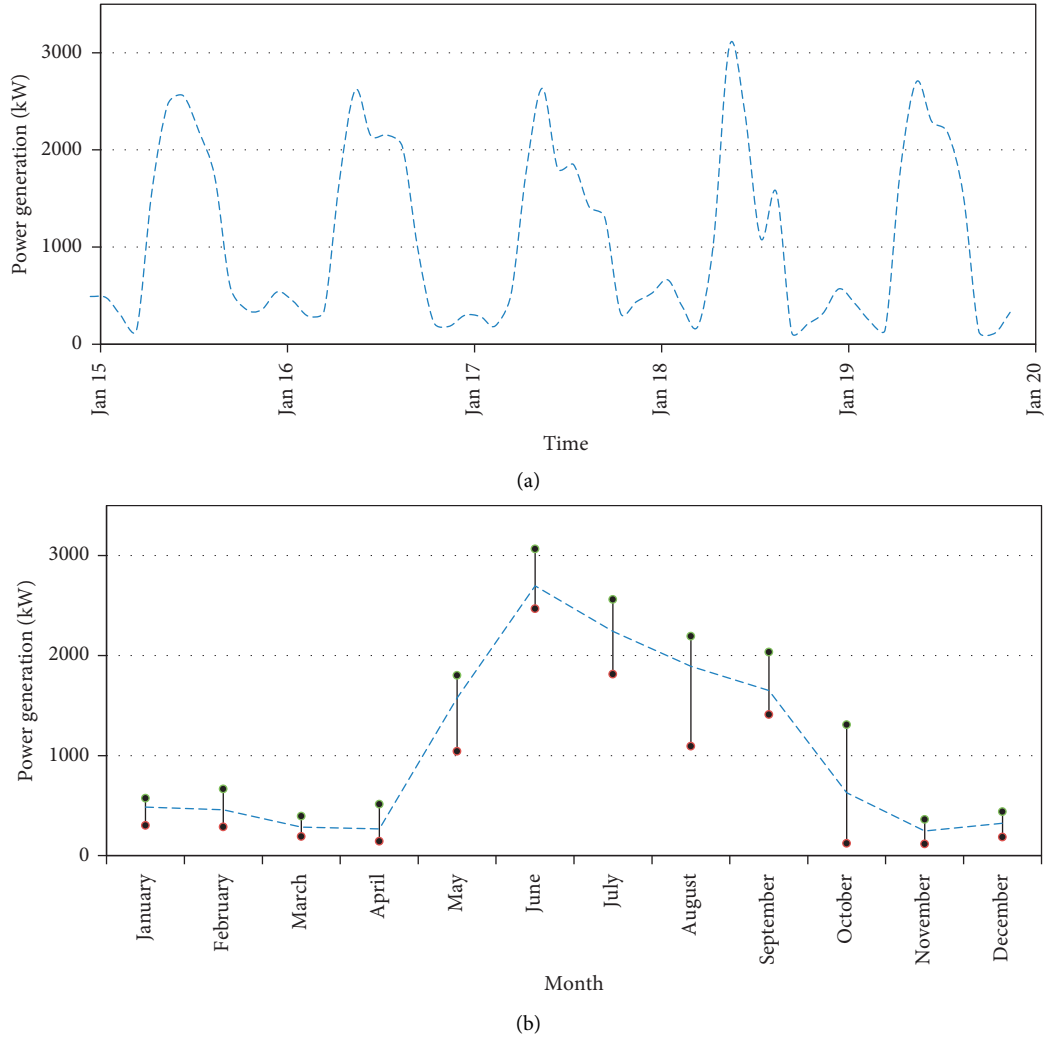


FIGURE 1: Variation of mean wind power generation. (a) Over the 5 years. (b) Averaged variation over a year.

TABLE 1: Variation of mean wind power generation (in kW).

Statistical parameter	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	485	457	283	266	1575	2697	2243	1893	1650	627	245	324
Minimum	298	283	188	142	1040	2466	1811	1091	1409	120	113	184
Maximum	570	663	391	512	1801	3065	2560	2191	2033	1305	360	435

rose diagram developed based on the wind direction and speed data over the past 5 years is illustrated in Figure 3. It reveals that the dominant wind direction for the highest wind contribution lies between 157.5° and 202.5° (south southeastern to south southwestern) while the wind with speed higher than 7.5 m/s mainly lies between south and south southeastern.

The 3D surface plot of the power generation with respect to average wind speed and ambient temperature can be seen in Figure 4. The relationships to the wind power generation with respect to average wind speed and ambient temperature

can be clearly observed. When the ambient temperature and the wind speed are higher, the power output is also higher.

2.4. Modeling Techniques

2.4.1. Feed-Forward Artificial Neural Network Based on Levenberg-Marquardt Method. Error back-propagation mechanism is used to adjust the learning parameters in the feed-forward neural networks. This process goes until the stopping criterion is achieved. The learning parameters can

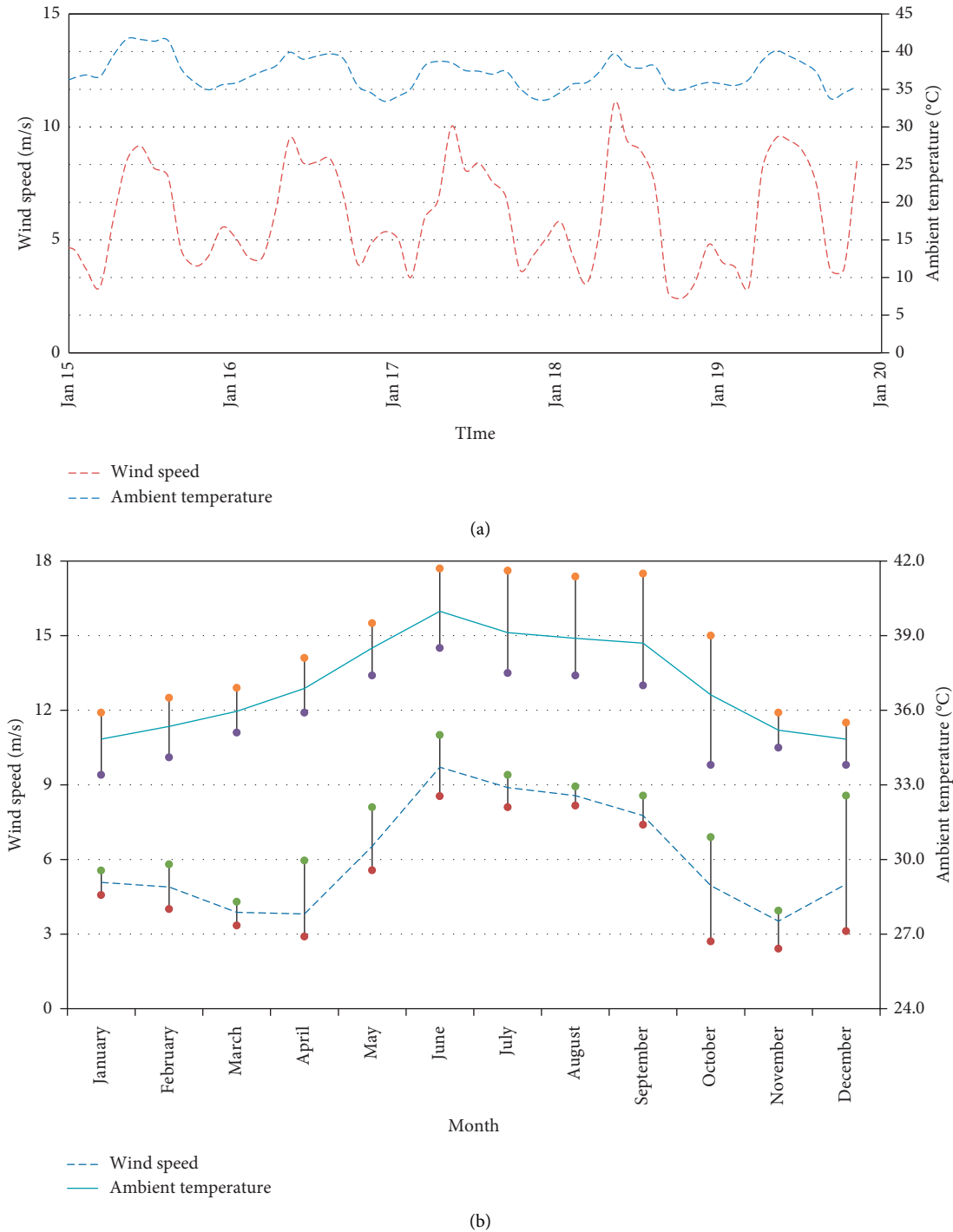


FIGURE 2: Variation of independent variables. (a) Over the 5 years. (b) Averaged variation over a year.

be adjusted based on the information used in the error surface. The information also can be local or global based on the problems. In addition, the curvature of the error surface is optional in use. Simple Gradient Descent (GD) was used for a long time as the most widely used learning algorithm. Modifications like momentum and adaptive learning coefficients were incorporated to overcome the poor convergence rates of the learning algorithm [31]. However, the learning algorithms based on Gauss–Newton (GN) method

have shown significantly improved convergence rates. These 2nd-order methods use the information on error surface for the convergence [32].

Levenberg–Marquardt method (LM) is one of the widely used learning algorithms in the related research to train the neural networks. This algorithm interpolates between two methods discussed above, GD and GN. Therefore, the training algorithm takes the merits of both GD and GN methods to produce stable and fast converging solutions.

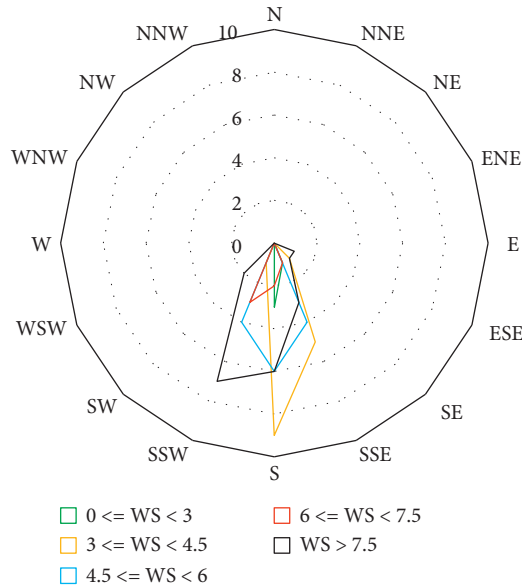


FIGURE 3: Wind rose diagram.

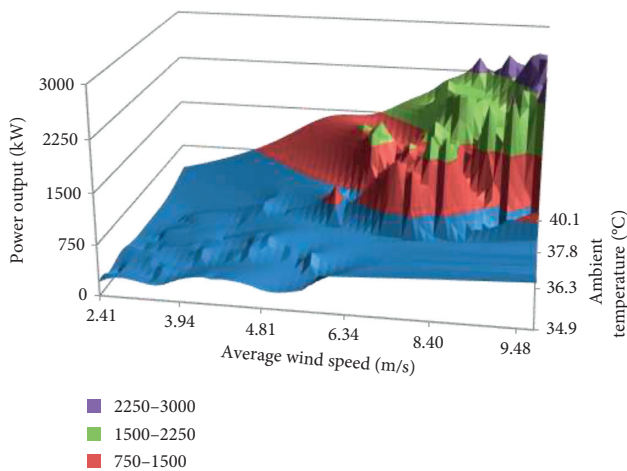


FIGURE 4: 3D surface plot of the variables.

The learning parameter optimization is used as the goal to minimize the objective function of the multilayer feed-forward neural networks. The iterative process is continuously carried out until the optimization meets the predefined value. Therefore, the technique gives excellent results for the smaller-scale problems; however, it is inefficient in larger-scale problems due to the computational complexity, memory usages, and error oscillations [32].

Nevertheless, the iterations in LM training algorithm always decrease the objective function unlike many other algorithms like GD, conjugate gradient, resilient back-propagation, etc. Therefore, from one iteration to the next iteration, there will be a decrease in the objective function. This is a merit of the algorithm; however, it also can lead to heavy computational costs. Within an iteration, the algorithm searches for the new parameters as required to decrease the objective function and therefore, the computational cost becomes higher. Nevertheless, LM

training algorithm is widely used by many researchers to solve nonlinear problems where the output is important rather than the methodology. More information on the LM algorithm can be found in Tomislav et al.'s work [33].

2.4.2. Multiple Linear Regression (MLR). The relationships which have reasons and result relations can effectively be estimated using the regression analysis. The linear relationship is formed as an equation in between the mainly one dependent variable with respect to the set of independent variables. Regression models with one dependent variable with a set of independent variables are the multilinear regression [34]. The multivariate regression model can be mathematically formed as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon, \quad (3)$$

where β_0 is the y -intercept (a constant term), β_i is the slope coefficient for each input variable; and ε is the model deviation. This showcases the progression of simple linear regression model to the multivariate linear regression (where the first two sections can be found in single regression ($y = \beta_0 + \beta_1 x_1$)). Therefore, the two linear models have to be carefully used in various calculations, however, while understanding variable behaviors. Therefore, simply the linear term emphasizes linearity on the variables without moving to any higher orders [35].

Thus, the multivariate regression assumes the data set to be normally distributed, without extreme values and no multiple ties between independent variables. These assumptions can be assessed by examining the residuals, having fitted a model.

2.4.3. Power Regression (PR). Power Regression (PR) is a nonlinear regression model, where the output is modeled in proportion to a power of the explanatory variables. The regression can be expressed as shown in the following equation:

$$y = ax_1^b x_2^c \dots x_i^p, \quad (4)$$

where i is the number of observations and a, b, c, \dots, p are constants. Therefore, the power regression plots are usually exponentially upward or downward in shape. The regression is reliable to predict values within interpolation limits; however, it is less reliable in extrapolation. The negative x -values are inappropriate in this setting because raising them to an arbitrary power can produce complex results that satisfy the minimization problem, but do not correspond to the real y -values. Power regression has been used by many researchers for a long time and it is continued to be used for many engineering applications [36–38].

2.5. Evaluation Criteria

2.5.1. Pearson Correlation Coefficient. Linear dependency of the random variables is usually measured by the Pearson correlation coefficient. The following set of equations showcase the definition of the Pearson correlation

coefficient. Two random variables are proposed as X and Y , with their expectations $E(X)$ and $E(Y)$, respectively. The variances $D(X)$ and $D(Y)$ are defined as given in the following equations:

$$D(X) = E[X - E(X)]^2, \quad D(X) > 0, \quad (5)$$

$$D(Y) = E[Y - E(Y)]^2, \quad D(Y) > 0. \quad (6)$$

The covariance ($cov(X, Y)$) of the variables can be formulated as follows:

$$cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\}. \quad (7)$$

Therefore, by definition, the Pearson correlation coefficient (R) can be presented as follows:

$$R = \frac{cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (8)$$

The numerical value of the Pearson coefficient is in between -1 and $+1$. When it is $+1$, there is a perfect, positive linear correlation in between the two variables. However, if it is -1 , the correlation is perfect, but it is negative. When this is 0 , there is no linear correlation between the variables. The Pearson correlation coefficient is easy and convenient to calculate; however, it is poor in performance for the non-linear relationships. Nevertheless, a zero correlation does not mean that the variables are totally independent. Therefore, a multivariate normal distribution is recommended to analyze true independence of the variables. More information can be found in Lee et al.'s work [39].

2.5.2. Spearman's Correlation Coefficient. The Spearman correlation coefficient is a nonparametric test, but similar to the Pearson correlation coefficient. It measures the strength and the direction of correlation of two variables [40]. Unlike the Pearson correlation coefficient, Spearman's correlation coefficient can assess the monotonic relationships and the relationships need not be linear. However, a similar presentation for the assessment can be made based on the numerical values of Spearman's correlation coefficient (-1 or $+1$). The coefficient has a higher numerical value when there is a similar rank between two variables and vice versa. Therefore, the coefficient is defined as the Pearson correlation coefficient of the ranked variables.

The coefficient (r_s) is formulated as follows:

$$r_s = \rho_{rg_X, rg_Y} = \frac{cov(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}}, \quad (9)$$

where n , X_i , and Y_i are the sample size and two raw sources of variables, respectively. rg_X , rg_Y , $cov(rg_X, rg_Y)$, σ_{rg_X} , and σ_{rg_Y} are the ranks corresponding to the two raw source variables, covariance of the rank variables, and the standard deviations of the rank variables, respectively. In addition, ρ denotes the usual Pearson correlation coefficient, but applied

to the rank variables. If all n ranks are distinct integers, then the coefficient can be computed as follows:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \quad (10)$$

where $d_i = rg(X_i) - rg(Y_i)$ is the difference between the two ranks of each observation and n is the number of observations. More details of the Spearman correlation coefficient can be found in Zar's work [40], and the coefficient was extensively used in related research [41–43].

2.5.3. Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). The squares of the errors are averaged to obtain the Mean Squared Error (MSE). The error is the difference between the estimated (or predicted) values and the corresponding actual (or observed) values. Therefore, the lower the MSE, the higher the accuracy of the prediction. The MSE is a risk function, and it is always positive and the zero is the lowest [44]. There is a significant disadvantage in the MSE. It sometimes has a higher order compared to the order of the data set. Therefore, as a solution to this disadvantage, Root Mean Squared Error (RMSE) is proposed and showcased in the following equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}}. \quad (11)$$

RMSE is a widely used assessor to analyze the accuracy of predicted nature against the observed nature of physical problems. Similar to MSE, RMSE is nonnegative and 0 represents the perfect match. However, it is impossible to have a perfect match in the real world. Lowering the RMSE produces better prediction models. However, comparisons across different types of data would be invalid because the measure is dependent on the scale of the numbers used. Therefore, RMSE is a measure of accuracy, to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent [45].

2.5.4. Bias and Nash Number. Bias measures the overestimation tendency or underestimation tendency of predictive parameters. Mean error in statistics is the bias. Therefore, it averages all possible errors. Uncertainty in a measurement or the difference between observed and predicted values are considered errors. The bias gives poor results for both statistical and physical methods. In addition, it does not showcase the degree of precision of the forecasting method [46]. The bias is mathematically presented in equation (12) in the usual notations described earlier:

$$bias = \frac{\sum_{i=1}^N (y_i - x_i)}{N}. \quad (12)$$

In addition, the Nash number is presented as follows:

$$Nash \text{ number} = 1 - \left[\frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - x_{average})^2} \right]. \quad (13)$$

It presents a measure of error variance of model (predicted) by the variance of observed. However, Nash number of 1 gives the perfect system and thus no error variance in the modeled (predicted) time-series [47].

2.6. Wind Power Model Development. Wind power generation data, wind speed, and temperature data were modeled based on equation (14) to identify the wind power generation prediction model. Three methods, ANN, MLR, and PR, were used to develop wind power generation models based on the available data.

$$\text{wind power} = f(\text{wind speed, ambient temperature}). \quad (14)$$

3. Results and Discussion

3.1. Correlation Results. Both independent variables, viz., wind speed and temperature, exhibit a normal distribution as can be seen in Figures 5(a) and 5(b). This is the most acceptable probability distribution for real-world scenarios. In addition, each of them has an acceptable linear correlation with the generated wind power, which is the dependent variable (refer to Figures 6(a) and 6(b)).

Although some outliers are present in Figure 6, both figures indicate an acceptable degree of linear behavior, wherein the wind power exhibits a more scattered distribution against the temperature compared to the better linearity between wind power and wind speed.

The correlation among each variable was investigated in terms of the Pearson correlation (refer to Table 2). The correlation between the wind power and the wind speed is the highest with 0.92, while the second highest correlation exists between wind power and the temperature. The correlation between the independent variables is 0.73. Similar observations can be seen for the Spearman correlation values (r_s) too. Therefore, the correlations among all the variables can be suggested strong.

Variation of the dependent variable around the mean power with reference to the variation of independent variables around their mean values was analyzed and its results can be seen in Figure 7. Both variations are positive. A unit increment (or decrement) in the wind speed (in m/s) increases (or decreases) the power generation by about 355 kW. The higher determination coefficient of the trend line (0.84) indicates its validity. The increase (or decrease) in temperature by 1°C increases (or decreases) the power generation by about 349 kW.

3.2. Wind Power Prediction Models. An ANN model was developed by applying the LM algorithm 70% of the data was used for training. The rest of the data were distributed equally for testing and validation. The coefficient of correlation for training, validation, and test was found as 0.99, 0.92, and 0.96, respectively. Predicted power values obtained by applying ANN were plotted against the actual wind power (refer to Figure 8(a)). The overall coefficient of correlation of

the ANN model was 0.97. The predicted power values (Figure 8(a)) prove the accuracy of the ANN model with slight or no deviations from the actual wind power values.

In contrast, some predicted power values of the regression models deviate considerably from the corresponding actual power values. However, MLR (refer to Figure 8(b)) and PR (refer to Figure 8(c)) models also exhibit a correlation coefficient of 0.95 and 0.94 proving their high accuracy. The distribution of error percentage too proves that the ANN-based model outperforms the regression-based models (refer to Figure 8(d)). Among the regression-based prediction models, MLR outperforms PR having a majority of errors with error percentages less than 25%.

The accuracy of the models was evaluated in terms of MSE, bias, and Nash number. The comparison results are presented in Table 3. All three models have very low RMSE and bias values. The Nash number of the ANN model is very close to 1, demonstrating its superior accuracy while that of the regression models is about 0.9, which is acceptable.

Table 3 showcases better Nash number performance in ANN. In addition, it has the minimum RMSE value among other two tests carried out. However, there is a drawback in the ANN architecture. The prediction model is incapable of developing a mathematical relationship between the dependent and independent variables. In other words, the mathematical behavior of equation (14) is inaccessible from ANN. Therefore, the developed ANN model has to be saved and rerun whenever needed.

This drawback can be eliminated with other two developed models, MLR and PR. Equations (15) and (16) were derived from those two models.

$$\text{Wind power}_{\text{MLR}} = -6053.4 + 252.7WS + 150.7T, \quad (15)$$

$$\text{Wind power}_{\text{PR}} = (2 \times 10^{-8}) \times WS^{1.69} T^{5.9}. \quad (16)$$

The ability to express the wind power generation in terms of the independent variables is a greater advantage in the prediction models developed based on the MLR and PR techniques. More importantly, the performance of the models is acceptable. Therefore, MLR- and PR-based models have a merit over the ANN model.

3.3. Comparison with Wind Power Models. Some research studies conducted on wind farms located in other countries express the power generation in terms of variables (Table 4). In most of the studies, a single variable of wind speed was considered to develop linear [48] or quadratic [48, 49] relationships. Turbine details were also incorporated to express the predicted power [50]. However, it is useful only when different types of wind turbines are used in a wind farm. The predicted power can be expressed in terms of the current and previous power generation data in time-series analysis [51]. The model proposed in this research is more comprehensive in this sense as the ambient temperature has also been incorporated in addition to the wind speed to predict the wind power.

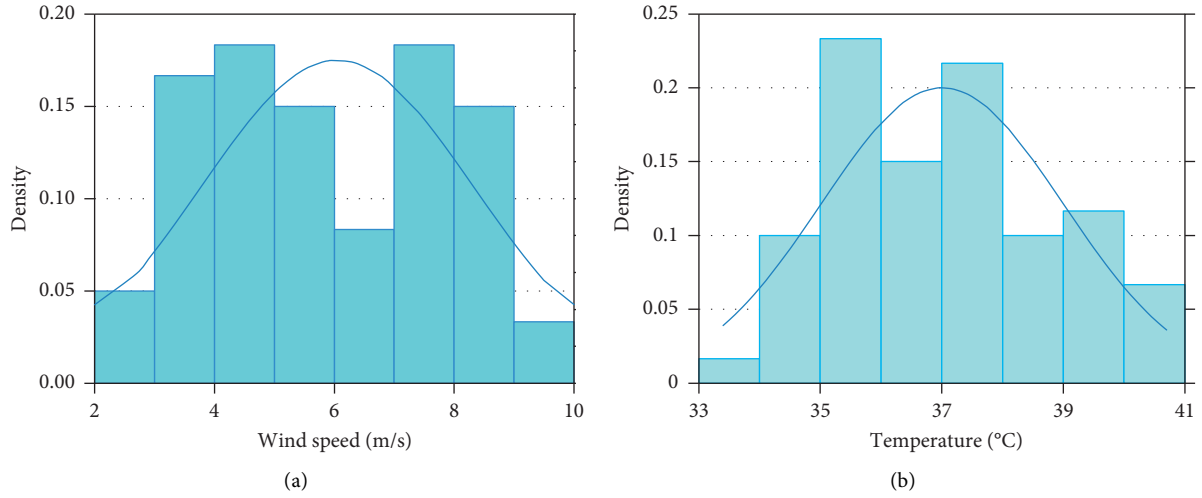


FIGURE 5: Normal distribution of the independent variables. (a) For wind speed. (b) For temperature.

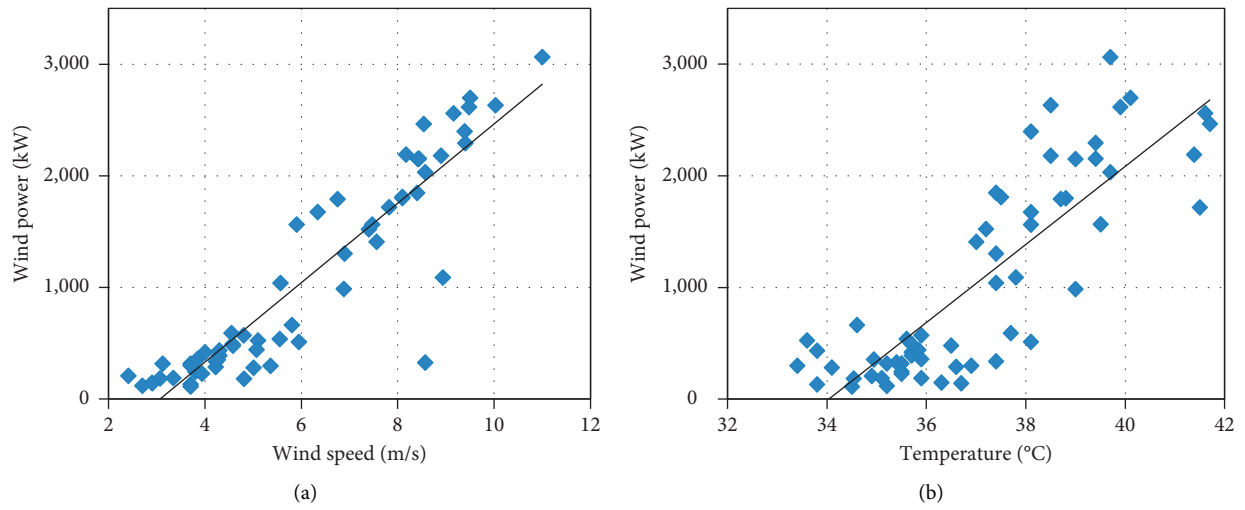


FIGURE 6: Correlation between the dependent variable and the independent variables. (a) For wind speed. (b) For temperature.

TABLE 2: Pearson’s and Spearman’s correlation matrix.

Wind speed	$R = 1, r_s = 1$	—	—
Temperature	$R = 0.73, r_s = 0.73$	$R = 1, r_s = 1$	—
Output power	$R = 0.92, r_s = 0.91$	$R = 0.84, r_s = 0.82$	$R = 1, r_s = 1$
—	Wind speed	Temperature	Output power

The ANN model presented in this research paper has better performance compared to the models developed by applying artificial intelligence algorithms [10]. However, the performance of the statistical methods presented in this study is slightly lower than that of the ANN model. Nevertheless, all the models proposed here are highly accurate with $R > 0.94$. The RMSE values of the models proposed in [48] are comparable to those of the ANN-based model of this

paper while the same of the Kalman filter based model in [49] is comparable to those of the MLR- and PR-based models introduced here. All three models of the present work produced coefficients of correlation higher than that generated by the Kalman filter-based model [49]. Accordingly, the performance of the proposed statistical models is also comparable to that of the wind power models presented in similar research studies. More importantly, they revealed

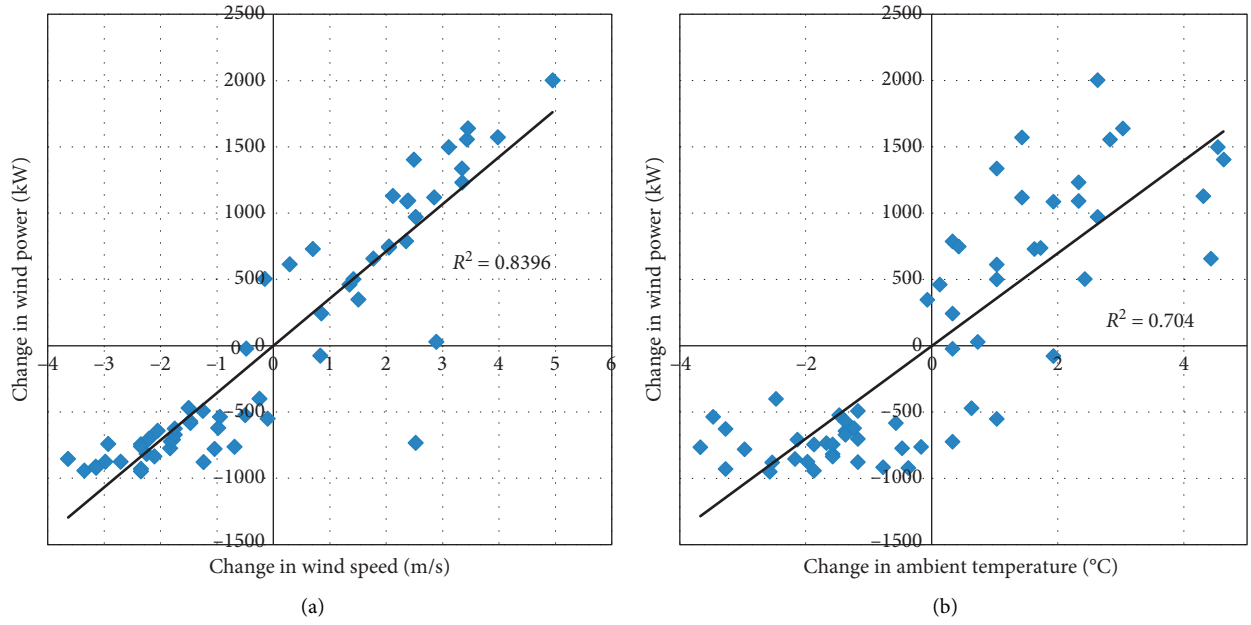


FIGURE 7: Variation of the dependent variable around the mean power. (a) With the variation of wind speed around its mean. (b) With the variation of ambient temperature around its mean.

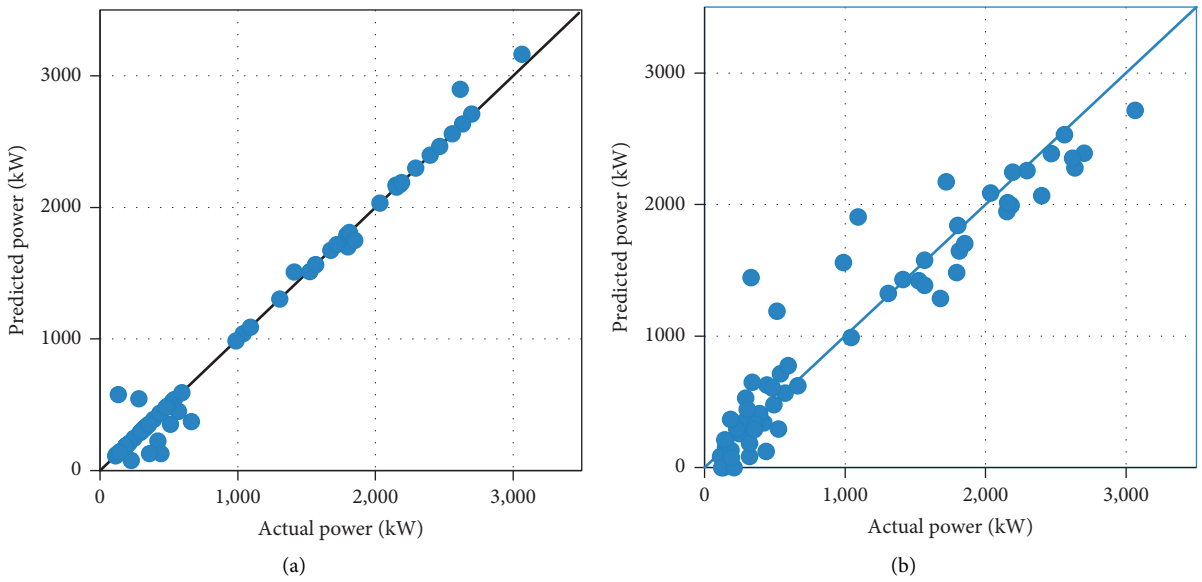


FIGURE 8: Continued.

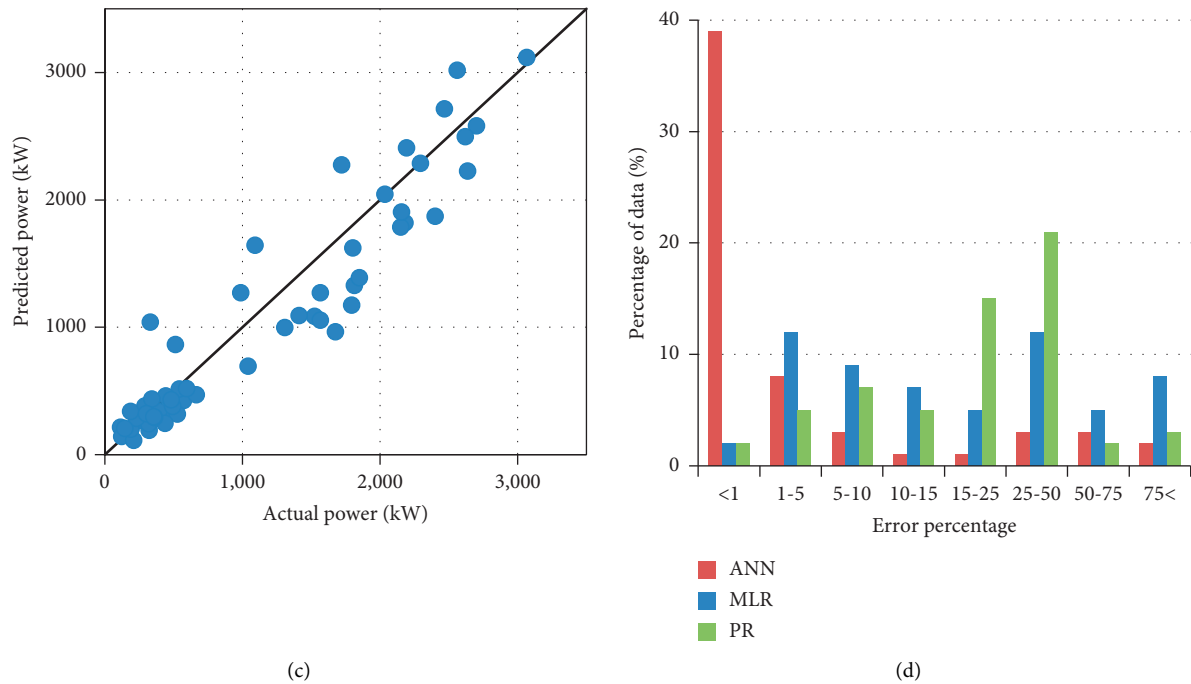


FIGURE 8: Performance of the models. (a) For ANN. (b) For MLR. (c) For PR. (d) Error percentage.

TABLE 3: Comparison of ANN and regression models.

	R	RMSE	Bias	Nash number
ANN	0.97	109	-0.0003	0.98
MLR	0.95	279	0	0.90
PR	0.94	287	0.0074	0.89

TABLE 4: Expression of wind power models.

Ref.	Area/country of study	Equation of the wind power model	Performance
48	Brazil	Power output = -130.662 + 37.760 wind speed	RMSE = 101.29 kW
48	Brazil	Power output = -18.500 - 0.201 wind speed + 2.710 (wind speed) ²	RMSE = 100.64 kW
49	The islands of Crete and Kefalonia, Greece	Power output = 10.45 wind speed + 7.08 (wind speed) ² - 108.5	RMSE = 267.45 kW R ² = 0.37 (R = 0.61)
		$P_{out} = P_{rated} * CF = P_{rated} [0.087 V - P_{rated}/D^2]$ where	
50	Florida, USA	P_{out} = power output P_{rated} = rated power of the turbine (kW) CF = capacity factor V = yearly averaged wind speed (m/s) D = turbine diameter (m)	Accuracy = 2.6-3.5%
51	Goodnoe Hills, Washington, USA	$U(t + 1) = 1.1044U(t) - 0.2273U(t - 1)$ where $U(t + 1)$ = one hour ahead power forecast	Forecast probability < 0.885

the advantage that the wind power could be expressed in terms of the independent variables of wind speed and ambient temperature. Therefore, MLR and PR models can also be readily used to predict the wind power generation when the future weather conditions are known or forecast.

4. Conclusions

Wind power forecasting is essential to the future world in numerous ways. This paper presented a forecasting (prediction) study on wind power generation based on the relevant climatic factors. The developed models were successfully applied to Pawan Danavi, a wind power farm in Sri Lanka. The results revealed that the three models developed based on ANN, MLR, and PR techniques showcase similar and acceptable performances. ANN is lacking on producing a mathematical function, though the other two models have successfully expressed the wind power in the form of mathematical functions. MLR- and PR-based equations proposed for Pawan Danavi wind power farm for its wind power generation are based on wind speed and the ambient temperature of the wind farm area. Therefore, in the event the wind speeds and ambient temperatures are projected for future years, the wind power generation can be readily predicted. This is a significant improvement to the current literature as this knowledge was not available for the Sri Lankan wind farms prior to this study. These results and conclusions can be used by the sustainable energy authorities in Sri Lanka.

Data Availability

The data used for the research are available from the corresponding author upon request subject to approval of the relevant authorities.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors are thankful to Lanka Transformers Limited (LTL Holdings), Sri Lanka, for providing data.

References

- [1] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, "The role of renewable energy in the global energy transformation," *Energy Strategy Reviews*, vol. 24, pp. 38–50, 2019.
- [2] X. Zhao, S. Wang, and T. Li, "Review of evaluation criteria and main methods of wind power forecasting," *Energy Procedia*, vol. 12, pp. 761–769, 2011.
- [3] Z. Lin and X. Liu, "Assessment of wind turbine aero-hydro-servo-elastic modelling on the effects of mooring line tension via deep learning," *Energies*, vol. 13, no. 9, p. 2264, 2020.
- [4] Global Wind Energy Council, *Global Wind Report*, Global Wind Energy Council, Brussels, Belgium, 2019.
- [5] Jones, D. (2020). Global Electricity Review. 1–75, <https://ember-climate.org/project/global-power-2020/>.
- [6] Global Wind Energy Council, *GWEC: Over 60GW of Wind Energy Capacity Installed in 2019*, Global Wind Energy Council, Brussels, Belgium, 2020.
- [7] Asian Development Bank, *Sri Lanka: Energy Sector Assessment, Strategy, and Road Map*, Mandaluyong, Philippines, 2019.
- [8] I. Okumus and A. Dinler, "Current status of wind energy forecasting and a hybrid method for hourly predictions," *Energy Conversion and Management*, vol. 123, pp. 362–371, 2016.
- [9] H. Liu, H.-Q. Tian, C. Chen, and Y.-F. Li, "A hybrid statistical method to predict wind speed and wind power," *Renewable Energy*, vol. 35, no. 8, pp. 1857–1861, 2010.
- [10] Y. Wang, Y. Yu, S. Cao, X. Zhang, and S. Gao, "A review of applications of artificial intelligent algorithms in wind farms," *Artificial Intelligence Review*, vol. 53, no. 5, pp. 3447–3500, 2020.
- [11] P. Gomes and R. Castro, "Wind speed and wind power forecasting using statistical models: autoregressive moving average (ARMA) and artificial neural networks (ANN)," *International Journal of Sustainable Energy Development*, vol. 1, no. 1/2, 2012.
- [12] T. Ouyang, X. Zha, and L. Qin, "A combined multivariate model for wind power prediction," *Energy Conversion and Management*, vol. 144, pp. 361–373, 2017.
- [13] G. Sideratos and N. D. Hatzigiorgiou, "An advanced statistical method for wind power forecasting," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 258–265, 2007.
- [14] B. Bilal, M. Ndongo, K. H. Adjallah et al., "Wind turbine power output prediction model design based on artificial neural networks and climatic spatiotemporal data," in *Proceedings of the 2018 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1085–1092, IEEE, Lyon, France, 2018, February.
- [15] M. N. Jyothi and P. R. Rao, "Very-short term wind power forecasting through adaptive wavelet neural network," in *Proceedings of the 2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE)*, pp. 1–6, IEEE, Bengaluru, India, 2016, January.
- [16] Y. Zhao, L. Ye, Z. Li, X. Song, Y. Lang, and J. Su, "A novel bidirectional mechanism based on time series model for wind power forecasting," *Applied Energy*, vol. 177, pp. 793–803, 2016.
- [17] M. Carolin Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ANN: a case study," *Renewable Energy*, vol. 33, no. 5, pp. 986–992, 2008.
- [18] P. F. Correia and J. M. Ferreira de Jesus, "Simulation of correlated wind speed and power variates in wind parks," *Electric Power Systems Research*, vol. 80, no. 5, pp. 592–598, 2010.
- [19] M. J. Durán, D. Cros, and J. Riquelme, "Short-term wind power forecast based on ARX models," *Journal of Energy Engineering*, vol. 133, no. 3, pp. 172–180, 2007.
- [20] C. Gallego, P. Pinson, H. Madsen, A. Costa, and A. Cuerva, "Influence of local wind speed and direction on wind power dynamics-application to offshore very short-term forecasting," *Applied Energy*, vol. 88, no. 11, pp. 4087–4096, 2011.
- [21] B. Zhou, C. Liu, J. Li, B. Sun, and J. Yang, "A hybrid method for ultrashort-term wind power prediction considering meteorological features and seasonal information," *Mathematical Problems in Engineering*, vol. 2020, Article ID 1795486, 12 pages, 2020.

- [22] R. Rahmani, R. Yusof, M. Seyedmahmoudian, and S. Mekhilef, "Hybrid technique of ant colony and particle swarm optimization for short term wind energy forecasting," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 123, pp. 163–170, 2013.
- [23] P. Mandal, H. Zareipour, and W. D. Rosehart, "Forecasting aggregated wind power production of multiple wind farms using hybrid wavelet-PSO-NNs," *International Journal of Energy Research*, vol. 38, no. 13, pp. 1654–1666, 2014.
- [24] S. Liu, G. Li, H. Xie, and X. Wang, "Correlation characteristic analysis for wind speed in different geographical hierarchies," *Energies*, vol. 10, no. 2, p. 237, 2017.
- [25] S. Hanifi, X. Liu, Z. Lin, and S. Lotfian, "A critical review of wind power forecasting methods-past, present and future," *Energies*, vol. 13, no. 15, p. 3764, 2020.
- [26] M. Narayana, "Validation of Wind Resource Assessment Model (WRAM) map of Sri Lanka, using measured data, and evaluation of wind power generation potential in the country," *Energy for Sustainable Development*, vol. 12, no. 1, pp. 64–68, 2008.
- [27] A. G. Amarasinghe and E. N. C. Perera, "Modeling predictive suitability to determine potential areas for establishing wind power plants in Sri Lanka," *Modeling Earth Systems and Environment*, vol. 7, no. 1, pp. 443–454, 2021.
- [28] M. Narayana and S. Witharana, "Adaptive prediction of power fluctuations from a wind turbine at Kalpitiya area in Sri Lanka," in *Proceedings of the 2012 IEEE 6th International Conference on Information and Automation for Sustainability*, pp. 262–265, IEEE, Beijing, China, 2012, September.
- [29] A. T. Peiris, J. Jayasinghe, and U. Rathnayake, "Forecasting wind power generation using artificial neural Network: "Pawan danavi"—a case study from Sri Lanka," *Journal of Electrical and Computer Engineering*, vol. 2021, Article ID 5577547, 10 pages, 2021.
- [30] V. Singh, "Application of artificial neural networks for predicting generated wind power," *International Journal of Advanced Computer Science and Applications*, vol. 7, 2016.
- [31] D. Brezak, T. Bacek, D. Majetic, J. Kasac, and B. Novakovic, "A comparison of feed-forward and recurrent neural networks in time series forecasting," in *Proceedings of the 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFER)*, pp. 1–6, IEEE, New York City, NY, USA, 2012, March.
- [32] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.
- [33] B. Tomislav, D. Majetic, and D. Brezak, "GPU implementation of the feedforward neural network with modified Levenberg-Marquardt algorithm," in *Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 785–791, IEEE, Beijing, China, 2014, July.
- [34] G. K. Uyanik and N. Güler, "A study on multiple linear regression analysis," *Procedia-Social and Behavioral Sciences*, vol. 106, pp. 234–240, 2013.
- [35] M. Tranmer and M. Elliot, "Multiple linear regression," *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, vol. 5, no. 5, pp. 1–5, 2008.
- [36] R. G. Shepherd, "Regression analysis of river profiles," *The Journal Of Geology*, vol. 93, no. 3, pp. 377–384, 1985.
- [37] S. Jong, B. Wise, and N Ricker, "Canonical partial least squares and continuum power regression," *Journal Of Chemometrics*, vol. 15, no. 2, pp. 85–100, 2000.
- [38] V. Gowariker, V. Thapliyal, R. Sarker, G. Mandal, and D. Sikka, "Parametric and power regression models: new approach to long range forecasting of monsoon rainfall in India," *MAUSAM*, vol. 40, no. 2, pp. 115–122, 1989.
- [39] D.-S. Lee, C.-S. Chang, and H.-N. Chang, "Analyses of the clustering coefficient and the Pearson degree correlation coefficient of chung's duplication model," *IEEE Transactions on Network Science and Engineering*, vol. 3, no. 3, pp. 117–131, 2016.
- [40] J. H. Zar, "Spearman rank correlation," *Encyclopedia of Biostatistics*, vol. 7, 2005.
- [41] P. Zhang, C. Li, C. Peng, and J. Tian, "Ultra-short-term prediction of wind power based on error following forget gate-based long short-term memory," *Energies*, vol. 13, no. 20, pp. 5400–5413, 2020.
- [42] H. Valizadeh Haghi, M. Tavakoli Bina, M. A. Golkar, and S. M. Moghaddas-Tafreshi, "Using Copulas for analysis of large datasets in renewable distributed generation: PV and wind power integration in Iran," *Renewable Energy*, vol. 35, no. 9, pp. 1991–2000, 2010.
- [43] R. Girard and D. Allard, "Spatio-temporal propagation of wind power prediction errors," *Wind Energy*, vol. 16, no. 7, pp. 999–1012, 2012.
- [44] List of probability and statistics symbols. Math Vault. 2020-04-26. Retrieved 2020-09-12.
- [45] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.
- [46] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [47] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations," *Transactions of the ASABE*, vol. 50, no. 3, pp. 885–900, 2007.
- [48] J. M. Lima, A. K. Guetter, S. R. Freitas, J. Panetta, and J. G. Z. de Mattos, "A meteorological-statistic model for short-term wind power forecasting," *Journal of Control, Automation and Electrical Systems*, vol. 28, no. 5, pp. 679–691, 2017.
- [49] C. Stathopoulos, A. Kaperoni, G. Galanis, and G. Kallos, "Wind power prediction based on numerical and statistical models," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 112, pp. 25–38, 2013.
- [50] C. L. Archer and M. Z. Jacobson, "Evaluation of global wind power," *Journal of Geophysical Research: Atmospheres*, vol. 110, no. D12, 2005.
- [51] B. G. Brown, R. W. Katz, and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power," *Journal of Climate and Applied Meteorology*, vol. 23, no. 8, pp. 1184–1195, 1984.