Forecasting Wind Power Generation Using Artificial Neural Network: “Pawan Danawi”—A Case Study from Sri Lanka

Amila T. Peiris, Jeevani Jayasinghe, and Upaka Rathnayake

1Department of Electronics, Faculty of Applied Sciences, Wayamba University of Sri Lanka, Kauliyapitiya, Sri Lanka
2Department of Civil Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

Correspondence should be addressed to Upaka Rathnayake; upakasanjeewa@gmail.com

Received 4 February 2021; Revised 7 March 2021; Accepted 17 March 2021; Published 24 March 2021

1.Introduction

The world’s energy demand is ever-increasing, and related research presents numerous models to forecast future energy demand [1]. As climatic reasons and environmental conditions have forced the world to move to green energies, many countries have shown their interest in this topic by implementing new policies, rules, and laws in their communities and at least in some of the selected areas [2]. Some cities are named green cities where there is a minimum level of greenhouse gas emissions [3, 4]. Some countries have already proposed annual targets to achieve certain percentages of their energy demand by renewable energy sources. Therefore, renewable energy is a highlighted topic in the twenty-first century. Solar, wind, biomass, geothermal, hydropower, and ocean waves are some of the readily available renewable energy sources to date in the world. On average, 26.2% of the 2018 world’s energy demand was supplied by renewable energies, and it is forecasted to increase the percentage up to 45% by the year 2040 [5].

Wind energy is one of the best solutions for global warming because it is completely pollution-free and causes no greenhouse effects [6]. It has become a cost-effective approach due to the rise in fossil fuel prices [7]. The wind is the only natural source of energy available everywhere and, therefore, it is the most promising renewable energy source.
Its contribution to the world’s energy demand is going to be increased in the future with the ending of the fossil fuel era [8].

Sri Lanka, as a country, is also in discussions for the generation of renewable energies, and it is proposed to generate renewable energy to meet 100% of the country’s demand by 2050 [9]. However, according to the statistics of the Ceylon Electricity Board [10], as of now, the majority (67.02%) of total electricity generation in Sri Lanka is done using thermal oil and coal followed by hydropower (30.16%). Wind power contributes only for 2.55% of the power generation in Sri Lanka. Therefore, Sri Lanka spends thousands of million dollars to import crude oil, which is about 25% of the expenses for imports and equals 45% of income from exports [11]. On the other hand, nearly 5,000 km² of windy areas was identified in Sri Lanka with good-to-excellent wind resource potential as reported by the Sri Lanka Sustainable Energy Authority [12].

However, the wind speed is intermittent and volatile, which makes influences on the safe operation of the power grid showing random, unstable, and antipeaking characteristics [13]. It may seriously affect the quality of electric power and the operation of the power grid. If the output of wind power generation can be accurately forecasted, the negative influences that wind power brings to the grid can be reduced by a large extent [14]. The accuracy of wind power prediction is important for balancing the power generation as well [15]. Wind power forecasting is important in ensuring the reliability of power systems and in reducing the cost of the power system too. Further, accurate predictions are helpful to the government, policymakers, and other responsible authorities for taking necessary actions.

Identifying the importance of forecasting wind power generation, researchers have applied numerous statistical, data mining, and machine learning techniques for developing prediction models [16–21]. Artificial neural network (ANN) was found to be highly popular among researchers in the field of wind power prediction. ANN was applied to assess the wind energy output of seven wind farms in Tamil Nadu, India, using data collected over a period of 3 years [22]. Three input variables, namely, wind speed, relative humidity, and generation hours, and one output variable (i.e., the energy output of wind farms) were used for modeling in MATLAB resulting in the root mean square error (RMSE) of 0.0806 and the overall percentage error of 4%. Another research based on ANN was conducted using the input parameters of average wind speed, average relative humidity, and generation hours of wind farms in Rajasthan, India [23]. The backpropagation algorithm was applied, and the mean squared error (MSE) becomes stable at 0.0070 after 300 iterations. Another ANN-based research with backpropagation was applied to develop a forecasting system for power generation in some wind fields in China [13]. According to a case study conducted in Tasmania, Australia, the prediction behavior of the ANN model is more accurate than the Similar Days approach when the performance is compared based on the daily mean absolute percentage error (MAPE) [24]. Research conducted in Prince Edward Island, Canada, concludes that ANN outperforms Fuzzy Predictor, Adaptive Neural Fuzzy Inference System, and Committee Machines in wind power prediction [25].

The prediction of wind power generation is possible for future years if the independent variables (climatic data) are available as projected climatic variables. Related research studies are extensively carried out [26–30], and projecting future climate under different scenarios is often reported. Representative Concentration Pathway (RCP) is one such climate projection scenario [31–36] widely adopted by the Intergovernmental Panel on Climate Change (IPCC). RCP2.6, RCP4.5, RCP6.0, and RCP8.5 are four scenarios based on the greenhouse gas concentration, and the projection of various weather factors is readily available.

Forecasting the energy output of wind farms by means of short-term, medium-term, or long-term prediction is reported. Hossain et al. presented extrapolation of wind speed and apply data on adaptive neurofuzzy inference systems to develop monthly and weekly wind power density prediction models in Malaysia [37]. Metaheuristic techniques such as ant colony optimization and particle swarm optimization were also incorporated to develop forecasting models [38]. In Iran, the empirical hourly wind power output of a wind farm over a year and data of wind speed and ambient temperature were collected, and a prediction model was introduced [38]. The meteorological data consisting of wind speed and ambient temperature is used as the inputs to the mathematical model. Both the statistical and the neural network-based approaches were applied to predict hourly wind speed data of the subsequent year for long-term prediction [39].

However, to the authors’ knowledge, only a few research studies have been carried out to forecast wind power in the context of Sri Lanka. Nevertheless, neural networks were used in the prediction of various energy-related research [40]. Therefore, it is highly important to carry out such a research to forecast the wind power of an already established wind farm in Sri Lanka. This will lead to understanding the importance in the context of Sri Lanka, while the results can be used for identifying the potential future wind power generations in the country. Fulfilling this research gap, the presented paper is focused on forecasting wind power generation in a wind farm in Sri Lanka: “Pawan Danawi”.

2. Methodology

2.1. Study Area and Data Collection. This research work is done based on the “Pawan Danawi” wind farm, which is located in Kalpitiya (around 08°02’56”N 79°43’08”E), North-Western area in Sri Lanka (Figure 1). This geographical area was identified as one of the best locations in Sri Lanka to establish a wind farm [41, 42]. It is an onshore wind farm, which has a 10.2 MW nameplate capacity with 12 wind turbines (model: Gamesa G90-850) for its operation. The height of each tower is around 65 m, and the diameter of the blades is around 58 m. The project was completed and connected to the national grid for its operation in August 2012.

Each wind turbine has three blades, and its rated power 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 national power grid as the transmission voltage in Sri Lanka is 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 the low voltage side before transformer input terminals is 0.95 at partial loads and 1 at nominal power.

Monthly average power generation data (MW) from January 2015 to December 2019 (5 years, 60 data sets) were obtained from the wind farm authorities. Figure 2 exhibits the variation of generated power over the 5 years. The maximum power generation is around 3.0 MW. The peak power generation can be seen in the months of June, July, and August, where most of the regions in Sri Lanka receive speedy winds. The current output was also calculated assuming a unity power factor and its variation over the year 2019 is illustrated in Figure 3.

Figure 4 shows the variation of average wind speed over the months for the previous 5 years. The variation over the months in each year looks similar, and speedy winds could be observed in the months of June, July, and August.

As the peaks of both wind power generation and wind speed are visible in the same months in each year, their relationships were plotted in Figure 5 for further analysis. It indicates that power generation is exponentially correlated with wind speed. Further, most of the time, the wind speed varied between the cut-in speed and the cut-out speed of 3.0 m/s and 20.0 m/s, respectively. The coefficient of determination obtained for the trend line of Figure 5 clearly shows that the predictive power exponentially varies with the wind power and wind speeds. However, the power generated by a wind turbine depends not only on the wind speed but also on the density of air [43]. Therefore, climate parameters of average wind speed (m/s), average wind direction (°), and average ambient temperature (°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.

2.2. ANN Algorithms. A feedforward artificial neural network model was developed to predict the power generation
SCG algorithm is a combination of the conjugate gradient approach and the model trust region approach. As it has avoided the line search technique by using the trust region approach, it is a much faster technique [49]. The SCG algorithm is widely used with problems that have a higher number of linear equations. It takes fewer memory and the training process stops when generalization deescalates, indicated by the increasing the mean squared error (MSE) [50].

In the BR algorithm, a conventional sum of the least square error function is combined with regularization. Therefore, it forces the neural network to converge to a set of weights and acquire minimum values for the biases, pushing the network to be smooth [51]. BR algorithm is based on the probability distribution concept. Hirschen and Schäfer proposed that “test set” or “validation set” is not necessary for BR, and the complete data set can use used for model fitting and model comparison [52]. In general, this algorithm takes a little more time to train, but this is ideal for relatively smaller and difficult data sets [26].

2.3. Development of Prediction Models. The ANN models were developed based on three major climatic factors in the area, which are directly affecting the power generation. The governing equation for the model development is given in

\[
\text{Power} = F(\text{wind speed, wind direction, ambient temperature}).
\]  

MATLAB version R2014b was used to develop the ANN framework for this study. The model has three independent variable matrices as shown in equation (3) with one dependent variable matrix. To complete the nonlinear formulation and architecture of the ANN, a single hidden layer was imposed on the modeling process. Training of the ANN was performed on 70% of the data set while a calibration process was carried out to reach a better validation data percentage. The calibration process was conducted in intervals of 5% from 5% to 25%. Therefore, the testing percentage of the process was automatically changed. In addition, a calibration was carried out in searching for a better training algorithm. Levenberg-Marquardt (LM), Scaled Conjugate Gradient (SCG), and Bayesian Regularization (BR) training algorithms were independently used in the training process. The process was conducted for 30 different runs for each case. In other words, 150 (= 5 × 30) separate runs were performed for one training algorithm resulting in 450 runs in total.

The performance of the training algorithms is evaluated by widely used techniques: Mean Squared Error (MSE), Coefficient of Correlation (R), and BIAS. They were determined for all three training algorithms with all five validation percentages. Root mean squared error ratio (RSR) and Nash number were also used to evaluate the performance of the ANN models. The calculation of MSE, R, RSR, Nash number, and BIAS is given in equations (4)–(8), respectively.
\[
\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (O_i - E_i)^2,
\]

where \(O_i\) is the observed power generation, \(E_i\) is the expected or predicted power generation, and \(N\) is the number of data points in the matrix:

\[
R = \frac{\sum_{i=1}^{N} (O_i - \bar{O})(E_i - \bar{E})}{\sqrt{\sum_{i=1}^{N} (O_i - \bar{O})^2 \cdot \sum (E_i - \bar{E})^2}},
\]

where \(\bar{O}\) and \(\bar{E}\) are the means of observed and predicted power generations, respectively. The smaller MSE values and \(R\) values closer to 1 indicate highly accurate models.

\[
\text{RSR} = \frac{\sqrt{\text{MSE}}}{\sigma},
\]

where \(\sigma\) is the standard deviation of the observed power values.

\[
\text{Nash number} = 1 - \frac{\sum_{i=1}^{N} (O_i - E_i)^2}{\sum_{i=1}^{N} (O_i - \bar{O})^2},
\]

\[
\text{BIAS} = \frac{\sum_{i=1}^{N} (E_i - O_i)}{N}.
\]

\section{3. Results and Discussion}

The results of the comparison of validation percentages (calibration of different training algorithms at various validation percentages) show that a validation percentage of 5\% gives the lowest MSE at epoch 2 for the LM algorithm. At the same time, the corresponding overall \(R\) value is 0.94, which is almost equal to the results of the other four percentage values (10\%, 15\%, 20\%, and 25\%). Figure 6 shows the variation of the predicted versus actual power for the LM algorithm at the validation percentage of 5\%. Overall, the LM training algorithm has successfully predicted the wind power generation according to equation (3).

Calibration of SCG training algorithm shows a better performance at validation percentage of 10\% with the minimum MSE of 0.017 at epoch 21. Figure 7 shows the variation of the predicted versus actual power of the corresponding model. According to the results of SCG, there is a strong correlation between input parameters and power output at all five validation percentages.

Unlike the LM and SCG training algorithms, BR training algorithms show the best performance at 15\% of validation percentage at epoch 922. However, the minimum MSE values of the BR algorithm are smaller than those of the other two algorithms and obtained at a much higher epoch. In addition, there is a strong correlation between input parameters and power output at all five validation percentages in the BR algorithm.

The performance of all models was analyzed in terms of \(R\), MSE, and BIAS. The results are summarized in Table 1. According to the results, the coefficient of correlation is higher than 0.94, MSE values are less than 0.17, and BIAS values are less than 1, which clearly demonstrates the accuracy of the models. Furthermore, MSE values obtained in the BR-based models are smaller compared to the other two algorithms; however, they are at higher computational costs.

When considering the computational efficiency of the three algorithms, LM and SCG produce minimum MSE at relatively lesser epochs. That means LM and SCG have reached the optimum result in less time. Among them, LM has the least number of epochs. That means the performance of LM is much better. However, in the case of BR, the number of epochs that had to be run to reach the optimum result is very high. It shows that BR is more time-consuming than the other two algorithms to reach the optimum.

Variation of the predicted and actual power levels for the LM algorithm with a validation percentage of 15\% during the past 5 years is illustrated in Figure 8. According to the results, the error is negligible in most of the months and numerically less than 0.9 MW demonstrating the accuracy of the LM-based ANN model. As per the calculations, it exhibits an RSR of 0.524 and a Nash number of −0.723.

Figure 8 is an interesting figure as it clearly shows the capabilities of the ANN model developed based on equation (3). In other words, the predicted power output is based on the wind speed, wind direction, and the ambient temperature of the area of interest. The prediction of wind power generation is possible for future years if the independent variables (climatic data) are available as projected climatic variables.

Therefore, the model development presented in this research is useful in predicting the future wind power generation by the Pawan Danawi wind farm. Such a model is not only useful for a particular wind farm but also highly valuable for the country itself in the path of developing renewable energy. Therefore, the developed prediction model can be considered as a starting point, and then, it can be further expanded and used across the country for forecasting the wind energy generations in future years. In addition, the model can be easily adopted by the regions and areas which have similar environmental and climatic conditions and develop the feasibility activities for future wind farms. This is highly essential in the development plans of a particular country. Similar studies conducted by other researchers based on data collected from wind farms located in other countries are summarized in Table 2. ANN and some other machine learning techniques such as support vector machine (SVM), neural network (NN), random forest (RF), and k-nearest neighbor (k-NN) and fuzzy logic techniques were used for developing wind power prediction models. Among them, ANN reported suitable to forecast the wind power generation accurately outperforming other techniques [13, 22–25]. Statistical parameters such as mean absolute percentage error (MAPE), root mean square error (RMSE), mean absolute error (MAE), and correlation coefficient were used to assess the performance of the prediction models. In contrast, the accuracy of the proposed model was proven by using five statistical parameters.
Figure 6: The predicted versus actual power for the LM algorithm at the validation percentage of 5%: (a) training, (b) validation, (c) test, and (d) overall.

Figure 7: Continued.
Table 1: Performance of the algorithms for different validation percentages.

<table>
<thead>
<tr>
<th>Validation percentage</th>
<th>ANN training algorithm</th>
<th>Number of epochs</th>
<th>R</th>
<th>MSE</th>
<th>BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LM</td>
<td>2</td>
<td>0.94</td>
<td>0.024</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>SCG</td>
<td>14</td>
<td>0.95</td>
<td>0.021</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td>BR</td>
<td>603</td>
<td>0.97</td>
<td>0.000</td>
<td>0.023</td>
</tr>
<tr>
<td>10</td>
<td>LM</td>
<td>4</td>
<td>0.91</td>
<td>0.114</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>SCG</td>
<td>21</td>
<td>0.96</td>
<td>0.017</td>
<td>-0.182</td>
</tr>
<tr>
<td></td>
<td>BR</td>
<td>314</td>
<td>0.94</td>
<td>0.006</td>
<td>-0.166</td>
</tr>
<tr>
<td>15</td>
<td>LM</td>
<td>2</td>
<td>0.97</td>
<td>0.219</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>SCG</td>
<td>19</td>
<td>0.95</td>
<td>0.163</td>
<td>-0.995</td>
</tr>
<tr>
<td></td>
<td>BR</td>
<td>992</td>
<td>0.98</td>
<td>0.008</td>
<td>-0.281</td>
</tr>
<tr>
<td>20</td>
<td>LM</td>
<td>2</td>
<td>0.97</td>
<td>0.127</td>
<td>0.536</td>
</tr>
<tr>
<td></td>
<td>SCG</td>
<td>17</td>
<td>0.95</td>
<td>0.160</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>BR</td>
<td>408</td>
<td>0.99</td>
<td>0.000</td>
<td>-0.276</td>
</tr>
<tr>
<td>25</td>
<td>LM</td>
<td>2</td>
<td>0.97</td>
<td>0.147</td>
<td>0.522</td>
</tr>
<tr>
<td></td>
<td>SCG</td>
<td>17</td>
<td>0.94</td>
<td>0.168</td>
<td>0.708</td>
</tr>
<tr>
<td></td>
<td>BR</td>
<td>413</td>
<td>1.00</td>
<td>0.000</td>
<td>-0.367</td>
</tr>
</tbody>
</table>

Figure 7: The predicted versus actual power for the SCG algorithm at the validation percentage of 10%: (a) training, (b) validation, (c) test, and (d) overall.

Figure 8: The error between predicted and actual power levels for LM algorithm with a validation percentage of 15%.
4. Conclusions

Forecasting models based on the artificial neural network were developed to predict the wind power generation of Pawan Danawi wind farm in Sri Lanka. The results showcased the capabilities and robustness of the prediction models developed under various training algorithms in which the Levenberg-Marquardt (LM) algorithm with a validation percentage of 15% produced the best results. Therefore, the LM-based ANN model is proposed to predict the potential wind power generation with the help of projected climatic scenarios in future years. Accordingly, the results of the research open a new era for the wind power generations in Sri Lanka to plan its energy demand and supply in the future. Sri Lanka, as a developing country, can effectively use this development to achieve its sustainable goals in renewable energy generation at a low cost. In addition, the results can effectively be used in similar climatic areas in the country to showcase the feasibility of wind power potential and its economic worth. However, research expansions are proposed with more data probably from the Kalpitiya area as the establishment of few more wind farms is in progress.

Data Availability

Prior request from Lanka Transformers Limited is needed for the availability of power generation data and climatic data.

Table 2: Comparison of similar research studies.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Techniques used</th>
<th>Location</th>
<th>Independent variables</th>
<th>Performance evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>[13]</td>
<td>ANN with BP neural network</td>
<td>China</td>
<td>Wind velocity and wind direction</td>
<td>Error</td>
</tr>
<tr>
<td>[14]</td>
<td>ANN with BP neural network</td>
<td>—</td>
<td>Wind speed, maximum, minimum, and average temperature, average humidity, air pressure, and rainfall</td>
<td>MAPE and RMSE</td>
</tr>
<tr>
<td>[16]</td>
<td>SVM, NN, RF and k-NN</td>
<td>Gansu province, China</td>
<td>Temperature, wind speed, wind direction, cloud cover, pressure, heat flux, radiation, precipitation, humidity, etc.</td>
<td>MAE, MAPE, RMSE</td>
</tr>
<tr>
<td>[17]</td>
<td>Wavelet method and the improved time series method</td>
<td>China</td>
<td>Wind speed</td>
<td>MAE, MAPE, MSE</td>
</tr>
<tr>
<td>[18]</td>
<td>Multilayer perceptron network and integrated k-NN</td>
<td>—</td>
<td>Wind speed, wind direction, air density, temperature difference, sensible heat flux, and vegetation</td>
<td>MAE and standard deviation of absolute error</td>
</tr>
<tr>
<td>[19]</td>
<td>ANN</td>
<td>Portugal</td>
<td>Wind speed</td>
<td>MAE, RMSE and mean relative error</td>
</tr>
<tr>
<td>[20]</td>
<td>Neural networks and fuzzy logic techniques</td>
<td>Aalborg, Denmark</td>
<td>Wind speed, wind direction, temperature</td>
<td>Normalize mean absolute error, normalized root mean square error</td>
</tr>
<tr>
<td>[21]</td>
<td>ANN</td>
<td>Tamil Nadu, India</td>
<td>Wind speed, relative humidity, and generation hours</td>
<td>RMSE and MAE</td>
</tr>
<tr>
<td>[22]</td>
<td>ANN with BP neural network</td>
<td>Rajasthan, India</td>
<td>Generation hours, relative humidity, and wind speed</td>
<td>MSE and MAE</td>
</tr>
<tr>
<td>[25]</td>
<td>ANN and radial basis function neural networks</td>
<td>Canada</td>
<td>Temperature, dew point temperature, relative humidity, wind direction, wind speed, pressure</td>
<td>MSE, absolute mean error, MAPE, and correlation coefficient</td>
</tr>
</tbody>
</table>

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

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

References


