Research Article

Hourly and Day Ahead Power Prediction of Building Integrated Semitransparent Photovoltaic System

S. Kaliappan,1 R. Saravanakumar,2 Alagar Karthick,3 P. Marish Kumar,4 V. Venkatesh,5 V. Mohanavel,6 and S. Rajkumar7

1Department of Electrical and Electronics Engineering, Kumaraguru College of Technology, Coimbatore, 641049 Tamilnadu, India
2Department of Wireless Communication, Institute of ECE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, 602105 Chennai, India
3Renewable Energy Lab, Department of Electrical and Electronics Engineering, KPR Institute of Engineering and Technology, Coimbatore, 641407 Tamilnadu, India
4Department of Electrical and Electronics Engineering, Easwari Engineering College, Chennai, 600089 Tamilnadu, India
5Department of Electrical and Electronics Engineering, Rajalakshmi Engineering College, Thandalam, Chennai, 602105 Tamilnadu, India
6Centre for Materials Engineering and Regenerative Medicine, Bharath Institute of Higher Education and Research, Chennai, 600073 Tamilnadu, India
7Department of Mechanical Engineering, Faculty of Manufacturing, Institute of Technology, Hawassa University, Ethiopia

Correspondence should be addressed to Alagar Karthick; karthick.power@gmail.com and S. Rajkumar; rajkumar@hu.edu.et

Received 11 July 2021; Revised 1 November 2021; Accepted 8 December 2021; Published 26 December 2021

Academic Editor: Francesco Riganti-Fulginei

Copyright © 2021 S. Kaliappan 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.

The building integrated semitransparent photovoltaic (BISTPV) system is an emerging technology which replaces the conventional building material envelopes and roof. The performance prediction of the BISTPV system places a vital role in the reduction of the energy consumption in the building. In this work, the artificial neural network (ANN) is used to predict the performance of this system by optimizing the important parameter of the feature selection. The Elman neural network (EN), feed forward neural network (FN), and generalized regression neural network model (GRN) are investigated in this study. The performance metrics of the errors are analysed such as the root mean square error (RMSE), mean absolute percentage error (MAPE), and mean square root (MSE). According to the findings, the model behaves consistently at the specified time and place in the experiment. Forecasters utilizing neural network models will have better accuracy if they use techniques like EN, FN, and GRN having the RMSE of 0.25, 0.37, and 0.45, respectively.

1. Introduction

The performance prediction of the solar photovoltaic system depends on the geographical location, meteorological condition, and also quality of data [1]. It plays a vital role in the best forecasting technique. Subhourly, it forecasts benefit from the use of satellite imagery. On the other hand, satellite data may predict PV output on a climatological time scale or provide projections up to six hours in advance [2]. When converting PV output, there are various options available, ranging from deterministic models with three or five parameters to parametric models or other machine learning techniques [3]. The prediction of the performance of the solar photovoltaic system plays a vital role in the energy sector. The need of the power demand may be compensated with this type of the power prediction need and its requirement. The artificial intelligence-based power prediction is an emerging tool to predict the system performance. The building integrated semitransparent solar photovoltaic system can be incorporated into the building structure by replacing the conventional building elements. The baseline model used here is a parametric PV output conversion model and often
used in PV output prediction for day-ahead time horizons using Numeric Weather Prediction (NWP) forecasts [4]. The deterministic and probabilistic intrahour predictions of solar irradiance are assessed using machine learning approaches (k-nearest neighbours and gradient boosting regression trees (GBRT)) on the other [5]. Regression trees are used in solar forecasting [6]. As inputs, the model employs expected climatic factors from a (NWP) model and real power measurements from photovoltaic (PV) plants [7] to forecast power production. Short-term forecasting of PV power production is done with the help of a Quantile Regression Neural Network (QRNN) [8]. For intrahour horizons, a real-time hybrid probabilistic model is created [9]. To summarise, it is critical to enhance weather variable forecasts by choosing the most informative predictors and rejecting the uninformative ones via effective model selection. The best use of available input data is required to achieve this. This has been done effectively using ensemble techniques in probabilistic energy forecasting [10]. The PV system’s solar cell directly transforms the sun’s energy into an unstable electric current.

To put it another way, the amount of electric power that can be generated is dependent on the changing weather conditions. In the open, you have the most energy. A daytime sky that is shady from nearby trees, buildings, and birds. The authors have studied the degradation of electrical power caused by the PV module’s hot spot in [11]. And [12] look into temperature’s effect on current-voltage, while [13] shows how the PV system’s inverter is unstable. These hiccups can affect the grid and alter its state of instability. To enhance and ensure that PV electricity is effectively distributed in the grid, power forecasts are crucial for managing the system’s utility and ensuring that reserve capacity is utilized [14]. Several papers were presented and developed for forecasting PV power in this context. Available data and their specific forecast horizon influence the forecast technique chosen [15]. Sunny, cloudy, and rainy days are possible, so the forecast is flexible [16]. With this method, only the PV system’s major design characteristics are taken into account for calculating the system’s actual power output. This is an excellent resource for PV plant owners who already have their facility’s design documents filled with all the information they need. Physical modelling surpasses machine learning techniques in both accuracy and efficiency [17, 18]. In general, hybrid models outperform either solely physical or statistical approaches, although some researchers suggest that incorporating physically determined features [15] or even simple clear sky irradiance [19] enhances performance. Two critical applications rely heavily on physical PV power forecasting models: (1) the prediction of power for new PV installations when the historical PV generation data is not present and (2) the best PV system prediction model using a combination of physical and data-driven modelling [20–22]. Using a model chain with numerous computation stages is critical to forecasting PV power production based on projected irradiance data accurately. Solar energy conversion may be broken down into three steps: beam separation, diffuse irradiance translation, and PV performance modelling, as described [23]. For the day-ahead, hourly regional forecast of German PV production based on the predictions of the global model used a numerical model with four design phases [24]. There are four steps of computation in the physical model (ECMWF). Each of the four processes (transposition of irradiance, temperature of cells, performance of solar panels, and kind of inverter) has its [20] section-contrasted support vector regression (SVR) with numerical modelling for cloud motion vector (CMV) and near-surface irradiance projections using satellite-derived CMV and near-surface irradiance estimates. The physical PV simulation contains all of these. Based on the irradiance data from the NWP, the researchers found that physical models were utilizing simple linear regression beat SVR in terms of accuracy. By combining historical data on PV output with projections of future solar radiation, Saint-Drenan and colleagues [25] established a method for computing the fundamental parameters of photovoltaic (PV) systems.

From the literature review, it is found that the performance prediction of the solar photovoltaic system yields maximum performance of the system. There are few researches found on the prediction of the building integrated semitransparent photovoltaic system with respect to the day and the hourly prediction of the system. In this work, the artificial neural network tool is used to optimize and predict the performance the building integrated semitransparent photovoltaic system. The Elman neural network (EN), feedforward neural network (FFN), and generalized regression neural network model (GRN). Finally, the system performance metrics is analysed with respect to the error analysis of the system such as root mean square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE), and the correlation coefficient presented in the subsequent section.

2. Materials and Method

Geographical location and climatic conditions affect BISTPV’s efficiency. According to the findings of this research, the grid-connected BISTPV system performs well in the hot and humid region of Kovilpatti, Tamilnadu. The climate is hot and humid in the southern section of India. For three years, the system was observed, and its output data was given. Figure 1 depicts the BISTPV generating set-up. The gathered data is put to good use in estimating the BISTPV system’s performance.

The geographical coordinate site latitude and longitude are $9110^\circ 0^\prime 0^\prime\prime \text{N}$ and $7752^\circ 0^\prime 0^\prime\prime \text{E}$. The experimental dataset is split into the three-dataset trained dataset, validated dataset, and testing dataset using the optimization tool of the MATLAB 2021 version. The artificial neural network machine learning toolbox is used to predict and optimize the performance of the system. The methodology is presented in Figure 2. As the quality of the data utilized to make the predictions increases, so does forecasting accuracy. Researchers have used previously reported dataset for PV power production to replicate the particular system performance and their geographical locations. However, data sets are sometimes disrupted by sudden fluctuations or static due to intermittent
meteorological conditions, power system oscillations, and outages, to name a few causes. Events that contradict patterns and are the result of random occurrences are known as statistical outliers. They have a substantial impact on forecasting and decision-making. Sensor failures or incorrect recordings, which might occur on occasion, could also corrupt or destroy data. Before further processing the damaged input data, it is essential to use decomposition, interpolation, and seasonal adjustments to recreate the distorted input data. Preprocessing procedures that are detailed and exact should be followed. In the middle of the day, data may be lost due to a failure of the solar radiation and temperature sensors connectivity and network of the location. To avoid bias in future studies, it is best to disregard them entirely.

3. Machine Learning Algorithm

3.1. Elman Neural Network Algorithm (EN). Jeffrey L. EN proposed the Elman neural network in 1990. It is a potent neural network with a lot of feedback. The input, hidden, context, and output all exist as separate layers in EN’s neural network. The input layer’s job is to transmit signals. Only linear weighting is used in the output layer. There is an extra layer called the context layer in EN that varies from regular BPNN. Detection of the current time frame’s output is done using input signals from the earlier reported frequency. Due to the context layer neuron’s output being stored in the hidden layer, the hidden layer receives it before any other layer. Figure 3 shows the study architecture, which incorporates the EN learning algorithm and the BP learning algorithm.

The following is a flowchart created by EN. The sensitivity to the source datasets is responsible for the increase in model processing capacity that arises from using these datasets. Prediction is performed using NN EN, a better variant of BPNN. Simulations use data gleaned from PV power plants, historical power databases, and multivariate meteorological factors. A data gap may be caused by power plant maintenance or failure. In this case, the next step is to remove any abnormal or missing data. Historical datasets must be standardized for the following reasons:

4. Feedforward Neural Network (FFN)

A single-layer perceptron, a FFN in its simplest form, is a familiar sight. The inputs into the layer are multiplied by the weights in this model. The weighted input values are then joined together to get a final result. A value of 1 is typically created if the total of the values exceeds a threshold, often set at zero; a value of -1 generally is produced if the sum falls below the threshold. The single-layer perceptron is an essential FFN model often employed in classification problems. Machine learning may also be included in a single-layer perceptron. The NN may compare its nodes’ outputs with the intended values using a trait known as the delta rule, enabling the network to fine tune its weights over time to provide more accurate output values. Gradient descent is created throughout the training and learning process. The technique of updating weights in multilayered perceptron is roughly the same, although the process is more
precisely characterized as backpropagation. Such networks have all hidden layers that are modified following the final layer’s output values. Although Feed Forward Neural Networks have a simple design, this might benefit some machine learning applications because of the reduced complexity. When using feed-forward neural networks, it is possible to set them up to operate independently of one another, but with a minor intermediate to help with the moderating process. More extensive tasks are handled and processed by this mechanism using numerous individual neurons, much as in the human brain. To provide a composite and coherent output, the findings from the various networks may be integrated after each job. Figure 4 shows the schematic of feedforward neural network.

5. Generalized Regression Neural Network Model (GRN)

Donald Specht introduced a probabilistic neural network, the extended regression neural network. It is possible to approximate any probability distribution function using GRNN’s neural network design. GRNN can solve any function approximation and estimate any continuous variable issue. Because of its parallel nature, this method only has to be run once. GRNN is shown schematically in Figure 5. There are four levels to the GRNN, as depicted in Figure 5: input, pattern, summation, and output. The two neurons of the summation layer, the S-summation neuron and the D-summation neuron, are coupled with the weights from the pattern layer’s summation layer. The S-summation neuron adds up the pattern layer’s weighted outputs, while the D-summation neuron does so unwrapped. An unknown input vector \( x \) is anticipated by dividing the outputs of each S-summation neuron by the outputs of each D-summation neuron in the output layer.

6. Energy Performance Metrics

System performance is assessed using the root mean square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE), and the correlation coefficient. The root mean square error (RMSE) is defined as the residual value’s standard deviation (prediction errors). The residuals quantify the distance between the regression line and the data points, and the RMSE evaluates how the residuals are
The following equations are taken from [29–32].

\[ RMSE = \frac{1}{m} \sum_{i=1}^{m} \left( \frac{Z_p}{C_0/C_1} i - (Z_m)_i \right)^2, \]  
\[ MSE = \frac{1}{m} \sum_{i=1}^{m} \left( \frac{Z_p}{C_0/C_1} i - \left(\frac{Z_m}{(Z_m)_i}\right)^2 \right)^2, \]  
\[ R = \frac{n(\sum zX) - (\sum z)(\sum X)}{\sqrt{[n\sum z^2 - (\sum z)^2] \left[n\sum X^2 - (\sum X)^2\right]}}. \]  
\[ MAPE = \frac{1}{k} \sum_{i=1}^{k} \left| \frac{(Z_m)_i - (Z_p)_i}{(Z_m)_i} \right|, \]  
\[ MAE = \frac{1}{m} \sum_{i=1}^{m} \left| \frac{(Z_p)_i - (Z_m)_i}{(Z_p)_i} \right|. \]  

6.1. Input Dataset. The input dataset adopted for the study is obtained from the work of [29–32]. The obtained dataset are split into the train data and the test data which is plotted in Figures 6–8 where the solar radiation wind speed and the ambient temperatures are presented. The obtained datasets are then feed in the MATLAB 2020b version optimisation tool to obtain the best results. There are a number of features considered for the study which is considered and optimized using the correlation plot heat map which is plotted in Figure 9.

The Pearson plot of the chosen characteristics for the prediction of the system can be seen in Figure 6. When it comes to prediction processing, the error ranges for the other two models are [-0.8] and [0.8], but the error goes for the EN-FFN-GRN models shift dramatically. The projected error for the GRN model is about 30% higher. The EN-FFN-GRN model’s mistakes fall within this range shown in Figure 9. Artificial neural network EN-FFN-GRN and a generalized regression neural network model are the machine learning techniques under consideration. Detailed model information may be found in the section on machine learning algorithms. Using the supplied experimental data, this will be achieved. PV panels’ power output must be adjusted to take account of site-specific environmental variables. The numbers are delivered to the audience every five minutes. Figures 6–8 show the climatic circumstances and the components that contributed to incoming solar radiation, which are the study’s input parameters taken into account.

7. Result and Discussion

The effectiveness of solar energy applications is location-dependent. In the daylight, we have access to solar radiation, which comes in bursts. The ability to anticipate the solar photovoltaic system’s output power is critical when building a large-scale PV power plant. Algorithms for machine learning assist in the forecast and assessing the system’s performance. The hot and humid climatic conditions were used
to generate the experimental dataset [29–32]. Figures 6–8 depicts the tested feature selection parameters such as insolation, ambient temperature, and wind speed to forecast STPV output power. The Pearson plot of the chosen characteristics for the BISTPV system power generation prediction can be seen in Figures 10(a) and 10(b). When it comes to prediction processing, the error ranges for the other two models are [-0.8] and [0.8], but the error runs for the EN-FFN-GRN models shift dramatically. The projected error for the GRN model is about 30% higher. The EN-FFN-GRN model’s mistakes fall within this range. We are looking at artificial neural network EN-FFN-GRN model for machine learning techniques. Detailed model information may be found in the section on machine learning algorithms, using the supplied data.

The experimental data will be achieved. PV panels’ power output must be adjusted to take account of site-specific environmental variables. The numbers are delivered to the audience every five minutes. Figure 6 shows the ambient circumstances, which are the study’s input parameters.
taken into account while calculating incoming solar radiation. Figure 7 shows how the solar photovoltaic system’s output power for the hourly variation changes using the ML algorithm of EN-FFN-GRN models shown in Figures 10(a) and 10(b). Figure 10 shows the hourly variation of the BISTPV system prediction curves for the output of sunlight power from the EN-FFN-GRN models on a sunny day. Furthermore, the ANN model’s estimated early-observation error is 32%. The EN-FFN-GRN model has prediction errors in the range of 2, which predicts -1% and 1%. A typical day’s worth of data is used to evaluate the EN-FFN-GRN model. Figure 8 shows the predicted daily PV panel output power variation for EN-FFN-GRN ML models on a sunny day. In addition, the ANN model’s estimated early-observation error is 30%. Of total, the EN-FFN-GRN model’s prediction errors are within a two-standard deviation of the actual value. In 94% and 98% for the tested samples, the EN-FFN-GRN model has relative error values of -1.5% and 1.5%. The EN-FFN-GRN model is evaluated using weather data from typical summer days. The model is viewed as being tested on the training data, while the model is considered to be being trained on the testing data. Figure 9 shows the weekly variation in the output power of the solar photovoltaic module in a similar pattern (Figures 11(a) and 11(b)). Prediction models are graded using MAPE and RMSE (RMSE). The MAPE is shown in Figure 12; in the EN-FFN-GRN models, the average root mean square error (RMSE) is 0.285 in dry, 0.301 in partly wet, and 0.426 in wet conditions. The EN-FFN-GRN model has the best predictive stability in the MSE setting because it is the smallest. For PV facade and roof installations, the model’s output varies widely, and the inaccuracy is much more significant. The MAPE typically falls between a value of 0.189 and a value of 0.241. The EN-FFN-GRN model has a lower MAPE value.

8. Conclusion

The predictions regarding the solar photovoltaic system’s performance are based on the experimental data. A repeatable model was created since it only uses environmental data and ignores the user’s physical location. Data sets for training, validating, and testing machine learning models are chosen depending on the nature of the available dataset. According to the findings, the model behaves consistently at the specified time and place in the experiment. Forecasters utilizing neural network models will have better accuracy if they use techniques like EN, FFN, and GRN. Finally, the final model made reliable predictions with an RMSE of 0.25 in EN and FFN and 0.42 in GRN. Since the model does not rely on its simulation location or configuration, these training method properties show that the model is independent of the simulation context. Power grid management systems are projected to function and operate more reliably thanks to greater computerization. As a result of the new rules, renewable energy producers and aggregators will have a more significant say in the electricity market.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] A. Mellit, A. Massi Pavan, E. Ogliari, S. Leva, and V. Lughin, “Advanced methods for photovoltaic output power forecasting: a review,” Applied Sciences, vol. 10, no. 2, p. 487, 2020.
[2] V. Kallio-Myers, A. Riihelä, P. Lahtinen, and A. Lindfors, “Global horizontal irradiance forecast for Finland based on geostationary weather satellite data,” Solar Energy, vol. 198, pp. 68–80, 2020.
[3] H. Boök, A. Poikonen, A. Aarva, T. Mielonen, M. R. A. Pitkänen, and A. V. Lindfors, ”Photovoltaic system modeling: a validation study at high latitudes with implementation of a novel DNI quality control method,” Solar Energy, vol. 204, pp. 316–329, 2020.
[4] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, “A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization,” Renewable and Sustainable Energy Reviews, vol. 124, p. 109792, 2020.
[5] H. T. C. Pedro, C. F. M. Coimbra, M. David, and P. Lauret, “Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts,” Renewable Energy, vol. 123, pp. 191–203, 2018.
[6] C. Voyant, F. Motte, G. Notton, A. Fouilloy, M. L. Nivet, and J. L. Dauhaut, “Prediction intervals for global solar irradiation forecasting using regression trees methods,” Renewable Energy, vol. 126, pp. 332–340, 2018.
[7] L. A. Fernandez-Jimenez, S. Terreros-Olarte, M. Mendoza-Viliena et al., “Day-ahead probabilistic photovoltaic power forecasting models based on quantile regression neural
networks,” in 2017 European Conference on Electrical Engineering and Computer Science (EECS), pp. 89–294, Bern, Switzerland, November 2017.

[8] Y. Chu, M. Li, H. T. C. Pedro, and C. F. M. Coimbra, “Real-time prediction intervals for intra-hour DNI forecasts,” *Energy*, vol. 83, pp. 234–244, 2015.

[9] Y. Ren, P. N. Suganthan, and N. Srikanth, “Ensemble methods for wind and solar power forecasting—a state-of-the-art review,” *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 82–91, 2015.

[10] J. le Gal la Salle, J. Badosa, M. David, P. Pinson, and P. Lauret, “Added-value of ensemble prediction system on the quality of solar irradiance probabilistic forecasts,” *Renewable Energy*, vol. 162, pp. 1321–1339, 2020.

[11] S. Chattopadhyay, R. Dubey, S. Bhaduri et al., “Correlating infrared thermography with electrical degradation of PV modules inspected in all-India survey of photovoltaic module reliability 2016,” *IEEE Journal of Photovoltaics*, vol. 8, no. 6, pp. 1800–1808, 2018.

[12] Y. Hishikawa, T. Doi, M. Higa et al., “Voltage-dependent temperature coefficient of the I-V curves of crystalline silicon photovoltaic modules,” *IEEE Journal of Photovoltaics*, vol. 8, no. 1, pp. 48–53, 2018.

[13] C. Li, “Unstable operation of photovoltaic inverter from field experiences,” *IEEE Transactions on Power Delivery*, vol. 33, no. 2, pp. 1013–1015, 2018.

[14] A. A. Babatunde and S. Abbasoglu, “Predictive analysis of photovoltaic plants specific yield with the implementation of multiple linear regression tool,” *Environmental Progress & Sustainable Energy*, vol. 38, no. 4, p. 13098, 2019.

[15] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, “Solar photovoltaic generation forecasting methods: a review,” *Energy Conversion and Management*, vol. 156, pp. 459–497, 2018.

[16] J. Kleissl, *Solar Energy Forecasting and Resource Assessment*, Academic Press, 2013.

[17] S. Theocharides, G. Makrides, A. Livera, M. Theristis, P. Kaimakis, and G. E. Georgiou, “Day-ahead photovoltaic power production forecasting methodology based on machine learning and statistical post-processing,” *Applied Energy*, vol. 268, article 110203, 2020.

[18] J. Antonanzas, D. Pozo-Vázquez, L. A. Fernandez-Jimenez, and F. J. Martinez-de-Pison, “The value of day-ahead forecasting for photovoltaics in the Spanish electricity market,” *Solar Energy*, vol. 158, pp. 140–146, 2017.

[19] M. P. Almeida, M. Muñoz, I. de la Parra, and O. Perpiñán, “Comparative study of PV power forecast using parametric and nonparametric PV models,” *Solar Energy*, vol. 155, pp. 854–866, 2017.

[20] B. Wolff, J. Kühnert, E. Lorenz, O. Kramer, and D. Heinemann, “Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data,” *Solar Energy*, vol. 135, pp. 197–208, 2016.

[21] M. Schmelas, T. Feldmann, J. da Costa Fernandes, and E. Bollin, “Photovoltaics energy prediction under complex conditions for a predictive energy management system,” *Journal of Solar Energy Engineering*, vol. 137, no. 3, pp. 1–10, 2015.

[22] E. Ogliari, A. Dolara, G. Manzolini, and S. Leva, “Physical and hybrid methods comparison for the day ahead PV output power forecast,” *Renewable Energy*, vol. 113, pp. 11–21, 2017.