International Journal of Power Electronics and Drive System (IJPEDS)
Vol. 16, No. 4, December 2025, pp. 2868~2878
ISSN: 2088-8694, DOI: 10.11591/ijpeds.v16.i4.pp2868-2878 ❒ 2868

Predictions of solar power using ensemble machine learning techniques

Arangarajan Vinayagam[1], R. Mohandas[1], R. Jeyabharath[2], B. S. Mohan[3], Srinivasan Lakshmanan[4], C. Bharatiraja[5]
[1]Department of Electrical and Electronics Engineering, New Horizon College of Engineering, Bengaluru, India
[2]Department of Electrical and Electronics Engineering, KSR College of Engineering, Namakkal, India
[3]Department of Electrical and Electronics Engineering, SJB Institute of Technology, Bengaluru, India
[4]Department of Computer Science and Engineering, Dr. NGP Institute of Technology, Coimbatore, India
[5]Center for Electric Mobility, Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur, India

Article Info

Article history:

Received Feb 9, 2025
Revised Nov 5, 2025
Accepted Nov 14, 2025

Keywords:

Extreme gradient boosting
Gradient boosting
Photovoltaic
Prediction
Random forest
Solar power

ABSTRACT

Predicting solar power production accurately is becoming more and more crucial for efficient power management and the grid's integration of renewable energy sources. Using data from an Australian photovoltaic (PV) power station, this study employs a variety of machine learning (ML) ensemble techniques, such as gradient boosting (GB), random forest (RF), and extreme gradient boosting (XGBoost), to forecast solar power production. ML models are developed utilizing pertinent information from electricity and meteorological data in order to forecast solar power. The predictive performance of trained ML models is verified in terms of metrics like mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient (R^2). With higher R^2 values and lower error results (MAE and RMSE), XGBoost performs better than GB and RF. Optimizing the hyperparameters of the XGBoost model significantly improves its performance. The tweaked XGBoost model shows a significant improvement in R^2 (more than 5% to 10%) and error results (reduced MAE and RMSE by 0.01 to 0.06), when compared to other ensemble approaches. Compared to other ensemble approaches, the tuned XGBoost methodology is more robust and generates more accurate forecasts in solar power.

Corresponding Author:

C. Bharatiraja
Center for Electric Mobility, Department of Electrical and Electronics Engineering
SRM Institute of Science and Technology
Kattankulathur, Tamil Nadu 603203, India
Email: bharatic@srmist.edu.in, bharatiraja@gmail.com

1. INTRODUCTION

The integration of solar energy to power networks imposes challenges as a result of its fluctuation in response to weather conditions, complicating the maintenance of a dependable power supply. Erroneous predictions of solar power can cause operational difficulties, possibly resulting in grid instability and financial detriment [1]. Accurate predictions are crucial for efficient spinning reserve management and power supply control during contingency events, and exact estimates are necessary for solar energy to be sold competitively in power markets. Consequently, forecasting solar power has become a significant difficulty in the renewable energy sector due to the intermittent and weather-dependent nature of solar energy. Accurate forecasting of solar power output is essential for efficient grid management, optimum resource allocation, and successful energy trading [1]-[3].

There are different methods that researchers have investigated to predict the generation of solar power, including both statistical and machine learning-based approaches. A lot of statistical methods have been used, such as linear regression, exponential smoothing, and autoregressive integrated moving average (ARIMA) [4], [5]. The above methods are based on historical data and mathematical models to make predictions; however, they can confront issues in capturing the intricate and nonlinear relationships that are inherent in solar power generation. ML approaches, on the other hand, are effective for solar power forecasting because they can capture complex patterns and non-linear connections in data from solar power generation [6]. Research utilizing support vector machine (SVM) models [7], the k-nearest neighbors (KNN) technique [8], and tree-based models, including random forest (RF) [9] and decision trees [10], exemplifies the efficacy of machine learning methods for this issue. Recent decade-emerged deep learning (DL) models that predict solar power. The advancement of DL has allowed for accurate problem-solving [11]. Compared to statistical and ML methods, feedforward neural networks (FFNN), recurrent neural networks (RNN), and long short-term memory (LSTM) models are increasingly used in photovoltaic (PV) power predictions [12], [13]. Even though deep learning methods are more robust and have demonstrated enhanced accuracy compared to traditional methods, they are only effective on large datasets and encounter challenges with training complexity [14], [15]. Additionally, DL models are susceptible to overfitting, which can result in inadequate generalization on unseen data if the appropriate regularization method is not implemented during the model training process [3].

Furthermore, solar energy forecasting research increasingly emphasizes ensemble approaches over individual machine learning techniques [16]-[18]. Antonanzas *et al.* [19] tested ensemble learning approaches, including gradient boosting and AdaBoost, to predict solar power. They found that ensemble models outperformed single models by minimizing biases and uncertainties. Ensemble approaches showed promise for more accurate solar energy estimates. Using random forest, an ensemble learning technique, numerous decision trees enhance prediction accuracy and handle high-dimensional data [20]. In addition, LGBM and other gradient boosting algorithms use boosting techniques to improve model performance [21]. Ahmad *et al.* [10] forecasted short-term solar energy output using real radiation data and basic radiation forecasts. The random forest predicting method was used to predict solar power in [22]. RF accuracy, extra trees (ET), computational cost, and stability were examined for hourly PV generating output prediction. The performance was compared to support vector regression (SVR). Model performance evaluation using RMSE, MAE, and R^2 revealed that RF outperformed other models on both testing and training datasets. Furthermore, a variety of algorithms, including AdaBoost [23], gradient boosting machine (GBM) [24], [25], extreme gradient boosting machine (XGBM) [26], and light gradient boosting machine (LGBM) [27] also been utilized for solar power predictions. Ensemble models that are based on boosting and bagging were examined in order to forecast short-term solar irradiation in [28]. Likewise, multiple tree-based ensemble approaches (DT, GBM, XGBM bagging, and RF) were employed to forecast solar energy, with performance confirmed using MAE, RMSE, and RMSE metrics [29].

The literature clearly indicates that the traditional statistical methods, like ARIMA, provide only rudimentary forecasting capabilities. Solar power generation involves complex and nonlinear interactions, which can be difficult for statistical approaches to capture. On the other hand, DL-based methods offer advanced capabilities, but they come with data requirements and interpretability problems. Combining numerous learners into an ensemble model improves prediction accuracy and robustness compared to using a single ML model. They are less likely to experience overfitting, work effectively with small to medium datasets, and require less computational resources than deep learning. As a result, this study proposes ensemble machine learning methods for predicting solar power generation. Among the models tested, XGBoost outperforms gradient boosting and random forest in prediction accuracy. Furthermore, XGBoost's hyperparameters are tuned to improve performance even more. The key contributions of this study are summarized below:

– A predictive analysis is carried out to predict solar power generation using ensemble ML models (GBM, RF, and XGBoost).
– After preprocessing the PV power generation data, the ensemble models are trained and evaluated for their predictive performance in terms of metrics like MAE, RMSE, and R^2.
– To get higher predictive accuracy, an improved XGBoost ensemble model is developed by tuning the hyperparameters of the XGBoost model.

2. MATERIALS AND METHODS

In this study, the employed ensemble ML models (gradient boosting, random forest, and XGBoost) are used to predict solar power generation using a dataset obtained from an Australian DKASC photovoltaic power station [30]. With a temporal resolution of five minutes (288 samples daily and a total of 95281 samples), the dataset comprises PV power (active power) and weather data (wind speed, temperature, global horizontal radiation, wind direction, and air pressure) with periods ranging from 01-01-2024 to 30-11-2024. Each experimental investigation is conducted within the Python (3.10) environment provided by the Google Colaboratory (Colab) platform. The primary platform is Colab, which includes pre-installed Python libraries, cloud storage, and free access to GPU and TPU resources. These resources are highly beneficial for the rapid training and processing of ensemble ML models. Pre-exploratory data analysis (EDA) and fundamental tasks are conducted on a personal computer (Intel® Core™ i5-1135G7 @ 2.40 GHz with 16 GB of RAM). This section outlines the forecasting framework and machine learning pipeline for solar power forecasting. The methodology is intended to guarantee reproducibility by providing explicit information on data preprocessing, model architectures, and evaluation protocols. The workflow of the solar power prediction process is shown in Figure 1.

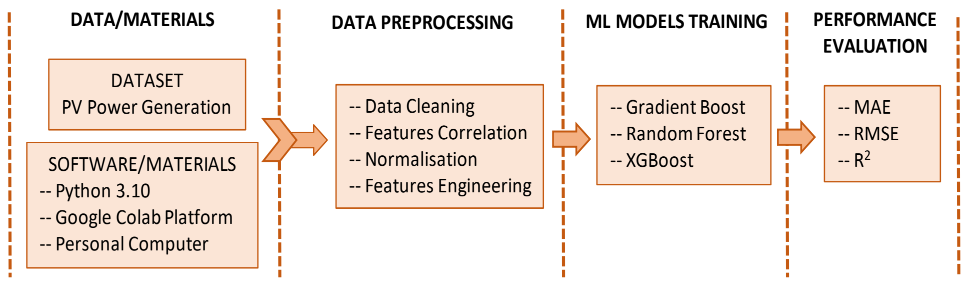

Figure 1. Workflow of solar power prediction

2.1. Data preprocessing

During the data cleaning phase, missing values in the dataset, such as null values, can be filled in using an appropriate imputation standard, such as mean, interpolation, or regression. Next, a range of outliers is then computed using the interquartile range (IQR) and scoped out to finally take the cleaned dataset, which is then used to train the ML model. After the data is cleaned, modelling and scaling are carried out through Min-Max scaling so that all features can be in the same range before the training of the model. Furthermore, Python's correlation analysis has been conducted using a heatmap visualization tool from the Seaborn library. It generates a correlation matrix that helps to find patterns among variables linked to patterns. The correlation heatmap in Figure 2 shows quite high connections between solar power production (Active Power) and worldwide horizontal radiation (including pyranometer), temperature, and wind speed. On the other hand, air pressure and wind direction have hardly any relationship with electricity production. In the end, the less closely connected variables are eliminated; the other variables are kept as the key ones used in training the ML model. While carrying out feature engineering, the most essential weather-based feature columns and the time-based feature columns (day, month, and year) are changed and updated appropriately as per the requirements. ML models have been shown to improve their solar power forecasting skills with feature selection. Enhanced solar forecasting will lead to better planning and energy management.

2.2. Ensemble ML models

This research investigates several ensemble methods including gradient boosting, XGBoost, and random forest techniques, in predicting the solar power output based on the dependent meteorological variables and the previous performance measurements. Every one of these ensemble methods seeks to utilize the benefits within various models in order to enhance forecasting capabilities and to minimize overfitting and therefore provide better forecasts across various environmental settings.

2.2.1. Gradient boosting (GB)

Gradient boosting, also referred to as GB, is a supervised machine learning technique, in particular, one of the ensemble methods. Sequential models are built in this approach, and each new model attempts to correct the errors in prior models. For regression tasks, the GB technique works best as it reduces the loss function using a gradient descent technique. Baskets of models may be used, focusing on the residuals from the previous models to identify more complex structures. This approach also proved effective in managing non-linear relationships and therefore improves the accuracy of solar power forecasting models, as has been shown in research. GB can be mathematically formulated as in (1) [31], [32].

$$y_p = F_o(x) = \sum(f\_i_{dt}(x)) \quad (1)$$

Where Y_P = prediction output, $F_O(x)$ = overall prediction, $f\_i_{dt}(x)$ = function of the individual decision tree model, and (x) = input features. Each learner fitted to the composite learner's output, during the model training procedure is trained to the negative gradient of the loss in relation to the output. Over time, this process improves the model by shifting biases and recessing intricate features in the provided data set.

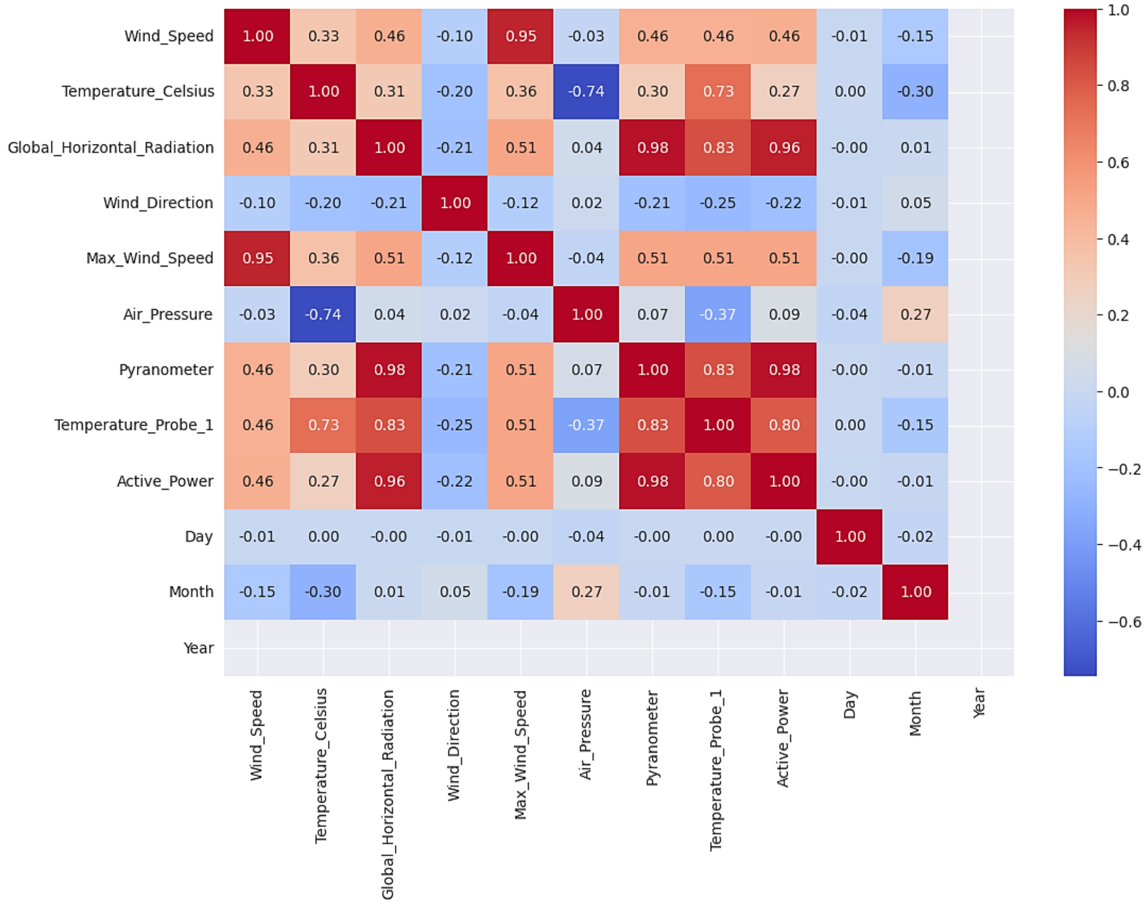

Figure 2. Correlation matrix

2.2.2. Extreme gradient boosting (XGBoost)

XGBoost is a superset of basic gradient boosting: it adds regularization methods in order to prevent overfitting. The algorithm uses multi-threading in order to speed up calculations without losing the quality of the predictions. Due to its efficiency and robustness, XGBoost is widely used within machine learning contests and in real-world applications such as forecasting of solar energy production [3], [33]. The mathematical expression of the functionality of XGBoost is given in (2).

$$y_{pt} = y_{pt-1} + \eta \sum_{it}^{n} F_{it}(x) \quad (2)$$

Where y_{pt} = predicted output at iteration t, y_{pt-1} = prediction at last iteration, η = factor of learning rate, $F_{it}(x)$ = decision tree function fitted to the residuals. Gradient boosting works well for predicting solar power because it can handle complicated data relationships, especially ones with non-linear trends and traits that interact with each other [34].

2.2.3. Random forest (RF)

Random forest is a powerful ensemble approach that creates many decision trees during the training phase and predicts the mean for regression issues. Random forest significantly reduces volatility and improves overall model stability by combining predictions from many trees. Its ability to accommodate large datasets with high dimensionality makes it particularly suitable for solar power forecasting, where several environmental factors influence output [35]. The random forest prediction formula may be represented as (3).

$$y_{pt}= \frac{1}{k}\sum_{k=1}^{k} T_k(x) \tag{3}$$

Where y_{pt}= predicted output for the input features *(x)*, k = total number of decision trees, $T_k(x)$= prediction of k^{th} decision tree. This approach mitigates variation by averaging estimates from several trees, making it resilient to overfitting and efficient for forecasting solar energy production based on diverse climatic variables [33], [36].

2.3. ML model training and evaluation metrics

During the training phase, the chosen ML model gets exposed to historical time series data. The data set is divided into training and testing sets to facilitate model training and assessment. The time series dataset is divided in such a manner that temporal order is preserved before training. In this study, the training and testing datasets are divided as follows: training set (80% data from 01-01-2024 to 25-09-2024) and testing set (20% data from 26-09-2024 to 30-11-2024). The training set is used to train the machine learning model, allowing it to recognize patterns and correlations between input features (weather-related variables) and the target variable (PV active power).

The trained model is assessed for performance using the testing data. Common assessment measures include mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient (R^2). These measures reveal the model's strengths and shortcomings by contrasting its predicted results with actual data [32]. The definition of performance metrics and their mathematical expressions are given as (4) [32], [37].

$$\text{MAE} = \frac{1}{k}\sum_{i=1}^{k}\left|y_{pi} - \mathrm{x}_{ai}\right| \tag{4}$$

Where k = total number of observations, y_{pi} = predicted value y_p at observation i, x_{ai} = actual value x_a at observation i.

$$\text{RMSE} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_{pi} - \mathrm{x}_{ai})^2} \tag{5}$$

$$R^2 = 1 - \frac{SSR}{SST} \tag{6}$$

Where SSR = sum of squares of residuals, $\sum_{i=1}^{k}(y_{ai} - y_{pi})^2$ (y_{ai} = actual values and y_{pi} = predicted values) and SST = sum of squares Total, $\sum_{i=1}^{k}(y_{ai} - y_M)^2$ (y_{ai} = actual values and y_M = mean of the actual values).

3. RESULTS AND DISCUSSION

The solar power generation was predicted in this work using a variety of ensemble ML approaches, including GB, RF, and XGBoost. The ML models were trained using the dataset's [30] power and weather-related attributes. Figure 3 shows the average distribution of solar power over the days (specific month) and over the months.

In Python, the Scikit-learn library is commonly used to implement ensemble algorithms. The Scikit-learn toolkit in Python is often used to develop ensemble methods. The default parameters for GB (learning rate = 0.1, estimators = 100, subsample = 1, max depth = 3), RF (estimators = 100, min samples split = 2, min samples leaf = 1), and XGBoot (learning rate = 0.3, estimators = 100, max depth = 6, min child weight = 1) were utilized during the training of the algorithms. The efficacy of ML models was assessed using important metrics, including MAE, RMSE, and R2. The outcomes of the assessed measures are shown in Table 1. The bar chart in Figure 4 clearly illustrates the comparative outcomes for MAE, RMSE, and R2 across all approaches.

Table 1. Results of performance metrics

Ensemble models	MAE	RMSE	R^2
Gradient boosting (GB)	0.04629	0.10351	0.86257
Extreme GB (XGBoost)	0.03880	0.08160	0.91460
Random forest (RF)	0.04010	0.08860	0.89920
Tuned XGBoost	0.02830	0.05550	0.96060

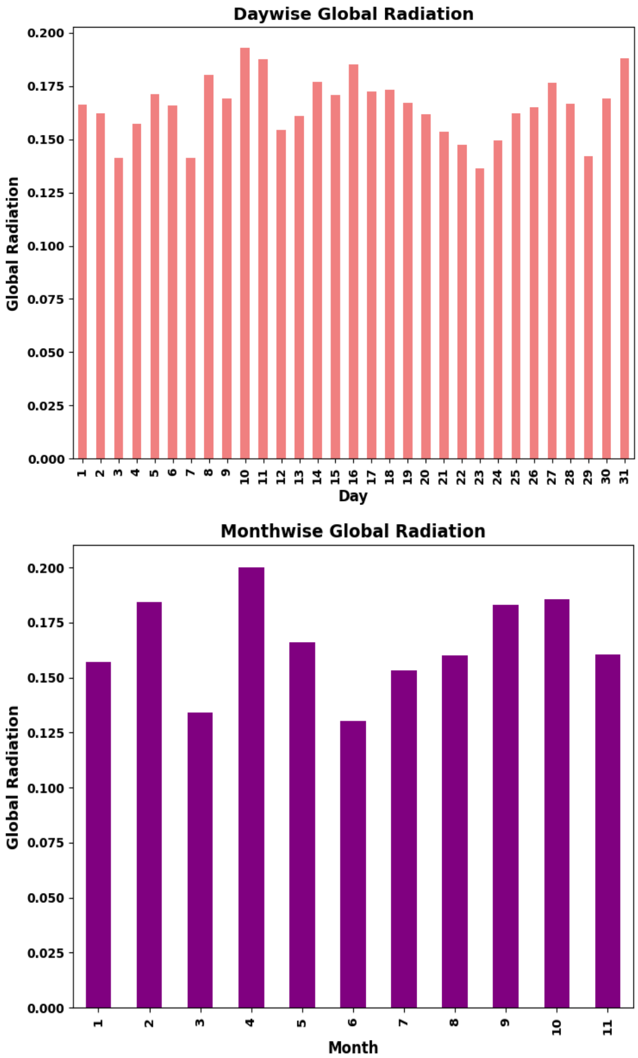

Figure 3. Solar power (average) distribution over days and months

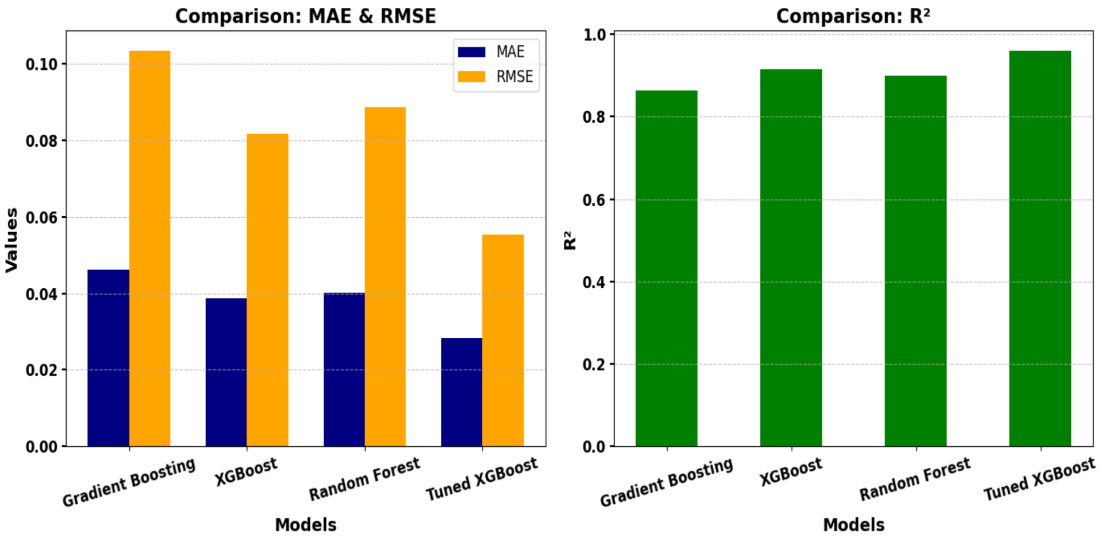

Figure 4. Comparative results of performance metrics

The findings show that RF outperforms GB in terms of correlation coefficients (R^2) and error results (MAE and RMSE). XGBoost provides superior R^2 and error outcomes than GB and RF. The XGBoost method performed better when the key hyperparameters were adjusted, such as the maximum depth (which may range from 3 to 7), the learning rate (which can range from 0.01 to 0.2), and the estimator (which can range from 50 to 150). With the tuned XGBoost model, a significant improvement in R^2 (7% to 10% over other ensembles) and error values (0.03 to 0.06 over other ensembles) was achieved.

Figure 5 depicts the relationship between actual and predicted active power, demonstrating that the predicted power of tuned XGBoost was more linear with actual power than other approaches. Furthermore, projected active power across a variety of time samples, as shown in Figure 6, clearly reveals that XGBoost and tuned XGBoost have lower error and are closer to the actual power values than other approaches. The overall findings show that the tuned XGBoost approach outperforms other ensemble algorithms in terms of robustness and accuracy of solar power prediction.

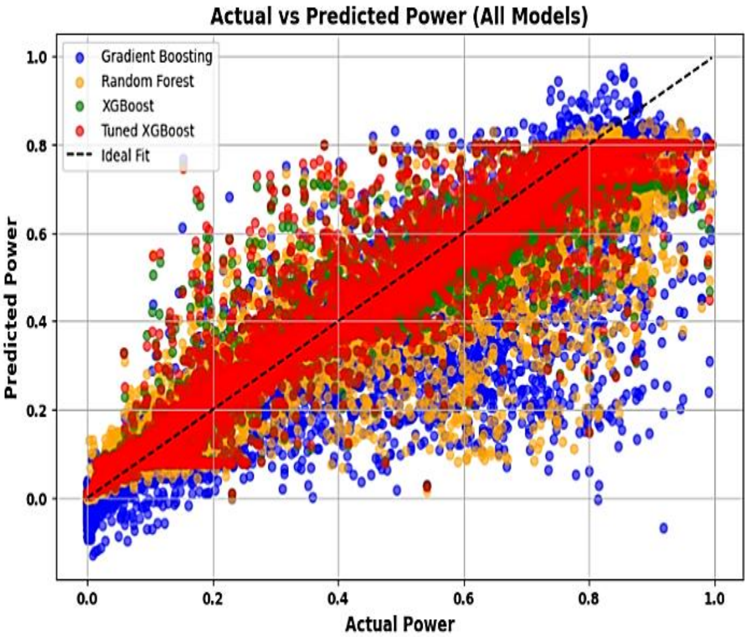

Figure 5. Actual versus predicted solar power

Furthermore, we conducted a comparative study between the tuned XGBoost predictive model and existing predictive methods. According to the comparison, out of all the methods listed in Table 2, the suggested tuned XGBoost model attained the highest prediction accuracy ($R^2 = 0.960$). Advanced methods like LSTM (0.938) and LightGBM (0.940) exhibit superior performance, whereas conventional models like linear regression ($R^2 = 0.718$) perform only moderately. The comparison clearly demonstrates that the proposed tuned XGBoost model offers superior performance and is more reliable for predicting the solar power of the given dataset. The overall findings of this study clearly show that the tuned XGBoost approach outperforms other ensemble algorithms in terms of robustness and accuracy of solar power prediction.

Table 2. Comparative (R^2) results between the tuned XGBoost model and the existing literature

Reference	Predictive method	R^2
[4]	Linear regression	0.718
[10]	Extremely randomized trees	0.927
[11]	LSTM	0.938
[17]	Light gradient boosting machine	0.940
[38]	Hybrid (CNN+LSTM)	0.845
Proposed	Tuned XGBoost	0.960

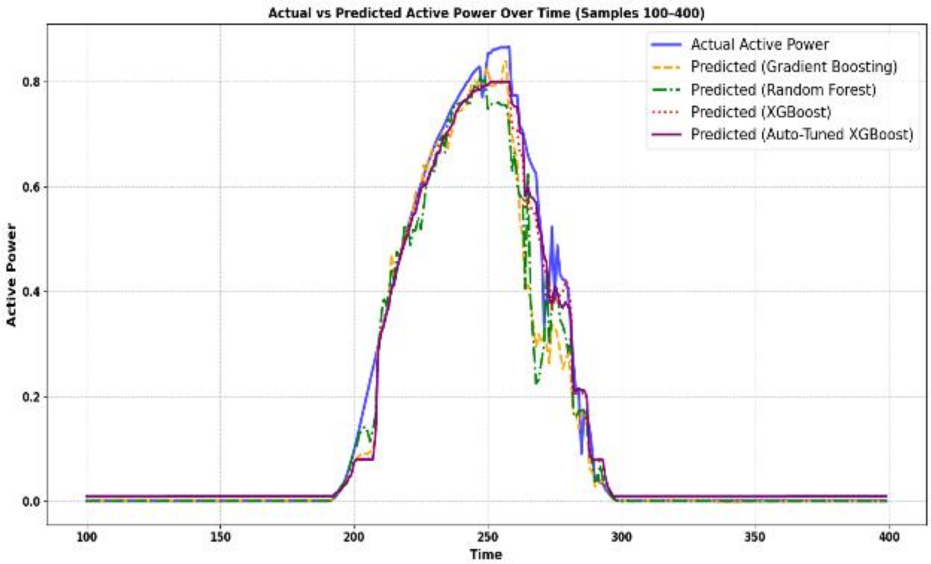

Figure 6. Predicted active power

4. CONCLUSIONS

In this study, a variety of ensemble techniques, including GB, RF, XGBoost, and optimally tuned XGBoost, were implemented to predict solar power. For the solar power forecasting, the dataset from the PV power facility in Australia was taken into account. The active power of solar PV was predicted by training the ensemble models with the appropriate features. The performance metrics MAE, RMSE, and R^2 were used to verify the predictive capabilities of ML ensemble approaches. In terms of correlation coefficient results (R^2) and error measurements (MAE and RMSE), the XGBoost model exhibits superior performance among ensemble approaches. By tuning the important hyperparameters, the performance of the XGBoost model was enhanced. Regarding performance assessments, the tuned XGBoost model presents better performance and forecasts solar power more precisely than other ensemble methods. It is shown that the tuned XGBoost model is more resilient and competent to provide possible contributions for the efficient power management and integration of solar energy into the power system.

FUNDING INFORMATION

This work was supported by the Scheme for Promotion of Academic and Research Collaboration (SPARC), MoE, under grant File No: 2A.184, and DST, Government of India, under Award PURSE Program grant File No: SR/PURSE/2021/65.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Arangarajan Vinayagam	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
R. Mohandas		✓				✓		✓	✓	✓	✓	✓		
R. Jeyabharath		✓				✓		✓	✓	✓	✓	✓		
B. S. Mohan	✓		✓	✓		✓		✓		✓			✓	✓
Srinivasan Lakshmanan		✓			✓		✓			✓		✓		✓
C. Bharatiraja					✓		✓			✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [CB], upon reasonable request.

REFERENCES

[1] H. I. Aouidad and A. Bouhelal, "Machine learning-based short-term solar power forecasting: a comparison between regression and classification approaches using extensive Australian dataset," *Sustainable Energy Research*, vol. 11, no. 1, p. 28, Aug. 2024, doi: 10.1186/s40807-024-00115-1.

[2] A. Zameer, F. Jaffar, F. Shahid, M. Muneeb, R. Khan, and R. Nasir, "Short-term solar energy forecasting: Integrated computational intelligence of LSTMs and GRU," *PLOS ONE*, vol. 18, no. 10, p. e0285410, Oct. 2023, doi: 10.1371/journal.pone.0285410.

[3] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," *Applied Computing and Informatics*, vol. 20, no. 3/4, pp. 231–250, Jun. 2024, doi: 10.1016/j.aci.2019.11.002.

[4] M. Y. ERTEN and H. AYDİLEK, "Solar Power Prediction using Regression Models," *Uluslararası Muhendislik Arastirma ve Gelistirme Dergisi*, vol. 14, no. 3, pp. 1–1, Dec. 2022, doi: 10.29137/umagd.1100957.

[5] P. Suanpang and P. Jamjuntr, "Machine learning models for solar power generation forecasting in microgrid application implications for smart cities," *Sustainability*, vol. 16, no. 14, p. 6087, Jul. 2024, doi: 10.3390/su16146087.

[6] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control, 5th Edition*. Wiley, 2015.

[7] W. VanDeventer *et al.*, "Short-term PV power forecasting using hybrid GASVM technique," *Renewable Energy*, vol. 140, pp. 367–379, Sep. 2019, doi: 10.1016/j.renene.2019.02.087.

[8] D. S. Tripathy, B. R. Prusty, and K. Bingi, "A k-nearest neighbor-based averaging model for probabilistic PV generation forecasting," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 35, no. 2, Mar. 2022, doi: 10.1002/jnm.2983.

[9] D. Niu, K. Wang, L. Sun, J. Wu, and X. Xu, "Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study," *Applied Soft Computing*, vol. 93, p. 106389, Aug. 2020, doi: 10.1016/j.asoc.2020.106389.

[10] M. W. Ahmad, M. Mourshed, and Y. Rezgui, "Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression," *Energy*, vol. 164, pp. 465–474, Dec. 2018, doi: 10.1016/j.energy.2018.08.207.

[11] M. Elsaraiti and A. Merabet, "Solar power forecasting using deep learning techniques," *IEEE Access*, vol. 10, pp. 31692–31698, 2022, doi: 10.1109/ACCESS.2022.3160484.

[12] I. Jebli, F.-Z. Belouadha, M. I. Kabbaj, and A. Tilioua, "Deep learning based models for solar energy prediction," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 1, pp. 349–355, Jan. 2021, doi: 10.25046/aj060140.

[13] S. S. Punyam Rajendran and A. Gebremedhin, "Deep learning-based solar power forecasting model to analyze a multi-energy microgrid energy system," *Frontiers in Energy Research*, vol. 12, Mar. 2024, doi: 10.3389/fenrg.2024.1363895.

[14] Kelvin Edem Bassey, "Solar energy forecasting with deep learning technique," *Engineering Science & Technology Journal*, vol. 4, no. 2, pp. 18–32, Jun. 2023, doi: 10.51594/estj.v4i2.1286.

[15] W.-C. Kuo, C.-H. Chen, S.-H. Hua, and C.-C. Wang, "Assessment of different deep learning methods of power generation forecasting for solar PV system," *Applied Sciences*, vol. 12, no. 15, p. 7529, Jul. 2022, doi: 10.3390/app12157529.

[16] S. K. Satyanarayana and A. N. Rao, "Solar power prediction using soft voting based ensemble machine learning classifier," in *Advances in Microelectronics, Embedded Systems and IoT*, 2024, pp. 529–539. doi: 10.1007/978-981-97-0767-6_44.

[17] D. Chakraborty, J. Mondal, H. B. Barua, and A. Bhattacharjee, "Computational solar energy – ensemble learning methods for prediction of solar power generation based on meteorological parameters in Eastern India," *Renewable Energy Focus*, vol. 44, pp. 277–294, Mar. 2023, doi: 10.1016/j.ref.2023.01.006.

[18] V. Raj, S.-Q. Dotse, M. Sathyajith, M. I. Petra, and H. Yassin, "Ensemble machine learning for predicting the power output from different solar photovoltaic systems," *Energies*, vol. 16, no. 2, p. 671, Jan. 2023, doi: 10.3390/en16020671.

[19] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, Oct. 2016, doi: 10.1016/j.solener.2016.06.069.

[20] S. Shahzad, M. A. Abbasi, H. Ali, M. Iqbal, R. Munir, and H. Kilic, "Possibilities, challenges, and future opportunities of microgrids: A review," *Sustainability*, vol. 15, no. 8, p. 6366, Apr. 2023, doi: 10.3390/su15086366.

[21] J. Zhang, Z. Liu, and T. Chen, "Interval prediction of ultra-short-term photovoltaic power based on a hybrid model," *Electric Power Systems Research*, vol. 216, p. 109035, Mar. 2023, doi: 10.1016/j.epsr.2022.109035.

[22] C. Persson, P. Bacher, T. Shiga, and H. Madsen, "Multi-site solar power forecasting using gradient boosted regression trees," *Solar Energy*, vol. 150, pp. 423–436, Jul. 2017, doi: 10.1016/j.solener.2017.04.066.

[23] H. Kim and D. Lee, "Probabilistic solar power forecasting based on bivariate conditional solar irradiation distributions," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 4, pp. 2031–2041, Oct. 2021, doi: 10.1109/TSTE.2021.3077001.

[24] J. Park, J. Moon, S. Jung, and E. Hwang, “Multistep-ahead solar radiation forecasting scheme based on the light gradient boosting machine: A case study of Jeju Island,” *Remote Sensing*, vol. 12, no. 14, p. 2271, Jul. 2020, doi: 10.3390/rs12142271.
[25] H. Tyralis and G. Papacharalampous, “Boosting algorithms in energy research: a systematic review,” *Neural Computing and Applications*, vol. 33, no. 21, pp. 14101–14117, Nov. 2021, doi: 10.1007/s00521-021-05995-8.
[26] J. Fan *et al.*, “Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China,” *Energy Conversion and Management*, vol. 164, pp. 102–111, May 2018, doi: 10.1016/j.enconman.2018.02.087.
[27] S. Choi and J. Hur, “An ensemble learner-based bagging model using past output data for photovoltaic forecasting,” *Energies*, vol. 13, no. 6, p. 1438, Mar. 2020, doi: 10.3390/en13061438.
[28] T. McCandless, S. Dettling, and S. E. Haupt, “Comparison of implicit vs. explicit regime identification in machine learning methods for solar irradiance prediction,” *Energies*, vol. 13, no. 3, p. 689, Feb. 2020, doi: 10.3390/en13030689.
[29] J. Moon, Z. Shin, S. Rho, and E. Hwang, “A Comparative analysis of tree-based models for day-ahead solar irradiance forecasting,” in *2021 International Conference on Platform Technology and Service (PlatCon)*, Aug. 2021, pp. 1–6. doi: 10.1109/PlatCon53246.2021.9680748.
[30] DKA Solar Centre, “Yulara Solar System”, Central Australia, 2025. [Online]. Available: https://www.dkasolarcentre.com.au/download?location=yulara.
[31] A. Shah, V. Viswanath, K. Gandhi, and N. M. Patil, “Predicting solar energy generation with machine learning based on AQI and Weather Features,” Oct. 2024, doi: 10.21203/rs.3.rs-3178713/v1.
[32] E. Raafat Maamoun Shouman, “Solar power prediction with artificial intelligence,” in *Advances in Solar Photovoltaic Energy Systems*, IntechOpen, 2024. doi: 10.5772/intechopen.1002726.
[33] B. Aljafari, D. S, B. C, P. K. Balachandran, and T. S. Babu, “Power enhanced solar PV array configuration based on calcudoku puzzle pattern for partial shaded PV system,” *Heliyon*, vol. 9, no. 5, p. e16041, May 2023, doi: 10.1016/j.heliyon.2023.e16041.
[34] N. Aksoy and I. Genc, “Predictive models development using gradient boosting based methods for solar power plants,” *Journal of Computational Science*, vol. 67, p. 101958, Mar. 2023, doi: 10.1016/j.jocs.2023.101958.
[35] M. Ali *et al.*, “Robust ANN-based control of modified PUC-5 inverter for solar PV applications,” *IEEE Transactions on Industry Applications*, vol. 57, no. 4, pp. 3863–3876, Jul. 2021, doi: 10.1109/TIA.2021.3076032.
[36] J. Wang, P. Li, R. Ran, Y. Che, and Y. Zhou, “A short-term photovoltaic power prediction model based on the gradient boost decision tree,” *Applied Sciences*, vol. 8, no. 5, p. 689, Apr. 2018, doi: 10.3390/app8050689.
[37] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” *PeerJ Computer Science*, vol. 7, p. e623, Jul. 2021, doi: 10.7717/peerj-cs.623.
[38] P. Mohanty, U. C. Pati, and K. Mahapatra, “Deep learning based framework for forecasting solar panel output power,” 2024, pp. 229–239. doi: 10.1007/978-3-031-45878-1_16.

BIOGRAPHIES OF AUTHORS

Arangarajan Vinayagam has over 15 years of industrial experience in the power plant and renewable energy industries. He also has academic experience (teaching and research) in smart grid and microgrid power systems, renewable energy technologies, power generation, and distribution control. His research interests include microgrids, electric vehicles, machine learning, power quality, fault analysis, and forecasting analysis of solar PV and wind generation. He received a Ph.D. from Deakin University, Australia, in 2017. He is currently working as a visiting professor and research fellow (Remote) at the Institute of Power Engineering, UNITEN, Malaysia. He can be contacted at email: arajanin@gmail.com.

R. Mohandas has received his Ph.D. degree in 2020 from Anna University, Chennai, India. He is currently working as an associate professor in the Department of Electrical and Electronics Engineering at New Horizon College of Engineering, Bangaluru, India. His research interests include microgrids, PWM techniques, power electronics, and renewable energy sources. He can be contacted at email: mohandas@newhorizonindia.edu.

R. Jeyabharath is a professor and director (accreditation & rankings) in the Department of Electrical and Electronics Engineering at K. S. R. College of Engineering and Technology, Tiruchengode, Tamil Nadu, India. He obtained his B.E. in Electrical and Electronics Engineering, M.E. in power electronics and drives, and Ph.D. in Electrical Engineering from Anna University, Chennai. His research interests include power electronics, special electrical machines, and embedded systems. He can be contacted at email: jeyabharath@ksriet.ac.in.

B. S. Mohan completed his undergraduate program in electrical and electronics engineering from SJB Institute of Technology, Bengaluru, and obtained his M.Tech. in power systems from the Department of EEE, Acharya Institute of Technology, Bengaluru. He is currently working as an assistant professor in the Department of Electrical and Electronics Engineering, SJB Institute of Technology, Bengaluru, with a total teaching experience of 13 years and 7 months. His areas of interest include grid integration of renewable and inverter-based resources, power electronics and converter technologies with system-level impact, fractional frequency transmission, and stability analysis of hybrid AC-DC microgrids integrating PV systems, electric vehicles, and battery energy storage. He can be contacted at email: sjbit.mohan@gmail.com.

Srinivasan Lakshmanan is working as an associate professor at Dr N.G.P. Institute of Technology, Coimbatore, Tamil Nadu. He received his Ph.D. degree from Anna University, Chennai, and M.Tech. (IT) from Anna University Coimbatore and B.E. (IT) from V.L.B Janakiammal College of Engineering & Technology, Coimbatore in 2020,2011 and 2004, respectively. His current research includes machine learning, text mining, large data analytics, and IoT. He has published 3 Science Citation Index Expanded (SCIE) Journals and 2 Emerging Sources Citation Index (ESCI) journals, 3 Scopus-indexed (SI) journals, 5 research articles in Scopus-indexed conference proceedings, and 3 papers as book chapters in reputed Scopus books. He had filed 4 Indian Patents and been granted 1 Indian and 1 Australian Patent. He had published 2 books titled ‘Mobile app development for beginners’ and Mobile application development using Android. He has more than sixteen years of professional experience in Various Engineering Colleges in Tamil Nadu and Karnataka. He can be contacted at email: srinivasanl.1982@gmail.com.

C. Bharatiraja is a professor in the Department of Electrical and Electronics Engineering at SRM Institute of Science and Technology (Kattankulathur Campus), Chennai, India. He has completed six sponsored projects from various government and private agencies. He also signed an MoU with various industries. Currently, he is running four funded research projects in wireless charging of EV and UAV, which include DST SERB Core Research Grant and DST PURSE, Government of India. He is currently working as a Professor with the Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Kattankulathur Campus, Chennai, India. He has authored more than 110 research articles, which are published in international journals, including various IEEE Transactions. His research interests include power electronics converter topologies and controls for PV and EV applications, PWM techniques for power converters and adjustable speed drives, wireless power transfer, and smart grid. He can be contacted at email: bharatic@srmist.edu.in or bharatiraja@gmail.com.