UNIABUJA Journal of Engineering and Technology
https://ujet.uniabuja.edu.ng/
ISSN: 2714-3236 (Online); 2714-3228 (Print)
Volume 2, Issue 2, 2025; 180-187

Enhancing the Performance of Mabushi Solar Power Plant through Machine Learning Forecasting Models: Generation and Demand Sides Management

Simeon C. OGBU[1*], Musa T. ZARMAI[2], Eli J. BALA[3]

[1*,2]*Department of Mechanical Engineering, University of Abuja, Abuja, Nigeria*
[3]*Department of Petroleum & Energy Resources Engineering. African University of Science and Technology, Abuja*

[1*]chijykeogbu@yahoo.com, [2]musa.zarmai@uniabuja.edu.ng, [3]ebala@aust.edu.ng

Abstract

This study explores enhancing the performance of Mabushi Solar Power plant through machine learning forecasting models: Generation and Demand Side Management. The research aims to develop models that can accurately forecast daily solar power output using environmental and operational data. A dataset of 108 daily records within 2023 to part of 2024 from the Mabushi solar power plant was used, incorporating variables such as solar power output (as dependent variable), irradiance, temperature, wind speed, humidity as (independent variables), and energy contributions from batteries, generators, and the grid. The methods adopted are data collection and processing, model development and training. Three machine learning models Random Forest, Gradient Boosting, and Prophet were developed and evaluated. Among these, the Gradient Boosting model proved most effective in optimal forecasting, achieving a Mean Absolute Error (MAE) of 0.006 MW, Root Mean Square Error (RMSE) of 0.008 MW, and a high R^2 score of 0.999. Its residual errors were tightly distributed around zero, indicating strong reliability and accuracy. The model was then used to forecast solar power generation into the future, with average daily predicted power output of 0.677 MW, with values ranging between 0.390 MW and 1.002 MW, culminating in a total predicted annual generation of approximately 247.04 MW. By applying ML algorithms to 365 days of data, this research produced accurate and interpretable forecasts of daily solar power generation. The results provided actionable insights for load scheduling, maintenance planning, and energy storage management. The forecast makes the plant predictive and not reactive, showcasing its practical relevance for energy planning and operational decision-making through demand side management.

Keywords: *Solar Power Plant, Machine Learning, Enhancing Performance, Forecasting Models, Demand Side Management*.

1.0 Introduction

The transition towards sustainable energy sources has emerged as a pressing global imperative, driven by the dual challenges of climate change and fossil fuel depletion (International Energy Agency, 2022). More research work on renewable energy is pertinent to achieve Nigeria's 29% reduction in greenhouse gas emission by the year 2030 (NDC 3.0) (UNDP OCT 2025). Among the various renewable energy sources, solar energy stands out due to its abundance, environmental friendliness, and declining installation costs (Markvart & Castaner, 2003). For developing nations like Nigeria, where erratic power supply has been a persistent bottleneck to economic growth, the harnessing of solar energy presents an opportunity for infrastructural stability and sustainable development (World Bank, 2020). However, the efficiency of solar power plants is significantly influenced by environmental variables and load demands, which exhibit high degrees of uncertainty and variability. This challenge necessitates the integration of intelligent prediction and optimization frameworks to ensure reliability and system responsiveness (Benitez & Singh, 2025).

In recent years, machine learning (ML) has demonstrated immense potential in modeling and forecasting complex, nonlinear relationships within solar energy systems. Traditional statistical methods, though useful, often fall short in capturing the intricate interactions between multiple environmental factors such as solar irradiance, temperature, humidity, and cloud cover. As reported by (Sedai et al 2025), ML algorithms such as Support Vector Machines (SVM), Random Forests (RF), and Deep Neural Networks (DNN) have outperformed classical models in long-term solar power forecasting by offering higher accuracy and adaptability to dynamic conditions. This advancement signifies a paradigm shift in the design and operation of photovoltaic systems. especially in data-scarce or data-volatile regions like Mabushi, Abuja. Photovoltaic systems convert sunlight directly into electricity using semi-conductor based solar cells typically arranged in modules. The Mabushi solar power plant represents a strategic node within Nigeria's decentralized energy network which aids the responsibility of the National Control Centre (NCC) in monitoring, controlling and optimizing national electricity grid. Yet, despite its potential, it is often plagued by inconsistent output, low

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

forecasting precision. These inefficiencies can be attributed to the lack of real-time adaptive systems that can respond proactively to environmental and operational fluctuations. Integrating machine learning into this context is therefore not merely a technological upgrade but a necessary evolution toward smarter energy infrastructures. As noted by (Idogho et al 2024), ML-based solar forecasting models can improve decision-making in energy dispatch, battery storage management, and load balancing, thus enhancing overall system efficiency and resilience.

Moreover, the global scientific community has increasingly acknowledged the role of data fusion and ensemble learning in boosting the reliability of solar power forecasting systems. The work of (Dhablia et al 2024), underlines the critical importance of combining multiple ML models to mitigate the weaknesses of individual algorithms, leading to more robust and context-aware solar energy systems. This approach aligns with the peculiar needs of urban settlements like Mabushi, where environmental variability and infrastructural constraints necessitate agile and intelligent energy solutions.

Consequently, the integration of machine learning into the Mabushi solar power system is not only innovative but also strategically aligned with Nigeria's renewable energy goals and the broader Sustainable Development Goals (SDGs). It enables a shift from reactive maintenance to predictive analytics, from static design to dynamic adaptability. As such, this study seeks to explore and demonstrate how machine learning techniques can be effectively employed to enhance the performance of solar power plant in real-world settings, using Mabushi as a pilot case.

2.0 Materials and Methods

The successful design and implementation of a machine learning-based solar power forecasting framework requires a thoughtful selection of both hardware and software materials. The materials are grouped into two main categories: (1) data-related inputs, which constitute the backbone of the forecasting model, and (2) computational tools and platforms, which support the model development, training, and evaluation processes.

2.1. Historical and Real-time Data from Mabushi Solar Power Plant.

The foundational material used in this study is the dataset obtained from the Mabushi solar power plant. These datasets comprise time-stamped records of key environmental and operational parameters including solar irradiance, ambient temperature, panel voltage and current, and historical power output. As emphasized by (P. Raghuwanshi 2024), the reliability of forecasting models is highly dependent on the quality and resolution of the input data. The data was sourced over a period spanning multiple weeks to capture both diurnal (active in daytime) and seasonal variability.

Table 1: Model results showing variations in performance across the algorithms:

S/N	Model	MAE (MW)	RMSE (MW)	R^2 Score
1	Random Forest	0.070	0.093	0.873
2	Gradient Boosting	0.006	0.008	0.999
3	Prophet	0.162	0.200	0.419

From table 1 above, it is evident that the Gradient Boosting Regressor achieved the highest forecasting precision with the lowest RMSE and MAE, and an almost perfect R^2 score of 0.999. This implies that the model was nearly able to explain all the variability in the solar power output using the provided system data and environmental parameters. It was adopted for one year forecast of solar power generated.

2.1.2 Overview of Mabushi Solar Power plant.

Mabushi, located within the Federal Capital Territory (FCT) of Nigeria, is a rapidly urbanizing district positioned at approximately latitude 9.062°N and longitude 7.440°E. It sits at an elevation of around 477 meters above sea level and experiences a tropical savanna climate characterized by distinct wet and dry seasons. Solar irradiance in this region typically ranges between 4.5 and 6.5 kWh/m^2/day annually, positioning it as a viable site for solar photovoltaic (PV) power generation. These geographic and climatic features provide the foundational rationale for selecting Mabushi as a suitable location for studying the integration of intelligent solar energy solutions. Figure 1 shows the aerial view of 1.54MWp/2.28MWh Storage Microgrid in Mabush Abuja.

It is part of a government-supported clean energy initiative aimed at supplementing the inconsistent grid supply with renewable alternatives. The system comprises a ground-mounted photovoltaic array of 3968 PV panels with a cumulative capacity of 1.54MWp coupled with a hybrid inverter of 1.17MW and a 2.28MWh Tesla's Power pack for energy storage. The installation is designed to serve a cluster of public facilities, including administrative offices and residential quarters.

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

Figure 1: Aerial view of the 1.54 PV / 2.28 MWh storage microgrid in Mabushi, Abuja

2.2. Meteorological Data Sources

To complement system-generated data, publicly available meteorological datasets particularly from the NASA POWER (which has wider coverage and better for academic research) database were integrated. These datasets provided additional variables such as solar irradiance, relative humidity, wind speed, and ambient temperature. These inputs have been validated in the literature as key predictors for solar radiation modeling (P. Setiawati et al 2025), (A. Sedai et al 2023). Furthermore, the geographical coordinates of Mabushi (Lat. 9.062°N, Long. 7.44°E) were used to align satellite data with the physical location of the solar plant.

2.3. Machine Learning Libraries and Programming Environment

For model development, Python was selected as the primary programming language due to its simplicity, extensive ecosystem of machine learning libraries. It provides a versatile and efficient environment for every state of the ML workflow from data preprocessing to model deployment. Key libraries used include:

- **Scikit-learn**: Employed for training classical models such as Support Vector Machine (SVM), Decision Trees, and Random Forests. This library also supports cross-validation, performance metrics, and feature engineering tasks.
- **TensorFlow and Keras**: Utilized for implementing Deep Neural Networks (DNNs), which are particularly effective for capturing complex non-linearity in solar power data (Idogho et al 2024)
- **Pandas and NumPy**: Used for data handling, cleaning, and numerical computation during processioning.
- **Matplotlib and Sea-born**: For data visualization and performance reporting.

These tools collectively enabled the seamless transformation of raw data into structured datasets and supported model training and evaluation in an efficient, reproducible manner.

2.4 Data Collection and Processing

The first stage involves the collection of historical data from the Mabushi solar power plant facility engineers and NASA POWER database. Key variables included solar irradiance, wind speed, PV output power, ambient temperature, time of day, and system voltage. These variables are selected based on their known influence on solar energy generation.

2.4.1 How the data was collected and processed

The dataset was collected on daily basis from the Ministry of works engineer's log, recording the average daily power generated or used from solar, battery, generator, and national grid. The datasets undergo series of processing steps including outlier detection, normalization, timestamp alignment, and missing data imputation to ensure quality and consistency. The cleaned datasets are then divided into two, 20% for training and 80% for testing. The 20% is used for training the machine learning models. The 80% is reserved for testing and validation, simulating real-world prediction tasks using unseen data.

2.4.2 Sensitivity Analysis

Sensitivity means how small changes in parameters affect output power.

$$S_X = \frac{\partial P}{\partial x} * \frac{X}{P} \qquad (1)$$

where X is a parameter, P is solar output power and S is Sensitivity.

1. Sensitivity to Solar Irradiance

$SG = \frac{\partial P}{\partial G} * \frac{G}{P} = 1$ (2)

Where G is solar irradiance

Meaning, power output changes directly proportional to irradiance

2. Sensitivity to panel efficiency

   $S\eta = 1$ (3)

   A 1% increase in panel efficiency gives 1% increase in power output.

3. Sensitivity to Temperature

   $S_T = \beta (Tc - T_{ref})$ (4)

   Higher temperature reduces power output.

For Mabushi Solar Power Plant

- Solar irradiance variability has the largest influence.
- Temperature effects slightly reduce efficiency due to high ambient heat.
- Improving panel efficiency or performance ratio significantly boosts output.

Sensitivity analysis using parametric equations shows that the most critical parameters affecting Solar Plant output are irradiance, panel efficiency, and temperature.

2.5 Model Development and Training

Selected machine learning algorithms are developed using the training datasets. Algorithms under consideration include Gradient Boosting Regressor (GBR), Random Forest Regression, and Facebook Prophet each chosen for their suitability in modeling nonlinear, time-dependent solar datasets. The training process involves feeding the model with known inputs (e.g., time, irradiance, temperature, humidity, wind speed) and expected outputs (e.g., solar power generation), allowing the model to learn the underlying relationships through iterative optimization.

2.5.1 Model Selection

The selection of machine learning models was guided by three principal criteria:

✔ **Suitability for Time-Series Forecasting**: The models had to effectively handle temporal patterns, including daily and seasonal trends, in solar irradiance and output power.

✔ **Robustness in Handling Noisy and Incomplete Data**: Given the real-world nature of the datasets (NASA POWER and facility engineer logs), tolerance to noise and adaptability to missing values was critical.

✔ **Interpret-ability and Ease of Integration**: Preference was given to models that offered transparency in their prediction mechanisms and could be easily implemented in Python-based environments.

Based on these criteria, the following models were selected:

Gradient Boosting Regressor: Machine Learning algorithm that predicts continuous values by combing many small decision trees, each one correcting the errors of the previous ones. It handles non-linear relationships, has strong predictive performance, works well with mixed features and widely used for real-world datasets like solar power plant datasets.

Facebook Prophet: A time-series forecasting model developed by Meta, known for its robustness in handling seasonality, holidays, and trend shifts. Prophet is particularly useful in energy forecasting due to its ability to decompose time series into additive components of trend, seasonality, and holiday effects. It requires minimal feature engineering and is ideal for univariate forecasting tasks using timestamps and target values alone.

Random Forest Regression: A tree-based ensemble learning method that aggregates predictions from multiple decision trees. Its strength lies in capturing nonlinear relationships between environmental variables (irradiance, temperature) and power output while resisting over fitting. It is also robust to outliers and works well with mixed-type features.

3.0 Results and Discussions

The process of training machine learning models to forecast solar power generation at the Mabushi Solar Power plant was carefully structured to ensure model reliability, interpret-ability, and predictive accuracy. The models were developed using the complete datasets comprising 108 daily records of solar power performance parameters. These parameters included solar power output, battery power, generator power, grid power, irradiance, ambient temperature, relative humidity, and wind speed. The controlling variables (independent) are solar irradiance, ambient temperature, relative humidity, and wind speed. The target variable for forecasting was the daily solar power output in megawatts (MW).

Three machine learning models were trained and compared in this study:

1. Random Forest Regressor (RF)
2. Gradient Boosting Regressor (GBR)
3. Facebook Prophet Time-Series Model

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

Each model was trained using the full datasets, after applying appropriate processioning steps such as date time parsing, feature engineering (e.g., cyclical encoding of date features), and lagged target generation. The datasets were split into training and testing segments (typically 20/80) to facilitate robust evaluation. For consistency and reproducibility, hyper parameters were initially set to default values, with the Gradient Boosting model later fine-tuned to exploit its higher predictive potential. The training phase involved evaluating the performance of each model on unseen test data using the following metrics: Figure 2 below shows the comparisons of the three (3) models performances.

- Root Mean Squared Error (RMSE) – for penalizing large deviations
- Mean Absolute Error (MAE) – for measuring average magnitude of forecast errors
- R-squared (R^2) – for evaluating the model's explanatory power

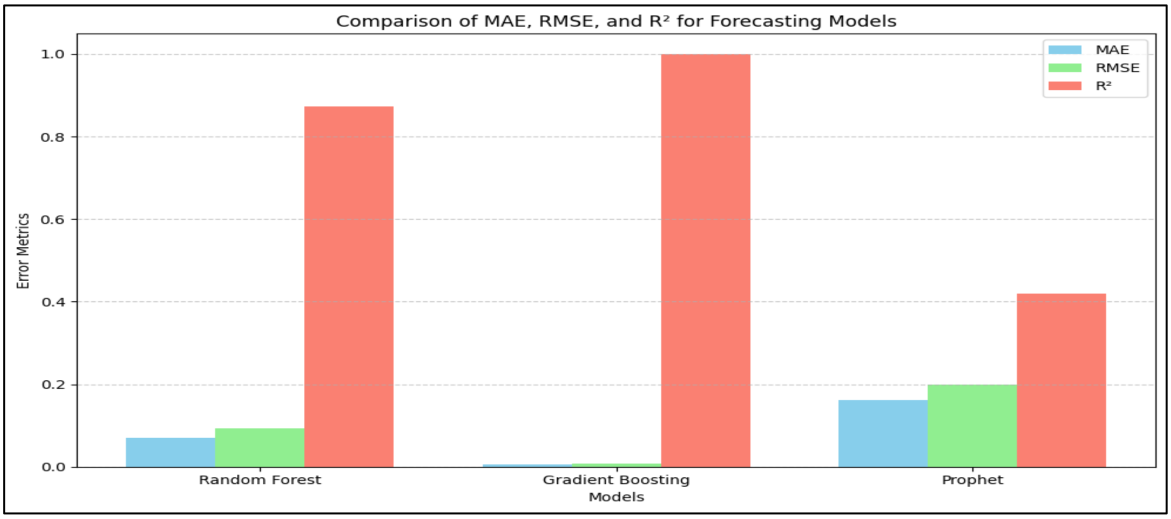

Figure 2: Bar Chart Comparing MAE, RMSE, and R² across Models

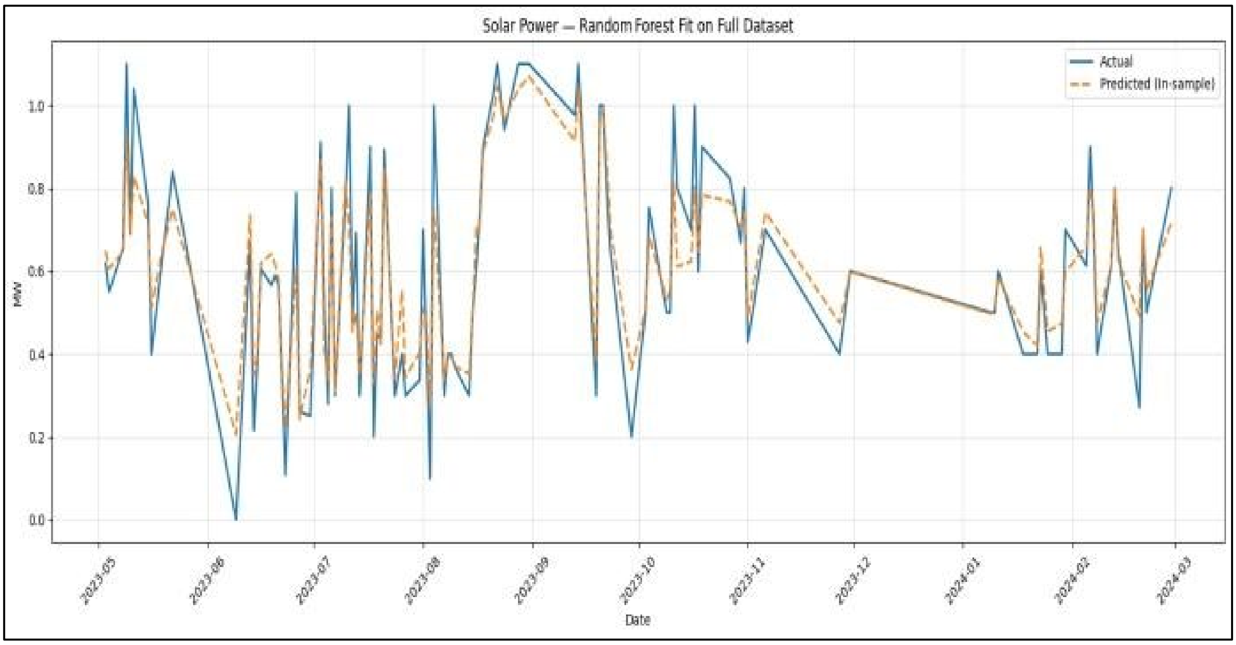

Figure 3: Actual vs Predicted Solar Power output (Random Forest)

Figure 3 above, visualizes how well the Random Forest model predicted the solar power output relative to the actual values. The prediction follows the trend of the actual data with moderate deviation.

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

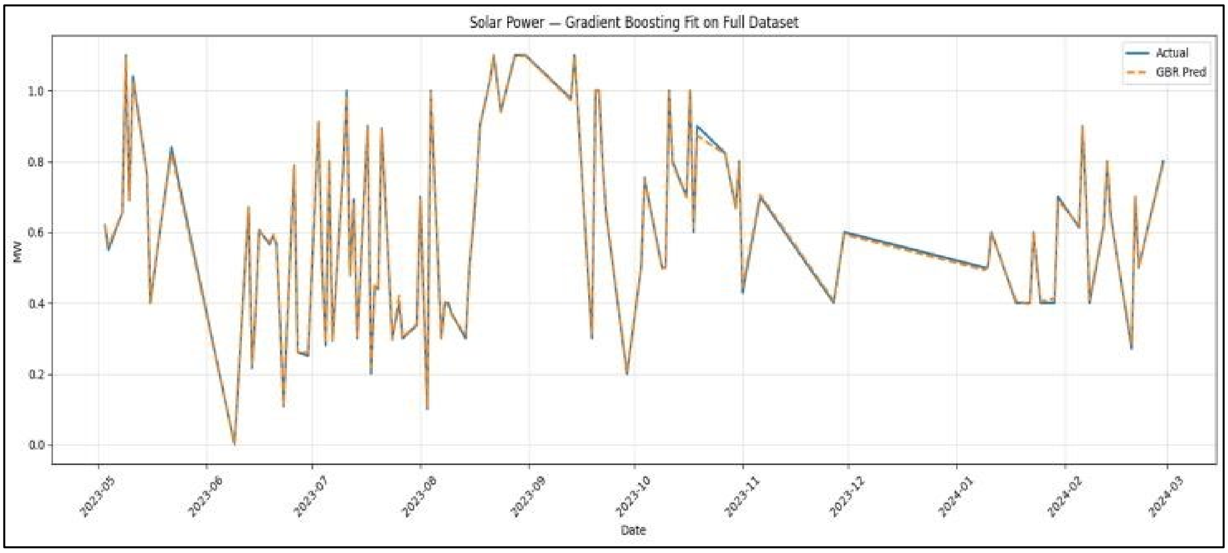

Figure 4: Actual vs Predicted Solar Power output (Gradient Boosting)

Figure 4, shows a near-perfect alignment between the predicted and actual solar power outputs, illustrating the Gradient Boosting model's superior fit.

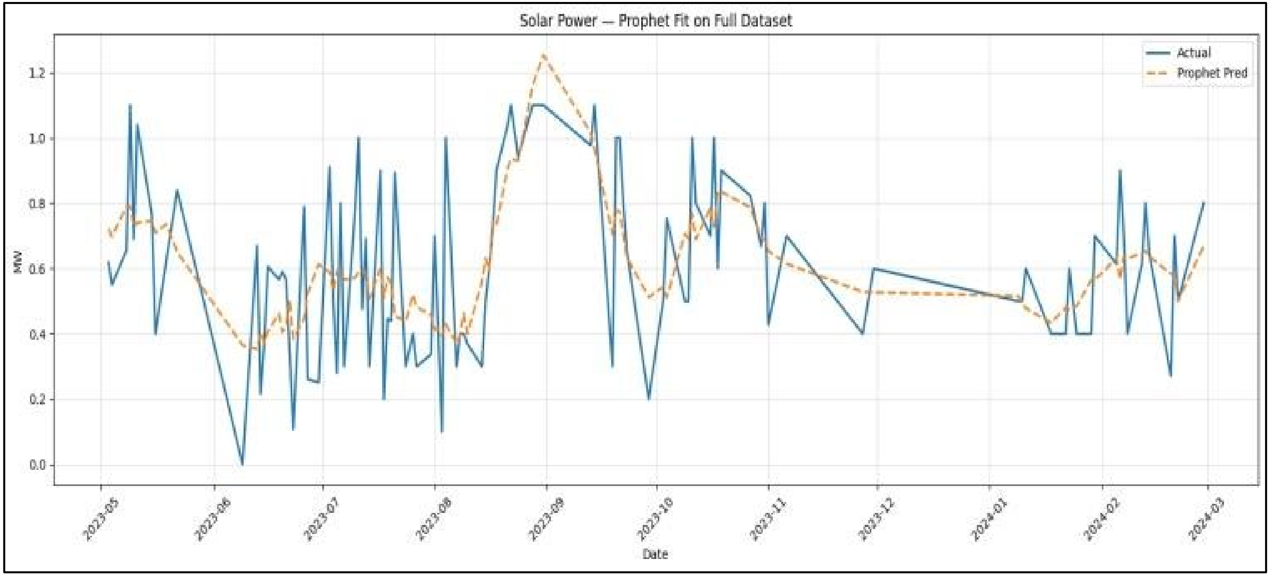

Figure 5: Actual vs Predicted Solar Power output (Prophet)

The Prophet model's performance in figure 5, deviates more significantly from the actual values, reflecting its lower R² and higher error values. Figure 6 below shows the forecasted daily solar power output for next 365 days using the best performed model, Gradient Boosting Regressor.

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

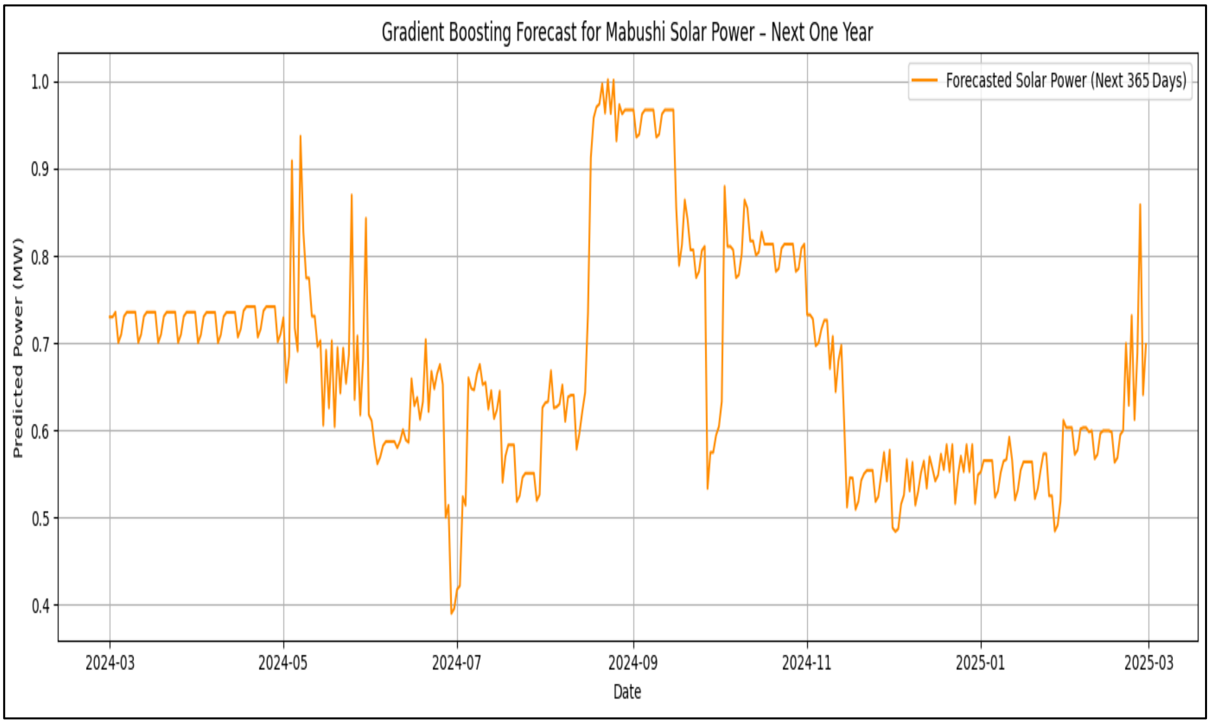

Figure 6: Forecasted Daily Solar Power Output for the Next 365 Days using GBR

4.0 Conclusion

The analysis of historical and forecasted data provided a clear understanding of the plant's generation dynamics. The Mabushi plant exhibited an average daily predicted power output of 0.677 MW, with values ranging between 0.390 MW and 1.002 MW, culminating in a total predicted annual generation of approximately 247.04 MW. These results demonstrate that the plant maintains consistent productivity within expected meteorological variations. This finding underscores the plant's capability to reliably supply renewable energy to the Abuja urban corridor, while also highlighting areas for operational fine-tuning based on seasonal demand and solar resource availability.

By applying ML algorithms to 365 days of data, this research produced accurate and interpretable forecasts of daily solar power generation. The results provided actionable insights for load scheduling, maintenance planning, and energy storage management. The predictive patterns enable plant operators to anticipate seasonal dips, plan maintenance during low-output periods, and maximize energy dispatch during high-generation days. This objective was therefore fully achieved, demonstrating that machine learning forecasting serves as an indispensable tool for strategic energy management and performance optimization.

References

Appiah, A. Y., Oduro, S.K., & Adjei, A.K. (2022). Machine learning applications in renewable energy: A review. *Renewable and Sustainable Energy Reviews*, 167, 1-18.

Dhablia, A., & Godil, S. A (2024). Machine Learning Integration for Enhanced Solar Power Generation Forecasting. *School of Business and Management, CHRIST (Deemed to be University)*, Bangalore Campus.

Dhablia, A., & Godil, S. A. (2023). Artificial intelligence in forecasting solar radiation: A Review. *Indonesian Journal of Electrical Engineering and Computer Science*, 30(2), 924–930.

Energy Commission of Nigeria. (2015). National renewable energy and energy efficiency policy (NREEEP). Federal Government of Nigeria.

Femia, N. A., Petrone, G., Spagnuolo, G., & Vitelli, M. (2008). Prediction techniques for maximum power point tracking of photovoltaic systems. *IEEE Transactions on Industrial Electronics*, 55(7), 2629–2638.

Hossain, M., Madlool, N., & Rahman, M. (2021). Artificial intelligence for renewable energy systems. Forecasting, control, and optimization. Energy AI, 5, 100-112.

Idogho, A.E., Agboola, A.O., & Akinyele, J.O. (2020). Hybrid solar-diesel power systems using machine learning-based forecasting models for Nigerian regions. International Journal of Renewable Energy Research, 10(3), 1097-1105.

Iloeje. N.B. (2021). The current status and future direction of solar PV in Nigeria. *International Journal of Renewable Energy Research*, 8(1), 83–92.

Ozoegwu, M., & Ugwuoke, R. O. (2020). Artificial intelligence techniques in smart grid and renewable energy systems: A review. *International Journal of Energy Economics and Policy* 10(4), 243–250.

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16

Raghuwanshi, A. (2020). A survey on artificial intelligence-based solar power prediction using machine learning models. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 7(5), 369–375.

Raghuwanshi, P. (2024). Predictive modeling and forecasting of solar power generation using machine learning techniques. *Journal of Advances and Scholarly Researches in Allied Education*, 21(5), 207–216.

Salim, A., & Zhang, T. (2020). Hybrid Renewable Energy Systems in Off-Grid Areas: A Review of Machine Learning Applications. *Renewable and Sustainable Energy Reviews*, 119, 109–132,

Setiawati, A., Hasibuan, T., Fadhil, A., & and Ramdhani, M.S. (2023). Comparison of machine learning methods for solar power forecasting in tropical region. *Indonesian Journal of Electrical Engineering and Computer Science*, 30(2), 924–930.

Setiawati, D., Istanto, I., & Mustika, I.W. (2022). The Influence of Solar Panel Tilt Angle on Solar Power Plant Performance in the Tropics. *IOP Conference Series: Earth and Environmental Science*, 1065(1), 1–7.

Setiawati, P. Karno, A.S.B., Hastomo, W., Sestri, E., Kasoni, D., Arif, D., & Razi, F. (2025) Predicting Solar Power Generation: A Machine Learning Approach for Grid Stability and Efficiency. *Pilar: Jurnal Informatika*, 21(1), 34–45.

Taylor, S. & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1) 37–45.

Umar, A.M., Oyedepo, S.O., & Yusuf, A. A. (2018). Solar energy development in Nigeria: Current status and challenges. *Nigerian Journal of Technology*, 37(4) 1132–1140.

Yang, Y., Li, H., & Xie, M. (2022). Ensemble learning approaches for solar power forecasting: A review. *Applied Energy*, 323, 119-564.

Zhang, Z., Wu, X., & Wang, C. (2022). Hybrid machine learning and numerical weather prediction for wind power forecasting. *Renewable Energy*, 195, 1291–1303.

Zhang, S., Hu, R., & Wang, J. (2023). Machine Learning for Lithium-ion battery safety management and thermal runaway prediction. Journal of Power Sources, 532, 231-291.

Received: 14-02-2026 / Accepted: 05-03-2026 / Published: 21-05-2026

https://doi.org/10.70118/ujet.2026.0302.16