International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-14 Issue-5, January 2026

Advanced Ensemble Machine Learning for Photovoltaic Production Forecasting in Tropical Microgrids: Application on the Katsepy Site, Madagascar

Linda Christelle Mevalaza, Andrianajaina Todizara, Rakotoarisoa Armand Jean Claude, Hervé Mangel, Jean Nirinarison Razafinjaka

Abstract*: This article presents a Machine Learning ensemble model for photovoltaic (PV) production forecasting, developed to address energy management challenges in rural tropical microgrid environments. The site studied is located in Katsepy, Madagascar, where the reliability of energy forecasts remains a significant challenge due to high climate variability. The main objective is to design a more accurate and adaptable forecasting tool to enable adequate energy planning and system operation. The methodology is based on historical photovoltaic production data and meteorological data collected between 2005 and 2023 from the PVGIS online platform. Several regression algorithms were evaluated, including Random Forests, Bagging, and Gradient Boosting, to identify the models best suited to the local context. Among these, Gradient Boosting showed the best performance according to RMSE, MAE, MAPE and R^2 measurements, followed closely by Random Forest and Bagging. The experimental process consists of two stages: first, validation using actual 2023 data, and then forward-looking forecasts for 2024 incorporating real temperature data. To improve accuracy and robustness, a Stacking ensemble model was constructed, combining the three best-performing algorithms as base estimators and the Extra Trees Regressor as the meta-model. This ensemble approach consistently outperformed the individual models and provided realistic production estimates for 2024, indicating a moderate decline in photovoltaic production compared to 2023, driven by observed climate variations. The proposed forecasting framework provides a solid foundation for future work on optimal energy management and fault diagnosis in the Katsepy microgrid system, with great potential for adaptation to other tropical coastal regions.

Keywords*: Forecasting, Photovoltaic Production, Tropical Microgrid, Machine Learning Model, Climate Variation.

***Nomenclature*:**
PV: Photovoltaic
PVGIS: Photovoltaic Geographical Information System
RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
R^2: Coefficient of Determination
MLP: Multi-Layer Perceptron
SVM: Support Vector Machine
KNN: K-Nearest Neighbours
SVR: Support Vector Regression
LSTM: Long Short-Term Memory
CNN: Convolutional Neural Network

Manuscript received on 20 October 2025 | First Revised Manuscript received on 07 November 2025 | Second Revised Manuscript received on 16 December 2025 | Manuscript Accepted on 15 January 2026 | Manuscript published on 30 January 2026.
*Correspondence Author(s)

Linda Christelle Mevalaza*, Student, Department of Electrical and Automation Engineering, Doctoral School of Renewable Energies and Environment, University of Antsiranana, Madagascar. Email ID: lindachristellemevalaza@gmail.com. ORCID ID: 0009-0000-4727-0989.

Dr. Andrianajaina Todizara, Researcher, Department of Electrical and Electronic Engineering, Higher Polytechnic School, University of Antsiranana, Maagascar. Email ID: todizara.andrianajaina@gmail.com. ORCID ID: 0000-0003-1240-8524

Dr. Rakotoarisoa Armand Jean Claude, Professor, Department of Electrical and Electronic Engineering, University of Antsiranana, Madagascar. Email ID: rajeanclaude@gmail.com

Hervé Mangel, Assistant Professor, Department of Electrical and Industrial Computing Engineering, Dupuy de Lôme, Research Institute (UMR CNRS 6027 IRDL). Email ID: herve.mangel@univ-brest.fr.

Jean Nirinarison Razafinjaka, Professor, Department of Electrical and Electronic Engineering, Higher Polytechnic School, Antsiranana, Madagascar. Email ID: razafinjaka@yahoo.fr, ORCID ID: 0009-0003-6450-4558.

I. INTRODUCTION

The integration of renewable energy into microgrids poses significant challenges in developing tropical regions, underscoring the need to predict PV output accurately. This study addresses these challenges using advanced ensemble Machine Learning techniques at the Katsepy site in Mahajanga, Madagascar. The forecasting framework developed aims to serve as a fundamental tool for intelligent energy management and fault diagnosis applications in this specific photovoltaic microgrid. Madagascar has strong solar potential, with over 2,800 hours of sunshine per year [1]. However, the country still lacks sophisticated forecasting tools capable of managing the variability of tropical climates. The Katsepy region (15°45'59.99'' S, 46°13'60.00'' E) is an example of a tropical coastal climate characterised by substantial seasonal variations. Consequently, these climate changes will significantly impact photovoltaic energy production [2].

This research presents a two-phase approach:

- Rigorous validation using 2023 data as an independent dataset through several ensemble methods.
- Forward-looking forecasts for 2024, integrating real temperature measurements with modelled weather scenarios.

The methodology is specifically designed for the Katsepy context while maintaining scientific rigour, and may be adapted to similar tropical coastal sites that lack sufficient resources.

Beyond traditional forecasting methods, this work applies ensemble learning techniques, including Random Forests, Bagging, and Gradient Boosting, with the latter demonstrating superior predictive performance. These techniques are integrated into a Stacked Regressor using an Extra Trees Regressor as a meta-model. Although these methods have been described in recent literature, they remain largely unexplored in tropical photovoltaic forecasting. Furthermore, combining PVGIS data with local temperature measurements constitutes an innovative hybrid approach for constrained environments.

There are several forecasting methods; however, recent advances in Machine Learning have significantly improved PV forecasting compared to traditional time series models, particularly through ensemble techniques such as Gradient Boosting, which offer a good compromise between accuracy and computational efficiency [3]. Deep Learning and Machine Learning applications have also advanced, particularly in modelling environmental factors such as solar radiation and temperature [3].

Ensemble methods, notably Random Forest and Gradient Boosting, are increasingly used in energy forecasting, making them effective for various PV configurations [4]. Stacked Regressors are particularly promising. For example, a Stacked model integrating LASSO, Random Forest, MLP, SVM, and XGBoost with Bayesian Regression as the meta-learner outperformed other models on Thai data [5].

Similarly, a Stack-ETR model combining Random Forest, XGBoost and AdaBoost with an Extra Trees meta-learner reduced the RMSE and MAE for monocrystalline thin-film PV systems [6].

Despite these scientific advances, few studies apply ensemble methods in developing tropical regions, where data scarcity and resource constraints further complicate matters. The integration of real temperature measurements into ensemble models for annual forecasts with seasonal scenarios remains underexplored, particularly in tropical Africa.

II. METHODOLOGY

A. Data Sources and Site Characterisation

This study is based on two primary data sources:

i. *Historical PV Data (2005-2023):* hourly meteorological data and simulated production data obtained from the Photovoltaic Geographic Information System (PVGIS). The free online PVGIS application is an excellent simulation tool developed by the European Commission and widely recognised for solar potential assessment and PV performance modelling [7].

ii. *Temperature Data (2024):* daily air temperature measurements collected from AccuWeather. The private American company AccuWeather is a commercial meteorological service provider [8].

The Katsepy site (-15.7667°S, 46.2333°E) is a tropical savanna (Köppen Aw) located on the coast at an altitude of 7 meters [9]. The dataset comprises 166535 hourly observations over 19 years, including solar irradiation, sun elevation, air temperature, wind speed and energy production (Wh). Located in northwestern Madagascar, Katsepy is at the mouth of the Betsiboka River and is mainly accessible by boat from Mahajanga [10].

B. Advanced Data Preprocessing and Feature Engineering

i. *Preprocessing Pipeline*

- **Temporal Feature Extraction**: Cyclical encoding (sine/cosine) of temporal variables; multi-scale features (hourly, daily, monthly, seasonal).
- **Year Normalisation** :

$$\text{normalized}_{\text{year}} = \frac{\text{year - min}_{\text{year}}}{\text{max}_{\text{year}} \text{ - min}_{\text{year}}} \quad \dots \quad (1)$$

- **Meteorological Scenarios for 2024** :
- Four scenarios: Conservative (historical weighted avg), Realistic (+0.2σ), Trend (+0.3σ), Optimistic (+0.6σ).
- Weighted historical averages (2019–2023 emphasised) integrated with real 2024 temperature data.
- **Data Quality**: Missing values imputed (median), outliers detected (IQR), and temporal consistency validated.

ii. *Ensemble Machine Learning Framework Data Splitting:*

a. *Case 1 (2023 Validation)* :
- Development set (2005–2022): Model training/hyperparameter tuning (TimeSeriesSplit cross-validation).
- Validation Set (2023): a holdout set for unbiased evaluation.

b. *Case 2 (2024 Prediction)* :
- Training Set (2005–2022) with temporal weighting (recent years: 2.0×; mid-recent: 1.5×; older: 1.0×).
- Validation on 2023 holdout.
- **Data Standardization:** Numerical features scaled (StandardScaler) using training data only to prevent leakage.
- **Model Categories**
- **Tree-Based**: Random Forest, Extra Trees, Gradient Boosting (XGBoost, LightGBM, CatBoost), AdaBoost.
- **Linear/Neighbour-Based**: Linear Regression, Lasso, Ridge, KNN.
- **Advanced Ensembles**: Bagging, Voting Regressor, Stacking Regressor, SVR, MLP.

Hyperparameter optimization (GridSearchCV) was applied to minimize RMSE.

C. Model Configuration and Ensemble Architecture

Table I shows the configuration of the three best individual models:

Table I: Configuration of the Three Best Individual Models

Model	Parameters	Values
Gradient Boosting	n_estimators	300
	max_depth	8
	learning_rate	0.05
	subsample	0.8
Random Forest	n_estimators	300
	max_depth	20
	min_samples_leaf	1
	min_samples_split	2
Bagging	(default parameters)	-

Two sophisticated ensemble approaches were implemented:

Retrieval Number: 100.1/ijrte.D830914041125
DOI: 10.35940/ijrte.D8309.14050126
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-14 Issue-5, January 2026

- **Voting Regressor:** A simple averaging approach combining the predictions of the three best-performing individual models
- **Stacked Regressor:** A meta-learning approach using Extra Trees Regressor as the final estimator

Table II: Extra Trees Regressor Parameter

Model	Parameters	Values
Extra Trees Regressor	n_estimators	300
	max_depth	15
	min_samples_leaf	5
	min_samples_split	10
	bootstrap	True

D. Performance Metrics

Model evaluation used several complementary metrics:

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \dots \quad (2)$$

Mean Squared Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|^2 \quad \dots \quad (3)$$

- Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y_i})^2} \quad \dots \quad (4)$$

- Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \quad \dots \quad (5)$$

Where n, y_i, $\hat{y}_i$ et $\overline{y_i}$ represent respectively the number of observations, the actual or observed value, the predicted value and the mean of the expected value at time i.

III. RESULTS AND ANALYSIS

A. Phase I: 2023 Validation Results

Validation on 2023 data demonstrated exceptional performance across all models:

Table III: Individual Model Performance

Model	RMSE [Wh]	MAE [Wh]	R²	MAPE [%]
Gradient Boosting	197.47	78.42	1.00	0.89
Random Forest	260.84	112.34	1.00	0.79
Bagging	292.71	127.70	1.00	0.89

Table IV: Ensemble Model Performance

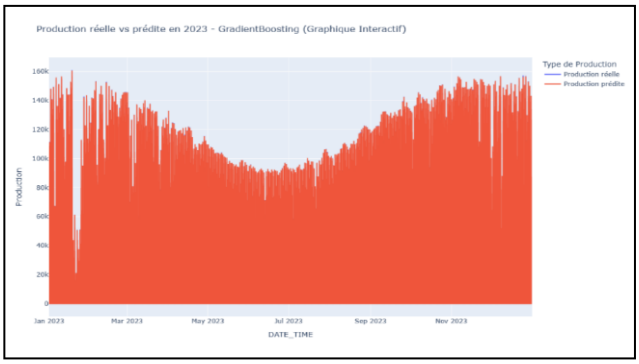
Model	RMSE [Wh]	MAE [Wh]	R²	MAPE [%]
Stacking Regressor	231.55	100.27	1.00	0.84
Voting Regressor	241.07	110.94	1.00	2.28

B. Performance Analysis and Overfitting Assessment

Exceptional R^2 values (>0.999) should be interpreted with caution to avoid any risk of overfitting.
Our rigorous methodology validates these results:

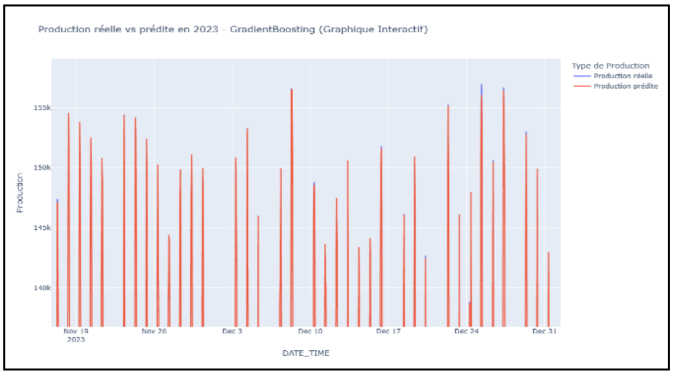
- Complete temporal separation, with the year 2023 used exclusively as an independent validation dataset;
- Cross-validation of the time series to prevent any data leakage;
- Consistent performance across several statistical indicators.

The high R^2 indicates a strong physical correlation between solar irradiance and PV production (p-value = 0.9992) in a tropical environment with stable weather.



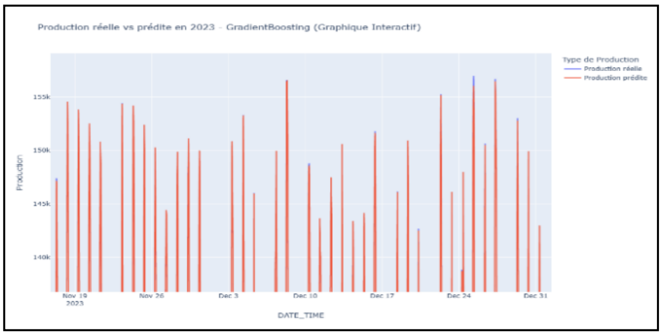
[Fig.1: Annual Comparison of PV Production Between Actual and Predicted Production]

The prediction curve in Figure 1 closely follows the actual production curve, confirming the Gradient Boosting model's excellent performance in predicting PV production.



[Fig.2: Correlation Graph Between Actual PV Production and Predicted PV Production]

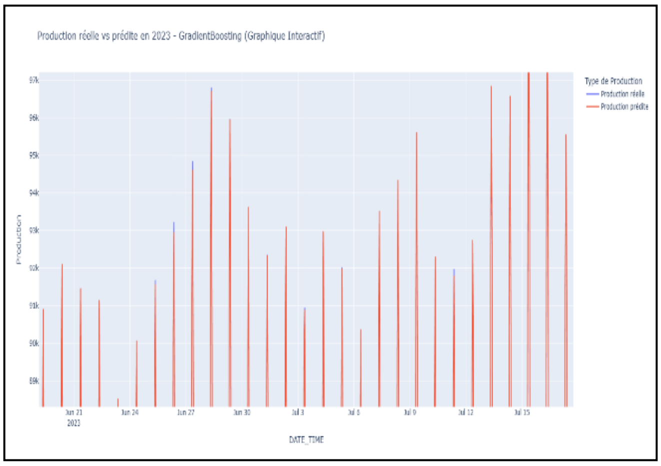
Figure 2 shows a strong positive linear correlation between actual and predicted PV production for 2023. This is evident from the scatter plot, which shows an upward trend from left to right. Thus, actual PV production and predicted production increase together, reinforcing the model's performance.



[Fig.3: Zooms on the Period of High PV Production, from November 19 to December 31, 2023]

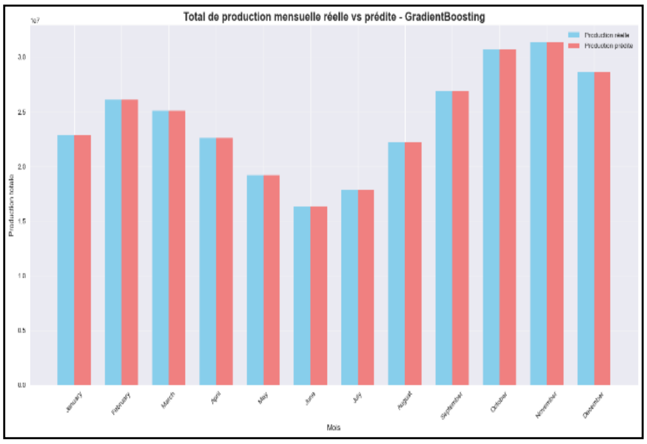
Retrieval Number: 100.1/ijrte.D830914041125
DOI: 10.35940/ijrte.D8309.14050126
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

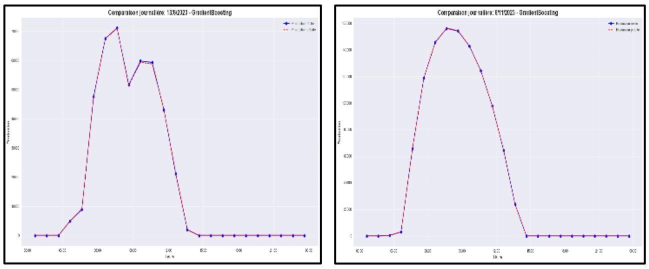


[Fig.4: Zoom on the Period of Low PV Production, from June 19 to July 17, 2023]

Figures 3 and 4 accurately reproduce seasonal trends, with some minor deviations due to unpredictable weather events. Thus, the model captures the overall dynamics while remaining sensitive to random and localized variations.



[Fig.5: Monthly Comparison of PV Production between Actual and Predicted Production]



[Fig.6: Daily PV Production Comparison on June 13 and November 6, 2023]

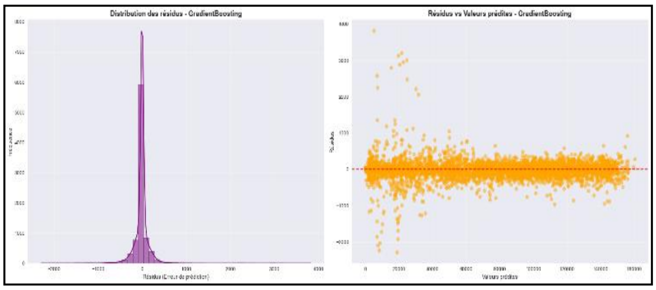
Monthly Performance Analysis (Fig. 5): The comparison between actual PV production and Gradient Boosting model predictions presented in Figure 5 and Table V for 2023 can be summarized as follows:

- Total Actual PV Production = 289 880 940 Wh
- Total Predicted PV Production ≃ 289 957 930 Wh
- Total Error: -76 990 Wh, or +0.026%

Table V: Monthly PV Production Comparison

Month	Real Production [Wh]	Predict Production [Wh]	Error [Wh]
January	22 861 650	22 861 425.89	224.11
February	26 081 960	26 081 181.30	778.70
March	25 112 642	25 113 490.88	-848.88
April	22 599 448	22 591 957.12	7 490.88
May	19 225 360	19 228 013.52	-2 653.52
June	16 337 014	16 346 206.99	-9 192.99
July	17 853 324	17 866 103.36	-12 779.36
August	22 212 592	22 227 132.25	-14 540.25
September	26 878 594	26 895 188.17	-16 594.17
October	30 743 388	30 745 695.21	-2 307.21
November	31 325 604	31 345 609.09	-20 005.09
December	28 649 364	28 655 926.25	-6 562.25
Total Annuel	289 880 940	289 957 930.03	-76 990.03

i. *Daily Analysis:* For June 13 and November 5 (Fig. 6), the actual and predicted PV production curves overlap almost perfectly, confirming a strong concordance and minimal bias. The model successfully reproduces the diurnal cycle, which begins with a gradual increase, peaks, and then drops to zero. This demonstrates its ability to capture intra-daily dynamics.



[Fig.7: A Error Distribution; b- Residuals Versus Predicted Values]

ii. *Residual Analysis:* The residuals of the Gradient Boosting model are centred around zero, follow a normal distribution, have constant variance, and exhibit few extreme values. This confirms the model's robustness and the absence of significant bias.

C. Phase II: 2024 PV Production Forecasting with Real Temperature Integration

Assessment of the Impact of Temperature:

Real 2024 temperature data revealed significant climatological changes:

- 2024 Average Temperature: 28.3°C
- Historical Average: 27.0°C
- Temperature Anomaly: +1.3°C (potentially reducing PV efficiency by 0.4-0.6%)

Table VI: Model Performance for 2024 Training

Model	MAE [Wh]	RMSE [Wh]	R^2
Stacking Regressor	**77.7**	**196.6**	**1,00**
Voting Regressor	96.8	229.8	1.00
Gradient Boosting	79.6	197.7	1.00
Random Forest	115.6	269.1	1.00
Bagging	117.5	272	1.00

Using already-optimised parameters, the Stacking model slightly outperforms

Retrieval Number: 100.1/ijrte.D830914041125
DOI: 10.35940/ijrte.D8309.14050126
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Gradient Boosting, as the meta-model learns to correct specific errors in the base models — Gradient Boosting, Random Forest, and Bagging — after they are trained.

Table VII: 2024 PV Production Forecasting Results

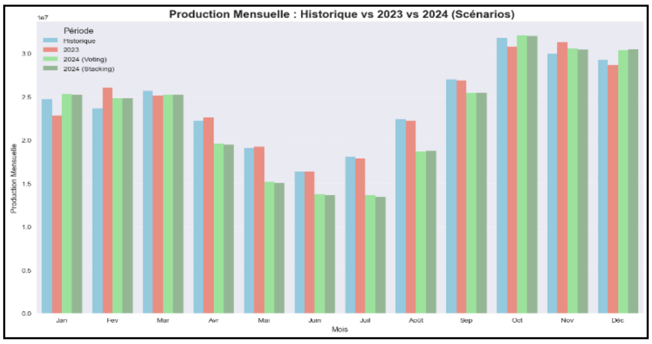
Scenario	Ensemble Model		Comparison Percentage	
	Stacking [Wh]	Voting [Wh]	Stacking vs 2023 [%]	Stacking vs Historical [%]
Optimistic	278 980 640	279 422 203	-3.8%	-3.8%
Realistic	**273 969 184**	**274 601 288**	**-5.5%**	**-5.5%**
Conservative	270 309 410	271 060 666	-6.8%	-6.8%
Trend	274 988 332	275 643 584	-5.1%	-5.2%

The selected scenario is realistic because it is considered the closest to reality. With this scenario and the Stacking model, the predicted PV production for 2024 is approximately 273,969,184 Wh. This decrease matches the observed temperature anomaly (+1.3°C) and confirms the physical validity of our models.

D. Seasonal Analysis

Table VIII: Monthly Production Comparison

Month	Historical Avg [Wh]	2023 Actual [Wh]	2024 Voting [Wh]	2024 Stacking [Wh]
January	24 704 737	22 861 650	25 290 318	25 267 135
February	23 650 334	26 081 960	24 798 759	24 805 692
March	25 664 440	25 112 642	25 257 876	25 222 912
April	22 242 634	22 599 448	19 535 826	19 448 687
May	19 040 185	19 225 360	15 188 027	15 041 855
June	16 349 165	16 337 014	13 779 302	13 670 274
July	18 072 266	17 853 324	13 637 867	13 430 043
August	22 390 786	22 212 592	18 670 424	18 724 688
September	26 97 3957	26 878 594	25 414 026	25 448 762
October	31 781 975	30 743 388	32 073 974	32 021 292
November	29 940 028	31 325 604	30 552 968	30 420 458
December	29 251 403	28 649 364	30 401 922	30 467 388
ANNUAL TOTAL	**290 061 910**	**289 880 940**	**274 601 288**	**273 969 184**



[Fig. 8: Monthly Comparison between Historical PV Production, Production in 2023 and Production in 2024 with Voting and Stacking]

According to Table VIII and Figure 8, the monthly PV production forecast for 2024 (Voting and Stacking) remains lower than historical levels and those of 2023. This decline reflects the impact of rising temperatures. Seasonal analysis highlights three distinct production patterns:

- High season (October–December): periods of maximum solar energy production, corresponding to high irradiation and favourable weather conditions.
- Low season (May–August): period of minimum PV production, linked to lower solar irradiation.
- Transition periods: intermediate phases with gradual variations that reflect the gradual climatic transitions between dry and wet conditions.

IV. DISCUSSION RESULTS AND ANALYSIS

A. Performance Benchmark

Our results surpass existing benchmarks:

- Study [11] (Eastern India): R^2=0.96, RMSE=313-315 Wh for Voting/Stacking models.
- Study [12]: LightGBM showed RMSE=231.9 Wh, MAE=119.6 Wh, R^2=0.87.
- Study [3]: Optimised Random Forest achieved RMSE=252.11 Wh, MAE=143.81, R^2=0.720.
- Study [13]: Gradient Boosting achieved R^2=0.827, RMSE=399.44, MAE=253.62.

Our models (Gradient Boosting for 2023 validation, stacking for 2024 prediction) significantly outperform these benchmarks, achieving R^2>0.99, MAE<79 Wh, MAPE<0.9%, and RMSE<198 Wh, demonstrating notable improvement for tropical regions with high meteorological variability.

B. Climate Change Impact Assessment

The +1.3°C temperature anomaly observed in 2024 provides empirical evidence of climate change's effects on tropical PV systems. Although such an anomaly can reduce PV yields by approximately 0.4-0.6% due to standard temperature coefficients, it has been successfully incorporated into the forecasting framework. The model predicts an annual decline in PV production of 5.3-5.5% (273.9-274.6 MWh, compared to 289.9 MWh in 2023). Thus, these phenomena reflect both direct thermal losses and the indirect effects of changes in atmospheric conditions influencing solar irradiance.

C. Model Selection for Operational Deployment

For the 2024 forecasts, ensemble methods were favoured for their robustness in the face of limited data. Their approaches reduce forecast variance through model diversity. In addition, they improve the quantification of uncertainty for risk assessment, thereby enhancing stability in the face of unpredictable weather conditions.

V. CONCLUSION

This study demonstrated the superior performance of advanced Machine Learning approaches, both individual and ensemble, for forecasting PV production in a tropical microgrid. Based on the Katsepy site in Mahajanga, Madagascar, the results confirm that Gradient Boosting and Stacking Regressor models are particularly well-suited to tropical environments where data is scarce and resources are limited.

A. Main Contributions

The main contributions of this work can be summarized as follows:

i. *Rigorous Validation for 2023:* The Gradient Boosting model achieved the highest accuracy (RMSE = 197.47 Wh, MAE = 78.42 Wh, R^2 >

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

0.99, MAPE < 0.9%), significantly outperforming comparable benchmarks in the literature.

ii. *Prospective Forecasting for 2024:* The Stacking Regressor demonstrated strong generalizability (RMSE = 196.6 Wh, MAE = 77.7 Wh, $R^2 > 0.99$), showing robustness in the face of climate variability.

iii. *Integration of Climate Impact:* The integration of a temperature anomaly of +1.3°C provided relevant empirical information for adapting PV systems to climate change in tropical regions.

iv. *Strategic Model Selection:* The complementary use of Gradient Boosting (for validation) and Stacking (for operational forecasting) demonstrates an effective balance between statistical accuracy and operational robustness.

v. *Hybrid Framework Design:* Combining 19 years of PVGIS data with real temperature measurements provides a flexible and reproducible approach. It also has the advantage of being suitable for resource-constrained tropical environments.

B. Implications and Potential Applications

The performance achieved (MAE < 79 Wh, MAPE < 0.9%) confirms the reliability of the forecasting framework proposed for the Katsepy site. This solid foundation paves the way for numerous future applications, but our research focuses primarily on optimising energy management and fault diagnosis in PV systems within microgrids. And that is precisely the next step in this study.

C. Acknowledged Limitations

Several limitations must be considered to ensure rigorous interpretation of the results, including:

i. *Site-Specific:* The analysis was limited to the Katsepy site, so applying the framework elsewhere will require local recalibration.

ii. *Temporal Resolution:* The use of daily data limits the accuracy of intraday forecasts, yet this is essential for optimal real-time energy management.

iii. *Simulated Production Data:* The use of PVGIS simulations excludes actual operational losses. To obtain results closer to reality, future work will need to incorporate field measurements, degradation and the effects of fouling.

This research establishes a robust, scalable forecasting framework explicitly validated for the Katsepy microgrid. Furthermore, it contributes to sustainable energy management in resource-limited tropical regions. The results confirm that ensemble Machine Learning is a promising approach for addressing operational challenges of isolated microgrids, provided that future studies continue with a view to its validation and deployment in real-world conditions.

While the model has demonstrated its full potential, integrating Deep Learning Models (such as LSTM, CNN, and others) investigated in several studies could improve these results in terms of accuracy and robustness.

ACKNOWLEDGMENT

The authors express their gratitude to PVGIS (Photovoltaic Geographical Information System) for providing comprehensive historical solar data and to AccuWeather for the 2024 temperature records. This work forms part of a doctoral research project on optimal energy management and fault diagnosis for PV microgrids, thus contributing to the sustainable energy transition in Madagascar.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/Competing Interests:** Based on my understanding, this article does not have any conflicts of interest.
- **Funding Support:** This article has not been funded by any organisations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. R. J. François, R. Achille and A. Lala, "Studies on the potential of rural electrification in the Soavinandriana district," International Journal of Progressive Sciences and Technologies (IJIPSAT), University of Antananarivo, vol. 35 No. 1, Oct. 2022, pp. 260-269. Available: https://ijpsat.org/index.php/ijpsat/article/view/4704/2894.
2. Latitude, "Satellite map of Katsepy, Madagascar," [Online]. Available: https://latitude.to/satellite-map/mg/madagascar/384571/katsepy. [Accessed: 05-09-2025].
3. J. Gaboitaolelwe, A. M. Zungeru, "Machine Learning Based Solar Photovoltaic Power Forecasting: A Review and Comparison," *IEEE Access*, vol 11, 2023. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10107594.
4. Vijayaganth V., M. Jayasri PK, Narmadha S., Pavithra D. S et al., "Ensemble Machine Learning based Weather Prediction System," 2023 International Conference on Communication, Security and Artificial Intelligence, *IEEE Xplor*e. DOI: https://doi.org/10.1109/ICCSAI59793.2023.10421274
5. K. P. Natarajan, J. G. Singh et al., "Solar Power Forecasting using Stacking Ensemble Models with Bayesian Meta-Learning," 2023 14th International Conference, Nov. 23, 2023, *IEEE Explore*. DOI: https://doi.org/10.1109/ICCCNT56998.2023.10308202
6. A. Abdellatif, H. Mubarak, S. Ahmad, T. Ahmed, G. M. Shafiullah, A. Hammoudeh, H. Abdellatef, M. Rahman and H. M. Gheni, "Forecasting Photovoltaic Power Generation with a Stacking Ensemble Model," *MDPI, Sustainability*, Vol. 14, Issue 17, Sep. 5, 2022. Available: DOI: https://doi.org/10.3390/su141711083.
7. "Photovoltaic geographical information system," [Online]. Available: https://re.jrc.ec.europa.eu/pvg_tools/en/. [Accessed: 05-09-2025].
8. "AccuWeather," [Online]. Available: https://www.accuweather.com/. [Accessed: 05-09-2025].
9. "DB-City", [Online]. Available: https://fr.db-city.com/Madagascar--Mahajanga--Boeny--Katsepy. [Accessed: 05-09-2025].
10. Wikipedia, "Katsepy," [Online]. Available: https://en.wikipedia.org/wiki/Katsepy. [Accessed: 05-09-2025].
11. D. Chakraborty, J. Mondal, H. B. Barua, A. Bhattacharjee, "Computational Solar Energy-Ensemble Learning Methods for Prediction of Solar Power Generation based on Meteorological Parameters in Eastern India," ResearchGate. Available: January 2023,

DOI: https://doi.org/10.1016/j.ref.2023.01.006
12. M. Boujoudar, I. Bouarfa, A. Dadda, M. Elydrissi, A. M. Taj, M. Abraim, H. Ghennioui and E. G. Bennouna, "Power PV Forecasting Using Machine Learning Algorithms Based on Weather Data in Semi-Arid Climate," BIO Web of Conferences 109, 01024 (2024). DOI: https://doi.org/10.1051/bioconf/202410901024
13. A. R. V. Babu, N. B. Kumar, R. P. Narasipuram, S. Periyannan, A. Hosseinpour and A. Flah, "Solar Energy Forecasting Using Machine Learning Techniques for Enhanced Grid Stability," Vol. 13, 04 June 2025, Digital Object Identifier 10.1109/ACCESS.2025.3574093, *IEEE Access*. DOI: https://doi.org/10.1109/ACCESS.2025.3574093

AUTHOR'S PROFILE

Linda Christelle Mevalaza is a PhD student at the Doctoral School of Renewable Energies and Environment, University of Antsiranana, Madagascar. She holds a Master's degree in Automation, Signal and Image Processing from the Faculty of Sciences at the University of Mahajanga. Her research focuses on developing advanced energy management and fault diagnosis methods for photovoltaic systems integrated into microgrids using Artificial Intelligence techniques. Her academic interests include renewable energy systems, predictive modelling, and intelligent control strategies to improve the performance, reliability, and efficiency of energy systems. She aims to advance sustainable energy technologies by integrating data-driven and innovative approaches.

Dr. Andrianajaina Todizara, is a project manager and researcher with a PhD in Science and Technology focused on renewable energy and artificial intelligence. With over seven years of experience, his work spans educational, agricultural, and digital innovation projects. He combines academic expertise with practical field experience to promote youth empowerment, sustainable development and community.

Dr. Rakotoarisoa Armand Jean Claude is currently a professor in the Department of Electrical Engineering at the Higher Polytechnic School of the University of Antsiranana, Madagascar. He received his PhD from the University of Paris 8, France, in 2011. He has 18 years of teaching experience in different renowned Technical and Engineering Institutes in Madagascar. His current research focuses on Electricity and Electrical Machines, Industrial Process Regulation, and Computer Programming.

Hervé Mangel holds a PhD in engineering. He is actually an assistant Professor at the University of Brest in France. His research interests are in Renewable Energy Technologies, Photovoltaic Systems, Wind Turbines, Control Theory, Network Stability, and Micro-Grid Control. He is also the project manager of the UBO and other Universities.

Jean Nirinarison Razafinjaka is an Automatic Specialist and currently a Professor in the Department of Electrical Engineering at the Higher Polytechnic School of the University of Antsiranana, Madagascar. He received his HDR from the Higher Polytechnic School of the University of Antsiranana, Madagascar, in 2016. He has 35 years of teaching experience at renowned engineering and technical institutes in Madagascar. He has published more than 71 Articles, 16 conference papers, two research papers, and three posters in reputable journals. His current research focuses on Advanced Controls for Electromechanical and Renewable Energy Systems.