

Comparing Deep Learning and Ensemble Models for Photovoltaic Power Forecasting

Edris Shamsulhaq¹, Fikri Arif Wicaksana²

^{1,2} Nusa Putra University, Jl. Raya Cibolang Cisaat - Sukabumi No.21, Sukabumi, 43152, Indonesia

ARTICLE INFO

Article history:

Received : 16/06/2026

Revised : 22/08/2026

Accepted : 09/09/2026

Keywords:

Deep Learning; Ensemble Learning;
LSTM; Photovoltaic Forecasting;
SHAP Analysis; XGBoost

ABSTRACT

Precise forecasting of photovoltaic (PV) power is crucial for maintaining grid stability and effectively incorporating solar energy. This study compares four forecasting models, LSTM, GRU, Random Forest, and XGBoost, using three years of real operational data (2023–2025) from the DKASC solar database. Fourteen input features were constructed from irradiance measurements, meteorological variables, rainfall, wind direction, and cyclical time features. A sliding window of 36 time steps captured temporal dependencies, while the dataset was chronologically split into training, validation, and testing subsets. Model performance was assessed using R^2 , RMSE, MAE, and MAPE. Results show LSTM achieved the highest accuracy with R^2 of 0.9647 and RMSE of 88.40 W, whereas XGBoost obtained the lowest MAPE of 19.18%. SHAP analysis and an ablation study were conducted to identify influential features and evaluate their contributions to forecasting performance. SHAP analysis identified global horizontal radiation, global tilted radiation, and cyclical temporal features, particularly `day_cos` and `hour_cos`, as the most influential predictors, while wind speed and daily rainfall showed limited impact. The findings demonstrate that deep learning models can provide reliable PV power forecasts when combined with informative feature engineering and appropriate temporal window selection.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Corresponding Author:

Edris Shamsulhaq
Nusa Putra University, Sukabumi, 43152, Indonesia
Email: edris.shamsulhaq_te23@nusaputra.ac.id

1. INTRODUCTION

The world's transition to renewable energy sources is accelerated due to environmental concerns and the depletion of fossil fuels, has established solar photovoltaic (PV) systems as the foundation of contemporary sustainable energy infrastructure [1][2][3]. However, the intrinsic fluctuations and sporadic nature of solar radiation—influenced by meteorological factors such as cloud cover, temperature, humidity, and wind—provide serious obstacles to energy dispatch optimization and grid stability, and the reliable integration of PV power into existing electricity networks [4][5][6]. Accurate short-term PV power output predictions is therefore not merely beneficial but essential for balancing supply and demand, lowering the cost of operations, and enhancing the resilience of power systems [7]. This necessity has motivated extensive research into various forecasting approaches. However, a clear understanding of which models and input features are most effective under real-world conditions remains an open question.

Data-driven methods, especially machine learning (ML) and deep learning (DL) models, have gained popularity in recent years, have demonstrated remarkable success in capturing the complex, nonlinear, and time-dependent patterns of solar power generation [8][9]. Among these, recurrent neural network architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are widely recognized for their ability to model temporal dependencies over extended sequences [10]. At the same time, ensemble methods like Random Forest (RF) and Extreme Gradient Boosting (XGBoost) offer robust predictive performance by leveraging multiple decision trees, often with lower computational overhead [11][12]. Despite the proliferation of hybrid and optimized models in the literature—many of which incorporate metaheuristic algorithms, signal decomposition, or external optimization routines [13]—and the increasing recognition that such complex pipelines often lack generalizability [14], there remains a notable gap in fair, direct comparisons of these four fundamental models under strictly identical conditions. Moreover, comprehensive interpretability analyses (e.g., SHAP) and feature ablation studies are seldom reported alongside such comparative evaluations.

To address these gaps, this study presents a systematic and rigorous comparative analysis of LSTM, GRU, Random Forest, and XGBoost for 3-hour-ahead PV power forecasting using real-world operational data. The main contributions of this work are:

- Utilization of three consecutive years (2023–2025) of 5-minute resolution data from the DKASC solar database, with a clear temporal split (2023 training, 2024 validation, 2025 testing) and three independent runs per model to ensure statistical robustness.
- A comprehensive feature engineering process that results in 14 input variables, including multiple irradiance components, meteorological parameters, wind direction (converted to sine/cosine), daily rainfall, and cyclical time features (hour and day of the year).
- Application of a 36-step sliding window (equivalent to 3 hours) to capture short-term temporal dependencies.
- A fair, non-optimized comparison of the four models under identical preprocessing and hyperparameter settings.
- SHAP (SHapley Additive exPlanations) analysis and an ablation study to interpret model predictions and quantify the contribution of each feature group.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes the dataset, preprocessing steps, and the methodological framework. Section 4 presents and discusses the results, including model performance, feature importance, ablation study, and SHAP analysis. The work is finally concluded in Section 5, which also suggests future research areas.

1.2. Related Works

A hybrid FHOGRLSTM model optimized with the Fire Hawk algorithm achieved R^2 values above 0.996 for both polycrystalline and monocrystalline PV arrays [15]. Several machine learning models, including Random Forest and XGBoost, were evaluated across forecast horizons from 0.5 to 24 hours, with Random Forest giving the best average R^2 of 0.89 [16]. A comparison of BiLSTM, LSTM, and XGBoost for solar energy forecasting showed that BiLSTM slightly outperformed LSTM with a test RMSE of 0.1186 versus 0.1187, while XGBoost exhibited strong training fit but weaker generalization [17]. A decomposition-based hybrid model combining EEMD, LSTM, XGBoost, and the Snake Optimizer achieved R^2 values around 0.996 after signal decomposition [18].

An MLPLSTMGA hybrid model for both electrical consumption and PV generation forecasting reported near-perfect R^2 values (0.9999) after genetic algorithm optimization [13]. A direct comparison between LSTM and GRU for solar power forecasting found that GRU outperformed LSTM, achieving a lower MAE (1.618 vs. 1.964) and a higher R^2 (0.9836 vs. 0.9733) [19]. A novel interval forecasting approach combining LSTM, XGBoost, kernel density estimation (KDE), and differential evolution achieved significantly narrower prediction intervals (PINAW of 0.122 at 95% confidence level) [20]. In the context of energy communities, Random Forest achieved the lowest RMSE (5.35 Wh), while GRU excelled in short-term irradiance forecasting [21]. A comparative review highlighted that LSTM offers superior accuracy but higher computational complexity, whereas Random Forest and XGBoost provide

a better balance [22]. Despite these advances, most existing works either rely heavily on metaheuristic optimisation, signal decomposition, or focus on a single model type. In contrast, the present study provides a direct comparison under identical preprocessing, feature engineering (14 features), and temporal windowing (36 steps) using three years of real-world DKASC data, supported by SHAP explainability and ablation analysis.

2. RESEARCH METHOD

2.1. Dataset Description

The dataset comprises three consecutive years of operational data (2023–2025) obtained from the DKASC photovoltaic system, with 100_DKA_M1_A_Phase_Active_Power selected as the target variable. The recorded variables encompass solar irradiance and core meteorological parameters, with data split chronologically into training (2023), validation (2024), and testing (2025) subsets as detailed in Table 1.

Table 1. Summary of the dataset across the three years (after preprocessing)

Year	Split	Samples	Min Power (W)	Max Power (W)	Mean Power (W)	Std Power (W)
2023	Training	101,175	0.0	1674.9	344.9	488.6
2024	Validation	103,356	0.0	1674.3	331.8	471.4
2025	Testing	105,179	0.0	1641.2	337.5	471.5

Note: The number of samples is approximate (assuming 5-minute intervals for a full year: $365 \times 24 \times 12 = 105,120$). Max power = $1.6413 \text{ kW} = 1641.3 \text{ W}$; mean power = $0.3373 \text{ kW} = 337.3 \text{ W}$.

2.2. Data Preprocessing

Several preprocessing steps were applied before model development:

- **Missing values** were handled using linear interpolation followed by backward-fill and forward-fill. Any remaining missing values were replaced with zeros.
- **Invalid or negative power values** (due to sensor noise) were clipped to zero: $y = \max(y, 0)$. Similarly, negative radiation values were set to zero.
- **Wind direction** (measured in degrees) was transformed into sine and cosine components to preserve its cyclical nature: $\text{wind_dir_sin} = \sin(\theta \times \pi/180)$, $\text{wind_dir_cos} = \cos(\theta \times \pi/180)$.
- **Timestamp information** was converted into cyclical temporal features. Hour of day: $\text{hour_sin} = \sin(2\pi \times \text{hour}/24)$, $\text{hour_cos} = \cos(2\pi \times \text{hour}/24)$. Day of year: $\text{day_sin} = \sin(2\pi \times \text{day}/365)$, $\text{day_cos} = \cos(2\pi \times \text{day}/365)$.

Fourteen input variables made up the final feature set: temperature, relative humidity, wind speed, wind direction (sine and cosine), global horizontal radiation, diffuse horizontal radiation, global tilted radiation, diffuse tilted radiation, daily rainfall, hour sine, hour cosine, day sine, and day cosine. To guarantee numerical stability, all features were normalized to the range [0,1] using Min-Max scaling.

2.3. Time-Series Window Construction

To capture temporal dependencies in PV generation, a sliding-window approach was employed. A fixed window size of 36 consecutive observations (approximately 3 hours, given a 5-minute sampling interval) was used. For each prediction sample, the previous 36 time steps were provided as model input, while the power value at the next time step was used as the prediction target. Formally, the input sequence is $X_t = \{x_{t-35}, x_{t-34}, \dots, x_t\}$ and the corresponding target value is y_{t+1} . This strategy enables the models to learn temporal patterns and short-term dynamics in solar power production.

2.4. Prediction Models

Four machine learning and deep learning models were evaluated. All models were trained using the same training/validation/test split (2023/2024/2025) and were run three independent times with different random seeds to ensure statistical robustness.

1. Long Short-Term Memory (LSTM)

LSTM is a sophisticated recurrent neural network that uses gating methods to capture long-term relationships. The LSTM cell incorporates three multiplicative gates—the forget gate, input gate, and output gate—that regulate the flow of information by selectively retaining or discarding past states, thus mitigating the vanishing gradient problem commonly encountered in standard RNNs[24][25]. One LSTM layer with 64 hidden units, a 0.1 dropout rate, and a dense layer with linear activation make up the implemented architecture. Using mean squared error loss, the model was optimized using the Adam method (learning rate = 0.001). In order to reestablish the optimal weights, training was limited to 50 epochs with an early stopping patience of five periods depending on validation loss.

2. Gated Recurrent Unit (GRU)

A streamlined version of the LSTM, the GRU combines the input and forget gates into a single update gate. Specifically, the GRU employs an update gate that determines how much of the past information to carry forward, and a reset gate that controls the integration of new input with the previous hidden state. It often requires fewer parameters and can be computationally more efficient [26][27]. The GRU model followed the same configuration as the LSTM: one GRU layer with 64 hidden units, a dropout layer (0.1), a dense output layer (1 neuron), and the same optimizer, loss function, and early stopping settings.

3. Random Forest (RF)

Random Forest is an ensemble learning technique that creates many decision trees and averages their predictions. The algorithm constructs each tree using a bootstrap sample of the training data (bagging) and randomly selects a subset of features at each split, which decorrelates the individual trees and enhances generalization. It is robust to overfitting and can capture nonlinear relationships without requiring feature scaling [28][29]. The RF regressor was configured with 50 decision trees (`n_estimators = 50`), a maximum tree depth of 12 (`max_depth = 12`), and parallel processing enabled (`n_jobs = -1`).

4. Extreme Gradient Boosting (XGBoost)

XGBoost is a tree boosting technique that uses regularization to avoid overfitting and produces trees in a sequential fashion, each of which fixes the mistakes of its predecessor by focusing on the residuals of the previous trees. Unlike Random Forest, which builds trees independently, XGBoost employs a gradient-based optimization framework that minimizes a regularized objective function combining a loss term and penalty terms for model complexity (L1 and L2 regularization), which controls overfitting and enhances predictive performance [30][31]. The XGBoost model was configured with 500 boosting rounds, maximum tree depth of 8, learning rate of 0.03, subsample ratio of 0.8, column sampling ratio of 0.8, and objective function: `reg:squarederror`.

2.5. Model Evaluation

Each model was trained and assessed three times independently using distinct random seeds in order to guarantee outcome stability. The obtained metrics' mean $\pm$ standard deviation matches the stated data. Four commonly used regression measures were used to evaluate the model's performance: Mean Absolute Error (MAE), Coefficient of Determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Samples with genuine power ≤ 0.01 W were not included in MAPE in order to prevent numerical instability (safe MAPE).

2.6. Feature Importance Analysis

The trained model's feature significance was derived using the built-in `feature_importances` attribute. Importance scores were aggregated across all 36 time lags to obtain a single importance value per original feature. This allowed the identification of the most influential variables affecting PV power prediction.

2.7. Ablation Study

To assess the contribution feature groups, an ablation study was carried out. Six experimental configurations were taken into account:

1. Full Features – all 14 variables.
2. Without Radiation – removal of the four radiation variables.
3. Without Time – removal of the four cyclical time features (hour_sin, hour_cos, day_sin, day_cos).
4. Without Weather – removal of temperature, humidity, wind speed, wind direction (sin/cos), and daily rainfall.
5. Without Wind – removal of wind speed and wind direction sin/cos.
6. Radiation + Time Only – retention of only the four radiation variables and the four time variables.

Table 3. Feature importance obtained from XGBoost (aggregated over time lags)

Feature	Importance
radiation_global_horiz	0.4293
radiation_global_tilted	0.1664
hour_sin	0.1384
hour_cos	0.1014
temperature	0.0249
day_cos	0.0242
humidity	0.0232
wind_dir_sin	0.0220
wind_dir_cos	0.0200
radiation_diffuse_horiz	0.0153
radiation_diffuse_tilted	0.0152
daily_rainfall	0.0115
day_sin	0.0083
wind_speed	0.0000

2.8. SHAP Explainability Analysis

To enhance model interpretability, SHAP (SHapley Additive exPlanations) [32][33] analysis was applied to the final XGBoost model. SHAP values were computed using a subset of the test dataset (500 samples) and subsequently aggregated across the temporal window to obtain a per-original-feature importance score. Three types of visualisations were generated: a SHAP bar plot (aggregated absolute SHAP values per original feature), a SHAP summary plot (scatter plot showing the distribution of SHAP values for the top 20 time-lagged features), and a SHAP waterfall plot (detailed decomposition of a single prediction).

2.9. Visual Diagnostics

For the XGBoost model, additional diagnostic plots were created to visually assess prediction quality: a scatter plot of actual versus predicted power (with a diagonal reference line), a residuals plot (sequence of prediction errors over the test set to detect systematic bias), and a time series comparison (overlay of actual and predicted power for the first 1000 test samples to visually evaluate temporal tracking).

3. RESULTS AND DISCUSSION

3.1. Comparative Model Performance

The performance of the four prediction models on the 2025 test set is summarized in Table 4. Each value is the mean $\pm$ standard deviation over three independent runs.

Table 4. Comparative performance of forecasting models

Model	R ²	RMSE (W)	MAE (W)	MAPE (%)
LSTM	0.9647 $\pm$ 0.0026	88.40 $\pm$ 3.30	36.20 $\pm$ 2.90	22.63 $\pm$ 1.24
GRU	0.9512 $\pm$ 0.0058	103.90 $\pm$ 6.30	39.70 $\pm$ 3.50	24.09 $\pm$ 2.25
Random Forest	0.9075 $\pm$ 0.0001	143.34 $\pm$ 0.04	52.89 $\pm$ 0.04	31.73 $\pm$ 0.02
XGBoost	0.9463 $\pm$ 0.0022	109.20 $\pm$ 2.20	39.80 $\pm$ 0.40	19.18 $\pm$ 0.14

As presented in Table 4, the LSTM model demonstrated the highest predictive accuracy ($R^2 = 0.9647$, $RMSE = 88.40$ W), confirming its superior capability in capturing temporal dependencies. Conversely, Random Forest yielded the weakest baseline ($R^2 = 0.9075$, $RMSE = 143.34$ W). Under the 36-step sliding window which generates 504 lagged inputs, tabular ensemble models faced an expanded feature space prone to high collinearity and the curse of dimensionality. However, empirical findings revealed a sharp divergence in mitigation: XGBoost showed exceptional resilience to overfitting, achieving a robust R^2 of 0.9463 and the lowest overall MAPE (19.18% plus or minus 0.14%) due to its built-in L1/L2 regularization and sparsity-aware splitting. In contrast, Random Forest's random feature sampling at each split point frequently selected sub-optimal, highly collinear lagged candidates, thereby degrading its structural predictive capacity.

3.2. Feature Importance

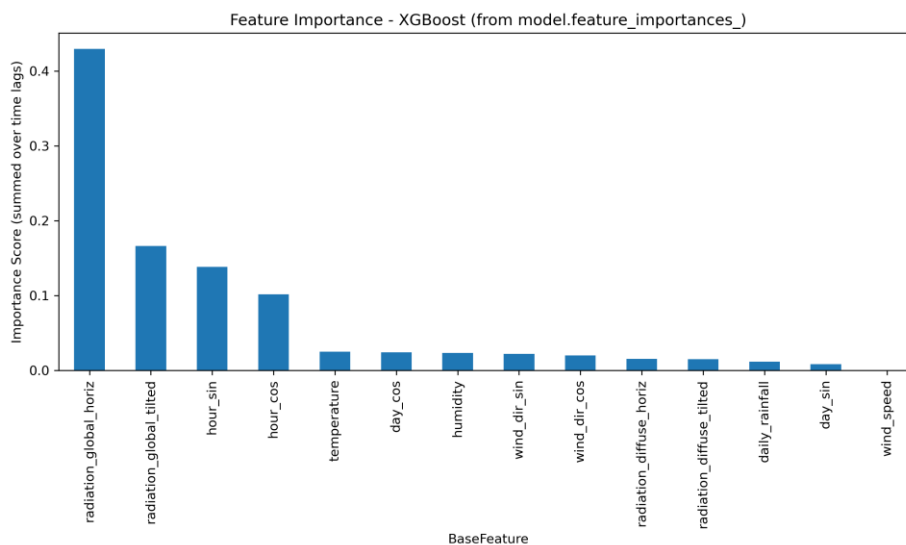


Figure 1. Feature importance from XGBoost (aggregated over time lags).

Figure 1 shows the feature importance from XGBoost (aggregated over time lags). Global horizontal radiation is the dominant factor (importance 0.4293), followed by global tilted radiation (0.1664) and hour sine (0.1384). Wind speed shows negligible importance (≈ 0.0000). These results are physically consistent with PV energy conversion principles.

3.4 Ablation Study

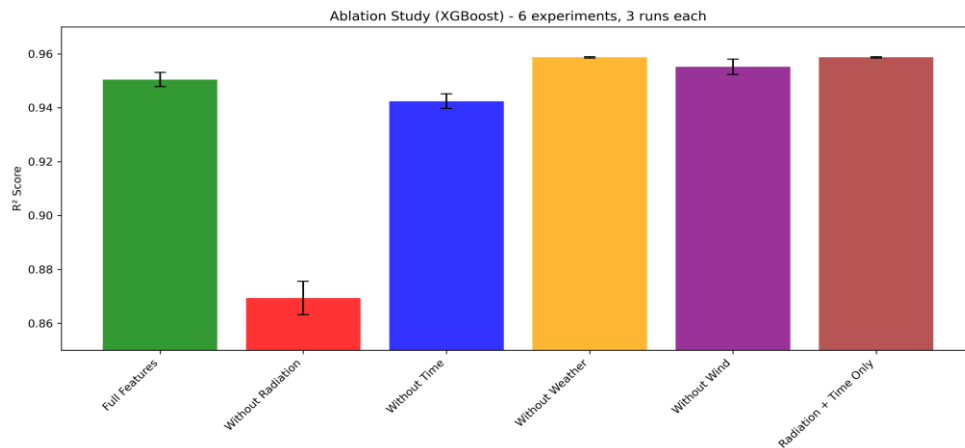


Figure 2. shows the ablation study results (R^2 for six feature sets).

Table 5. Mean absolute SHAP values

Rank	Feature	Mean SHAP
1	Radiation Global Horizontal	0.131
2	Day Cos	0.122
3	Radiation Global Tilted	0.078
4	Hour Cos	0.028
5	Wind Direction Sin	0.025
6	Radiation Diffuse Horizontal	0.018
7	Temperature	0.017
8	Radiation Diffuse Tilted	0.016
9	Humidity	0.014
10	Hour Sin	0.012

Table 6. Ablation study results (XGBoost, mean $R^2 \pm$ std over 3 runs)

Feature Set	Num Features	$R^2 \pm$ std	% Change
Full Features	14	0.9505 ± 0.0027	0.00
Without Radiation	10	0.8694 ± 0.0062	-8.53
Without Time	10	0.9424 ± 0.0027	-0.85
Without Weather	8	0.9587 ± 0.0003	+0.87
Without Wind	11	0.9552 ± 0.0029	+0.50
Radiation + Time Only	8	0.9587 ± 0.0003	+0.87

The most severe performance degradation occurs when radiation variables are removed (R^2 drops from 0.9505 to 0.8694, a -8.53% change). Removing time features causes only a minor decrease (-0.85%). Interestingly, removing weather variables (temperature, humidity, wind, rainfall) slightly improves R^2 (+0.87%), suggesting that some of these variables introduce noise. The Radiation + Time Only configuration achieves the highest R^2 (0.9587), indicating that irradiance and temporal information alone are sufficient for accurate prediction.

3.5 SHAP Explainability

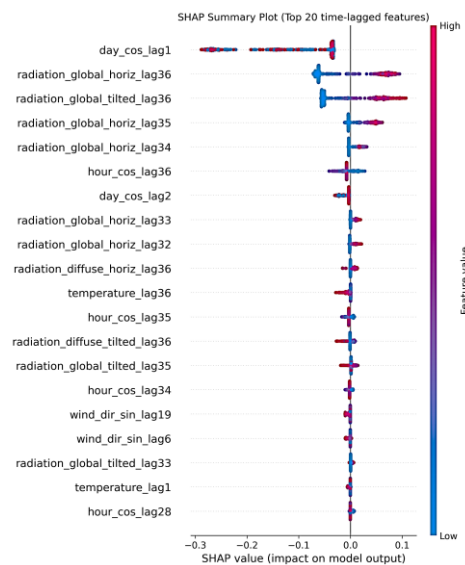


Figure 3. SHAP summary plot (top 20 time lagged)

Figure 3 presents the SHAP summary plot, which provides a comprehensive overview of the contribution of the most influential features to the XGBoost model predictions. The results indicate that solar irradiance variables, particularly global horizontal radiation and global tilted radiation, have the strongest impact on photovoltaic power forecasting. Higher irradiance values are generally associated with positive SHAP values, leading to increased predicted power output, whereas lower irradiance values contribute negatively to the prediction. Temporal features, especially `day_cos` and `hour_cos`, also exhibit substantial influence, reflecting the seasonal and diurnal patterns inherent in photovoltaic power generation. In contrast, variables such as wind speed and daily rainfall show relatively small SHAP contributions, indicating a limited effect on forecasting performance. Overall, the SHAP analysis confirms that irradiance and temporal information are the primary drivers of PV power prediction and supports the findings obtained from the feature importance and ablation analyses, thereby enhancing the interpretability and reliability of the proposed forecasting framework.

4. CONCLUSION

In this study, three years of actual operational data (2023–2025) from the DKASC database were used to create and assess four forecasting models for photovoltaic (PV) power prediction: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Random Forest (RF), and XGBoost. A sliding window of 36 time steps (three hours) was employed to capture temporal dependencies, and a comprehensive feature engineering process was carried out, resulting in fourteen input variables that included multiple irradiance components, meteorological measurements, wind characteristics, daily rainfall, and cyclical time features (hour and day of year). All models were trained on 2023 data, validated on 2024 data, and tested on unseen 2025 data, with each experiment repeated three times to report mean $\pm$ standard deviation.

The LSTM model achieved the highest overall performance, attaining an R^2 score of 0.9647, the lowest RMSE (88.40 W), and the lowest MAE (36.20 ± 2.90 W). XGBoost, while slightly lower in R^2 (0.9463), delivered the best MAPE (19.18%), demonstrating its strength in relative-error minimization. GRU performed competitively ($R^2 = 0.9504$), whereas Random Forest yielded the weakest results, indicating limited capability in capturing complex nonlinear temporal patterns.

SHAP analysis identified global horizontal radiation as the most influential variable, followed by `day_cos`, global tilted radiation, and `hour_cos`. Time features contributed significantly, while wind speed and daily rainfall had negligible impact. The ablation study confirmed that removing radiation variables caused a drastic drop in R^2 (−8.53%), whereas eliminating weather variables (temperature, humidity, wind, rainfall) slightly improved performance, suggesting that some meteorological variables may

introduce noise. The ‘Radiation + Time Only’ configuration achieved the highest R^2 (0.9587), indicating that irradiance and temporal information alone are sufficient for accurate PV power forecasting.

More sophisticated like transformer-based models (like Informer and Temporal Fusion Transformer), which have demonstrated potential in long-sequence forecasting, may be investigated in future research. Multi-step ahead forecasting and probabilistic forecasting to quantify prediction uncertainty are also valuable directions. Furthermore, future work will explore advanced feature extraction techniques to alleviate the dimensionality of the sliding window for tabular models. Finally, deploying the best-performing model (LSTM or XGBoost) in a real-time edge computing environment for on-site PV power prediction would constitute a practical extension of this research.

Acknowledgments

The authors thank Nusa Putra University for institutional support. We thank the Desert Knowledge Australia Solar Centre (DKASC) for releasing the dataset to the public.

REFERENCES

- [1] Z. A. Aljawary et al., “Sustainable Energy Generation through Photovoltaic Systems: Current Trends and Future Perspectives,” pp. 1–6, Dec. 2025, doi: 10.1109/MECON67253.2025.11276917.
- [2] T. Cang, “Comprehensive Exploration of Solar Photovoltaic Technology: Enhancing Efficiency, Integrating Energy Storage, and Addressing Environmental and Economic Challenges,” *Applied and Computational Engineering*, vol. 123, no. 1, pp. 10–16, Jan. 2025, doi: 10.54254/2755-2721/2025.19565.
- [3] S. Kalogirou, *Solar Energy Engineering: Processes and Systems*, Elsevier, 2024.
- [4] B. Sivadharshini et al., “Solar Energy Forecasting Models for Grid Integration and Power Balancing,” *E3S Web of Conferences*, vol. 540, p. 10022, Jun. 2024, doi: 10.1051/E3SCONF/202454010022.
- [5] C. Wan et al., “Photovoltaic and solar power forecasting for smart grid energy management,” *CSEE Journal of Power and Energy Systems*, vol. 1, no. 4, pp. 38–46, Jan. 2016, doi: 10.17775/CSEEJPES.2015.00046.
- [6] W. A. Beckman, N. Blair, and J. A. Duffie, *Solar Engineering of Thermal Processes, Photovoltaics and Wind*, 5th ed., Wiley, 2021.
- [7] N. Y. Hendriks et al., “All sky imaging-based short-term solar irradiance forecasting with Long Short-Term Memory networks,” *Solar Energy*, vol. 272, p. 112463, Apr. 2024, doi: 10.1016/J.SOLENER.2024.112463.
- [8] M. Jordan, J. Kleinberg, and B. Schölkopf, “Pattern Recognition and Machine Learning,” Springer, 2006.
- [9] A. Nadeem et al., “AI-Driven precision in solar forecasting: Breakthroughs in machine learning and deep learning,” *AIMS Geosciences*, vol. 10, no. 4, pp. 684–734, 2024, doi: 10.3934/GEOSCI.2024035.
- [10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [11] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer, 2021, doi: 10.1007/978-1-0716-1418-1.
- [12] R. Shwartz-Ziv and A. Armon, “Tabular Data: Deep Learning is Not All You Need,” *Information Fusion*, vol. 81, pp. 84–90, Jun. 2022, doi: 10.1016/j.inffus.2021.11.011.
- [13] C. F. Mbey et al., “A Novel Deep Learning-Based Data Analysis Model for Solar Photovoltaic Power Generation and Electrical Consumption Forecasting in the Smart Power Grid,” *Applied Computational Intelligence and Soft Computing*, vol. 2024, 2024, doi: 10.1155/2024/9257508.

- [14] T. N. Nguyen and F. Müsgens, "A meta-analysis of solar forecasting based on skill score," *Journal of Renewable and Sustainable Energy*, vol. 18, no. 2, Mar. 2026, doi: 10.1063/5.0300682.
- [15] P. N. L. M. Radzi et al., "Optimizing solar power forecasting with metaheuristic algorithms and deep learning models for photovoltaic grid connected systems," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-23822-1.
- [16] A. J. Lari et al., "Using Machine Learning Algorithms to Forecast Solar Energy Power Output," *Electronics*, vol. 14, no. 5, Mar. 2025, doi: 10.3390/electronics14050866.
- [17] H. Alizadegan et al., "Solar Energy Production Forecasting: A Comparative Study of Bi-LSTM, LSTM, XGBoost Models with Activation Function Analysis," May 2024. doi: 10.20944/preprints202405.0994.v1.
- [18] Y. Xu, X. Ji, and Z. Zhu, "A photovoltaic power forecasting method based on the LSTM-XGBoost-EEDA-SO model," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-16368-9.
- [19] A. D. Huynh and T. K. Nguyen, "The Comparison of GRU and LSTM in Solar Power Generation Forecasting Application," *International Journal of Science and Research Archive*, vol. 13, no. 1, pp. 1360–1370, Sep. 2024, doi: 10.30574/ijrsra.2024.13.1.1831.
- [20] D. K. Avoussou, A. A. Salami, and D. Barchiesi, "Interval forecasting of solar irradiation using LSTM, XGBoost and bandwidth optimization of KDE by differential evolution," Sep. 2026, *Solar Compass*, doi: 10.1016/j.solcom.2026.100171.
- [21] F. Aksan et al., "A Comparative Study of Machine Learning Models for PV Energy Prediction in an Energy Community," *Energies*, vol. 18, no. 22, Nov. 2025, doi: 10.3390/en18225980.
- [22] A. Mollasalehi and A. Farhadi, "Solar and Wind Power Forecasting: A Comparative Review of LSTM, Random Forest, and XGBoost Models," Sep. 2025. [Online]. Available: <http://arxiv.org/abs/2509.24059>.
- [23] Desert Knowledge Australia, "Desert Knowledge Australia Solar Centre (DKASC) Historical Dataset (2023–2025)," Desert Knowledge Australia Solar Centre.
- [24] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*, Cambridge University Press, 2023.
- [25] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, pp. 1735–1780, Jun. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [26] F. Chollet, *Deep Learning with Python*, 2nd ed., Manning, 2021.
- [27] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," Dec. 2014. [Online]. Available: <https://arxiv.org/pdf/1412.3555>.
- [28] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2009, doi: 10.1007/978-0-387-84858-7.
- [29] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [30] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.
- [31] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.
- [32] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 4766–4775, 2017.
- [33] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed., 2022. [Online]. Available: <http://leanpub.com/interpretable-machine-learning>.