Shahid Bahonar University of Kerman
Biomechanism and Bioenergy Research
Online ISSN: 2821-1855
Homepage: https://bbr.uk.ac.ir
Iranian Society of Agricultural Machinery Engineering and Mechanization

Accurate PV Power Forecasting Using a Lightweight LSTM Model: A Case Study

Soha Sami[1], Seyed Reza Hasan-beygi[1], Jafar Massah[1]

[1] Mechanics of Biosystems Engineering Department, College of Aburaihan, University of Tehran, Tehran, Iran.

Corresponding author: rhbeigi@ut.ac.ir

ARTICLE INFO

Article type:

Research Article

Article history:

Received 14 June 2025

Received in revised form 28 July 2025

Accepted 30 August 2025

Available Online 30 September 2025

Keywords:

Photovoltaic power forecasting; LSTM; short-term prediction; renewable energy; time series; feature selection; grid stability; computational efficiency.

ABSTRACT

This study presents a significant advancement in short-term photovoltaic (PV) power forecasting through the development and validation of a simple yet highly effective LSTM-based model tailored to address the operational demands of renewable energy integration. By harnessing a meticulously curated dataset and employing rigorous feature selection, the model achieved exceptional performance metrics an R^2 score of 0.9212 and an RMSE of 0.0650 on unseen data outperforming benchmark models such as CatBoost and GBR. These outcomes affirm the model's capacity to capture temporal dependencies in PV generation data while maintaining computational efficiency, making it well suited for real-time energy management applications. However, limitations such as dependence on high-quality input data and untested resilience under extreme weather conditions suggest areas for refinement. Future research could enhance the model by incorporating probabilistic forecasting, lightweight attention mechanisms, or transfer learning to improve adaptability across diverse geographic and climatic contexts. Ultimately, this work contributes a robust, practical tool to the evolving landscape of smart grid technologies, supporting the global transition toward sustainable energy systems with improved forecasting precision and scalability.

Cite this article: Sami, S., Hasan-beygi, S. R., & Massah, J.(2025).). Drying Food Waste Using Cabinet Dryer with a Conventional Tray. *Biomechanism and Bioenergy Research*, *4*(3), 68-80. https://doi.org/10.22103/bbr.2025.25811.1132

Publisher: Shahid Bahonar University of Kerman
DOI: https://doi.org/10.22103/bbr.2025.25811.1132

INTRODUCTION

Accurate forecasting of photovoltaic (PV) power generation could be fundamental to modern energy network management and the optimal integration of renewable energy sources. It mitigated the uncertainties associated with atmospheric variability, thereby enabling more effective planning for energy production, distribution, and storage. In smart energy systems that rely heavily on renewable energies, reliable forecasting enhanced grid stability and reduced operational costs (Liu et al., 2022; Rezvani et al., 2022).

The economic impact of precise PV forecasting could also be considerable. For example, in South Korea, the implementation of accurate forecasting models has contributed to reduce production fluctuations and operational costs, and ultimately increased the revenue of solar power plants. Furthermore, during peak demand periods, accurate forecasting's enable power plants to inject more energy into the grid, thereby improving profitability. In some cases, such as energy storage systems, accurate forecasting has been increased profitability up to 20% (Araya et al., 2023). In Poland, forecasting models predicted that PV generation would reach up to 31,219 GWh by 2028, potentially reduced national energy costs (Izdebski & Kosiorek, 2023). Moreover, big data-based forecasting models have been reported to reduce energy storage costs by approximately 15% (Stoliarov, 2024). These findings collectively are underscored the strategic importance of advanced forecasting to achieve operational efficiency and economic viability in renewable energy systems.

In recent decades, solar energy has emerged as a key driver in the transition toward sustainable and renewable power systems. With rising global energy demands and growing concerns over greenhouse gas emissions, PV systems have gained widespread adoption as clean, scalable, and increasingly cost-effective energy sources. However, the intermittent and weather-dependent nature of solar generation have presented significant challenges for ensuring a reliable energy supply. Therefore, accurate forecasting would be essential not only for operational planning but also for maintaining grid reliability and economic balance (Di Leo et al., 2025; Fan et al., 2024). As such, solar power forecasting has become a critical tool for both utilities and policy-makers in the deployment of intelligent energy management systems (Balal et al., 2023; Praveenraj et al., 2024).

In parallel, artificial intelligence (AI)-enabled forecasting methods have promised results in various regional implementations. For example, in Greece, AI-driven forecasting systems have significantly improved power distribution efficiency and predictive maintenance, leading to increase energy output and reduce operational costs (Alexakos et al., 2022). These outcomes could reinforce the strategic potential of AI-based forecasting in supporting reliable and cost-effective energy transitions.

Despite recent advancements in solar forecasting methodologies, several limitations are persisted that included, reducing predictive accuracy under highly variable weather conditions, dependence on high-quality input data, and the computational burden associated with many advanced algorithms (Sleiman & Su, 2024). Among the various approaches, deep learning models—particularly long short-term memory (LSTM) networks—have demonstrated significant potential in capturing the complex temporal dependencies inherent in PV generation data (Fraga-Hurtado et al., 2025). Nevertheless, the development of models that are both locally adaptable and computationally efficient remains an open research challenge (Jakoplić et al., 2023). The aim of this study is to address this gap by proposing an optimized LSTM-based framework that balances forecasting accuracy, computational feasibility, and practical deployment ability.

Owing to the stochastic nature of solar irradiance, PV forecasting continues to be an active area of research. Deep learning approaches—particularly LSTM-based architectures—have been widely employed to capture temporal dependencies and improve prediction performance (Jailani et al., 2023). Recent efforts have explored hybrid models, such as LSTM-transformer or CNN-LSTM networks, which combined different learning paradigms to increase accuracy. However, these models often involved high computational costs, making them unsuitable for real-time or resource-constrained settings (Salman et al., 2024).

The choice between forecasting methods is often depended on the specific application. Deterministic forecasts are typically suitable for short-term operational decisions, whereas probabilistic forecasts are increasingly critical for grid stability and long-term risk management. Studies have shown that probabilistic models outperform deterministic models in representing uncertainty, particularly during nonstationary periods (An et al., 2021; Bracale et al., 2016). This growing distinction is emphasized the relevance of probabilistic approaches in the evolution of modern, flexible energy systems.

To address the limitations of traditional models, research workers have integrated metaheuristic optimization techniques—such as Aquila optimization and the crested porcupine optimizer—to fine-tune LSTM hyperparameters, resulting in improved performance (Fan et al., 2024; Liu et al., 2022). Nonetheless, these enhancements are typically required large, high-resolution datasets and introduce additional model complexity. These constraints highlighted the pressing need for more efficient, scalable, and context-adaptive forecasting solutions.

Despite considerable progress, several key research gaps are remained. First, the most state-of-the-art forecasting models relied on high-resolution, location-specific meteorological data, which may not be available in all regions (Sabri & El Hassouni, 2023). Second, while hybrid deep learning models such as CNN-LSTM-Attention offered high accuracy, their computational overhead are often limited their applicability in real-time forecasting environments (Ahmed et al., 2025). Finally, there is a growing need for forecasting frameworks that are not only accurate but also lightweight and adaptable to diverse deployment contexts. In response, this study would propose a computationally efficient, locally adaptable LSTM-based architecture tailored for short-term solar energy forecasting.

MATERIALS AND METHODS

Data

The dataset used in this study was compiled from multiple reliable sources to ensure both accuracy and comprehensiveness. The primary data regarding PV production, electricity consumption, and other building-related features were obtained from the IPart platform, which provided high-resolution energy data from real residential and industrial buildings.

In order to enhance the performance of the forecasting model, this dataset was enriched by merging it with weather data retrieved from the Ninja platform, which included historical meteorological variables such as temperature, wind speed, and cloud cover.

In addition, several important solar radiation features—such as global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI) were computed via the PVLib library, which is a well-established tool for solar energy modeling and simulation.

As a result, the energy performance data and meteorological parameters were combined into a unified and time-aligned format, provided a solid foundation for training and evaluating the LSTM-based forecasting model.

Feature Selection

In order to identify the most relevant input features for the PV production forecasting model, the Pearson's correlation coefficient was employed as a statistical method to evaluate the linear relationships between variables. The strength and direction of the linear association between two continuous variables, X and Y is quantified by the Pearson's correlation coefficient, Equation (1) (Hapsari et al., 2025).

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (1)$$

where, X_i and Y_i represent the individual sample values of variables X and Y, respectively; $\bar{X}$ and $\bar{Y}$ denote their mean values; and n is the total number of samples. The coefficient r ranges from −1 to +1, where values close to +1 indicate a strong positive linear correlation, values near −1 indicate a strong negative linear correlation, and values around zero suggest little to no linear relationship.

Features demonstrating a strong or moderate correlation with the target variable, PV output, were selected as inputs to the LSTM model. This feature selection process was critical to reduce dataset dimensionality, eliminate irrelevant or redundant inputs, and enhance the training efficiency and generalizability of the forecasting model.

Data Preparation:

The initial dataset, composed of weather and PV-related features, was first augmented by incorporating lagged variables (shifted by one time step) to provide the model with a temporal context and recent trends. All the features were normalized via Min-Max scaler to ensure that the input values fell within the [0, 1] range, which could facilitate stable training.

The following variables were selected as the model inputs on the basis of their correlation with PV output: generation, Day_cos, zenith, elevation, irradiance surface, GHI, DNI, DHI, and AOI. Shifted versions of these features were also included to improve temporal representation.

In order to convert the input data into a suitable format for LSTM training, it was structured to overlap sequences of 24-time steps (corresponding to hourly data over a day). Each sequence was associated with the PV generation value at the next time step, forming the basis for supervised learning.

Model architecture

The meticulously designed LSTM model was crafted with two stacked LSTM layers, each housing 64 units. To ensure robustness, drop-out regularization (with a rate of 0.3) was meticulously applied after each LSTM layer to prevent overfitting. A dense layer with one output neuron was meticulously used at the final layer to produce the forecast value, further enhancing the model's robustness.

The model was compiled with the Adam optimizer and trained with the mean squared error (MSE) as the loss function. Early stopping and learning rate reduction strategies were employed to enhance training stability and convergence. Table 1 shows the details of the network. Also, Figure 1 shows the overall process of the model and its schematic.

Table 1. The developed LSTM Model Architecture

Layer (Type)	Output Shape	Parameters	Description
Input Layer	(24, 18)	0	24-time steps, 18 features (9 original + 9 shifted)
LSTM (1st layer)	(24, 64)	21,248	return_sequences=True
Drop-out	(24, 64)	0	Dropout rate: 0.3
LSTM (2nd layer)	(64)	33,024	return_sequences=False
Dropout	(64)	0	Drop-out rate: 0.3
Dense (Output)	(1)	65	Linear activation
Total Parameters	—	54,337	Summed from all layers

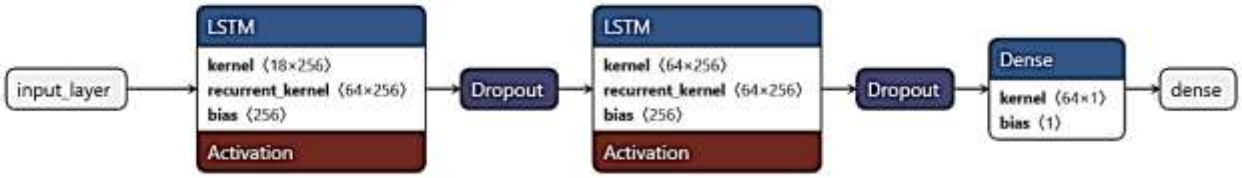

Figure1. LSTM model schematic

Model Evaluation

The performance of the LSTM model in forecasting photovoltaic energy production was assessed by four widely used error metrics, calculated on the both training and test datasets:

Mean Absolute Error (MAE)

This metric measures the average magnitude of the absolute differences between the predicted and actual values. It is straightforward to interpret and less sensitive to outliers.

$$MAE = (1/n) * \Sigma|y_i - \hat{y}_i| \tag{2}$$

Equation 2 describes how to obtain the MAE, where y_i is the actual value of sample i, $\hat{y}_i$ is the predicted value of sample i, and n is the total number of samples. The symbol Σ represents the sum of all samples from i=1 to n, and the absolute value $|y_i - \hat{y}_i|$ represents the absolute difference between the actual and predicted values. MAE, a simple and interpretable measure, is the average of the model error and is less sensitive to outliers, making it easy to understand and use (Robeson & Willmott, 2023; Willmott & Matsuura, 2005).

Mean squared error (MSE)

The MSE calculates the average of the squared differences between the predicted and actual values. It penalizes larger errors more heavily, making it useful for identifying large deviations in predictions.

In Equation 3, the formula for calculating MSE is stated as the square of the difference $(y_i-\hat{y}_i)$^2, which makes larger errors have a greater impact. As with MAE, y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of samples. MSE is not just a theoretical concept but a practical tool that is useful in identifying large deviations in predictions, making it highly relevant and useful in real-world scenarios (Zhao et al., 2023).

$$MSE = (1/n) * \Sigma(y_i - \hat{y}_i)^2 \tag{3}$$

Root mean square error (RMSE)

As the square root of the MSE, the RMSE retains the same unit as the original data and provides a more intuitive understanding of the magnitude of the prediction error Equation 4 shows the formula for obtaining RMSE (Chicco et al., 2021).

$$RMSE = \sqrt{(1/n) * \Sigma(y_i - \hat{y}_i)^2} = \sqrt{MSE} \tag{4}$$

Coefficient of determination (R^2 score)

This statistical metric indicates how well the model explains the variance in the target variable. An R^2 value closer to 1 suggests that the model has strong predictive ability. In equation 5, $\bar{y}$ is the mean of the actual values and the denominator is the sum of the squares of the deviations of the actual values from the mean. An R^2 close to 1 indicates a high accuracy of the model in predictions(Yadav et al., 2025).

$$R^2 = 1 - [\Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2], \bar{y} = (1/n) * \Sigma y_i \tag{5}$$

Implementation of different models for comparison

Two other models, CatBoost and gradient boosting regression (GBR), were also implemented on the data. The both models were optimized through trial and error. The results were compared and analyzed with those of the LSTM model via validation criteria.

RESULTS AND DISCUSSION

Correlation Analysis of Selected Features

Triangular Pearson's correlation heatmap that was illustrated the linear relationships between PV power generation and selected input features is shown in Figure 2. The Figure reveals that the PV output had strong positive correlations with irradiance-related variables, particularly global horizontal irradiance (GHI) and direct normal irradiance (DNI), indicating their direct influence on energy production. Additionally, solar geometric features such as the cosine of the day, elevation angle, and azimuth angle also showed moderate to high positive correlations with PV generation, suggesting their relevance in capturing seasonal and diurnal patterns of solar exposure. Conversely, features such as the angle of incidence (AOI) and zenith angle have negatively correlated with the PV output, reflecting the physical inverse relationship between these angles and effective solar irradiance on the panels. These correlations guided the feature selection process by identifying the most impactful predictors for model training.

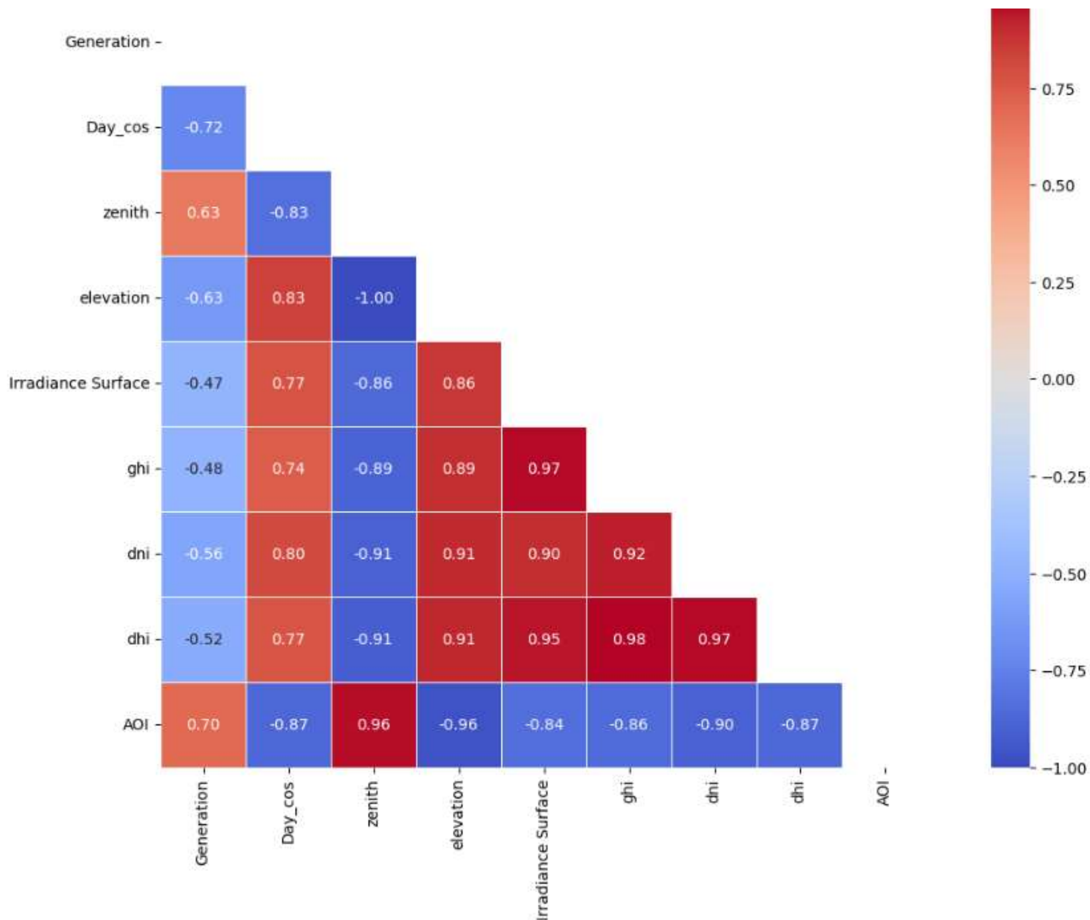

Figure 2. Pearson's correlation heatmap between PV power generation and selected input features

Performance Evaluation of Models for PV Production Prediction

In this study, the performance of three machine learning models i.e. CatBoost, LSTM and GBR were evaluated to forecast PV production. The results of modelling are summarized in Table 3.

Table 2. Evaluation of the Modelling Performance for PV Production Prediction

Model	Dataset	R^2 Score	MAE	MSE	RMSE
CatBoost	Training	0.9367	0.1097	0.0488	0.2208
CatBoost	Test	0.9179	0.1094	0.0535	0.2313
LSTM	Training	0.9270	0.0327	0.0044	0.0661
LSTM	Test	0.9212	0.0324	0.0042	0.0650
GBR	Training	0.7562	0.2287	0.1834	-
GBR	Test	0.7364	0.2325	0.1911	-

As given in Table 2, the CatBoost model demonstrated robust performance, achieving an R^2 score of 0.9367 on the training dataset and 0.9179 on the test dataset. Its error metrics, including the MAE (0.1094) and RMSE (0.2313) on the test set, indicate great predictive accuracy and reasonable generalizability, with only a slight performance drop on the test data. This suggested minimal overfitting and strong applicability for PV forecasting.

The LSTM model exhibited superior performance, with an R^2 score of 0.9270 on the training dataset and 0.9212 on the test dataset. Notably, its error metrics were exceptionally low, with an MAE of 0.0324, an MSE of 0.0042, and an RMSE of 0.0650 on the test set. These results highlighted the remarkable precision and excellent generalizability of LSTM, positioning it as the most effective model in this study.

In contrast, the GBR model yielded suboptimal results, with an R^2 score of 0.7562 on the training dataset and 0.7364 on the test dataset. Its greater error metrics (MAE of 0.2325 and MSE of 0.1911 on the test set) indicated less prediction accuracy compared with CatBoost and LSTM. The absence of RMSE values for GBR limited a complete comparison, but its overall performance suggested that it is less suitable for this application.

In conclusion, the LSTM model outperformed both CatBoost and GBR in terms of accuracy and generalizability, making it the preferred choice for PV production forecasting. CatBoost remains a viable alternative with strong performance, whereas GBR's greater errors and smaller R^2 scores render it less effective. Future work could focus on optimizing the LSTM model or incorporating additional features to further increase the prediction accuracy.

Model validation

The performance evaluation of the developed LSTM model was conducted on the both training and test datasets. As given in Table 3, the results indicated that the model achieved great predictive accuracy with minimal error across all the evaluation metrics.

Specifically, the model obtained an R^2 score of 0.9270 on the training data and 0.9212 on the test data, demonstrating a strong ability to capture the underlying patterns in PV energy production. Furthermore, the values of the MAE, MSE, and RMSE were consistently small and closely matched between the training and testing phases.

The close alignment of performance metrics across the both datasets could be confirmed that the model is not overfitting and generalizes well to unseen data. This would validate the robustness and reliability of the LSTM-based forecasting framework.

Table 3. The Developed Model Evaluation Metrics

Metric	Train	Test
R^2 Score	0.9270	0.9212
MAE	0.0327	0.0324
MSE	0.0044	0.0042
RMSE	0.0661	0.0650

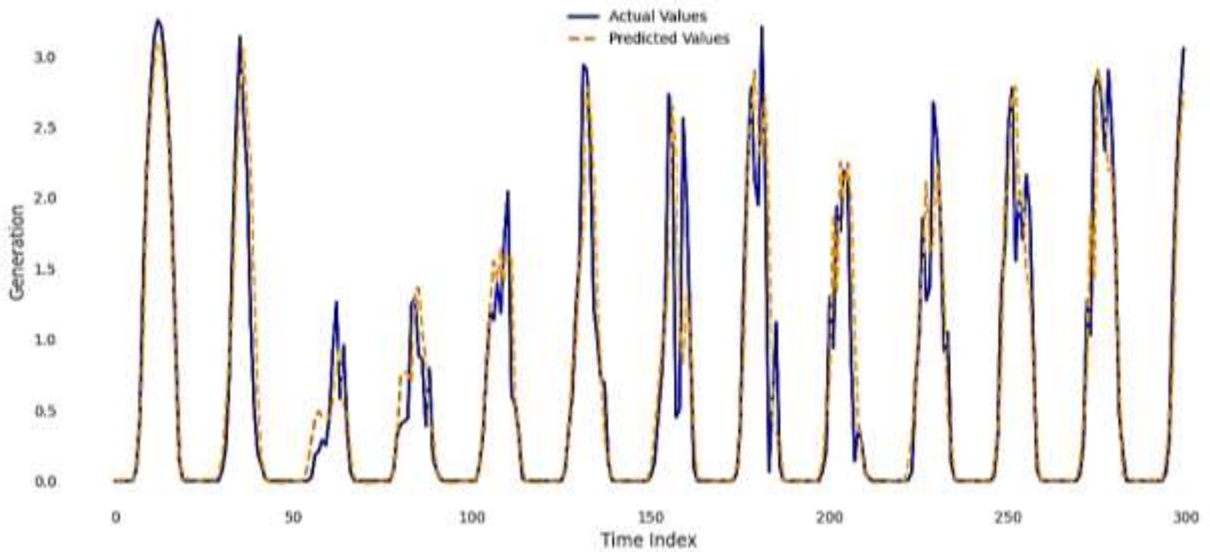

Figure 3. Comparison Between Predicted and actual PV output for 300 test samples

Analysis of LSTM Model Performance through Loss Curves

The performance of the LSTM model in predicting PV production would be further evaluated through the analysis of training and validation loss curves (Figure 4). The evolution of MSE loss over 30 epochs for the both training and validation datasets are depicted on the Figure 4. Initially, both training and validation losses sharply declined, which indicated suitable model convergence during the early stages of training. The training loss decreased from approximately 0.016 to a stable value of approximately 0.005, reflecting effective learning from the training data. Similarly, the validation loss decreased from an initial high value to stabilize at approximately 0.004–0.005 after approximately 10 epochs, which was revealed robust generalizability to unseen data.

The close alignment between the training and validation loss curves after the initial convergence phase indicated minimal overfitting, a critical factor for the reliability of the model in real-world applications. The slight fluctuations in the validation loss toward the later epochs (beyond 20 epochs) were minor and did not suggest divergence, further supporting the model's stability. These results are also corroborated the quantitative metrics previously reported, including a R^2 score of 0.9212 and a RMSE of 0.0650 on the test dataset, which could underscore LSTM's superior predictive accuracy and its suitability for PV production forecasting.

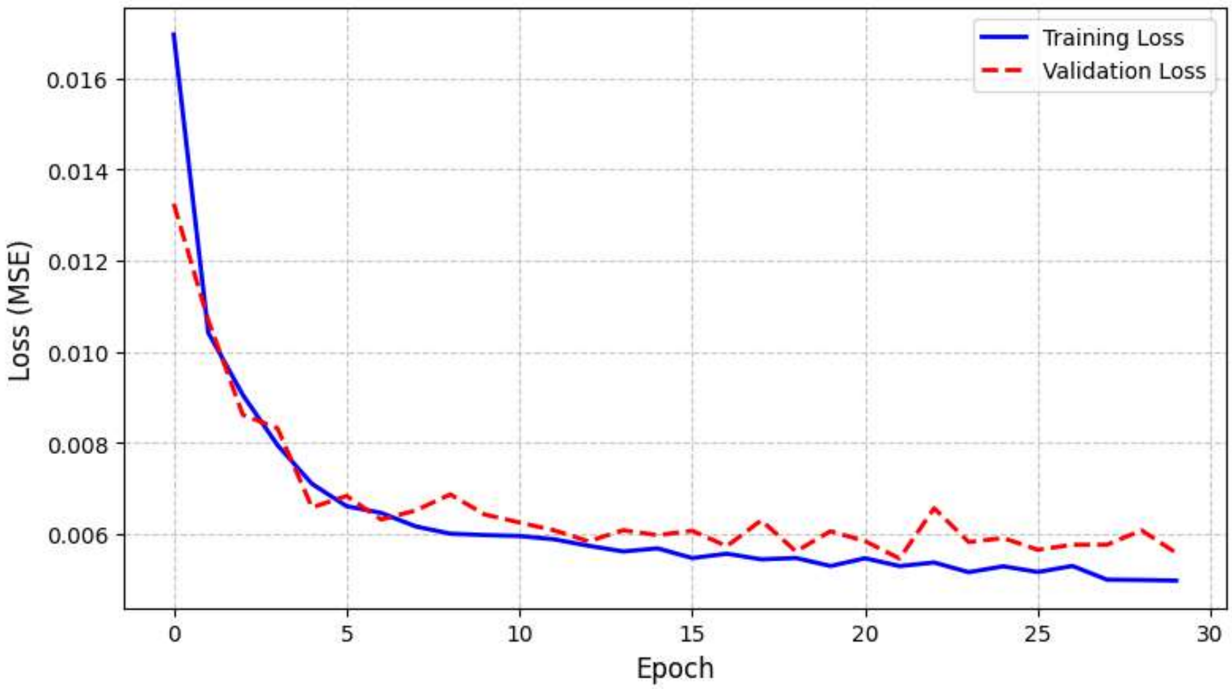

Figure 4. Training and validation loss curves for the LSTM model

Comparative Analysis with Recent Studies

The performance of the developed LSTM model for hourly PV power forecasting in Sydney was compared with that of recent models in the literature to assess the effectiveness of the model (Table 4). The developed model achieved a R^2 score of 0.9212 and a RMSE of 0.0650 on the test dataset, reflecting strong predictive accuracy and generalizability.

Compared with basic LSTM approaches such as that of Suyambu et al., which reported a R^2 of 0.866 and a RMSE of 0.0544 for energy load forecasting, the current model showed a greater correlation and comparable error metrics (Suyambu et al., 2024). While hybrid and attention-based models such as those by Abdullah et al. and Wu et al. achieve greater R^2 scores (above 0.99), they relied on more complex architectures (for example attention layers and dual decomposition) that significantly increased computational cost and implementation complexity (Abdullah et al., 2024; Wu et al., 2024).

The simplicity of the proposed LSTM model could be a key advantage, offering robust and consistent performance across the training and test phases without reliance on hybrid mechanisms. This balance between accuracy and efficiency made it well suited for real-time solar forecasting applications, especially in resource-constrained environments. These findings confirmed that while hybrid models might achieve marginally greater accuracy, a well-tuned standalone LSTM could provide a strong trade-off between performance and simplicity.

Table 4. Comparative Performance of Hourly PV Forecasting Models

Study	Model	R^2	MAE	RMSE	Notes
This Study	LSTM	0.9212	0.0324	0.0650	Accurate, generalizable, and computationally efficient
(Sharma et al., 2022)	LSTM	0.931	–	–	Focused on HVAC energy use prediction
(Abdullah et al., 2024)	Hybrid LSTM + Attention	0.992	0.007	0.012	High accuracy due to attention mechanisms
(Suyambu et al., 2024)	LSTM	0.866	0.0431	0.0544	Basic LSTM for load forecasting
(Wu et al., 2024)	BiGRU-BiLSTM (Hybrid)	0.9922	–	–	Uses dual time-domain decomposition
(Mazen et al., 2023)	GRU–TFT	–	1.19	1.44	Advanced loss functions applied to solar power forecasting

Assessment of LSTM Model Predictive Accuracy via Regression Analysis

The predictive capability of the LSTM model for PV production forecasting was rigorously assessed through a regression analysis of the test dataset, as illustrated in Figure 5. This scatter plot compared the actual PV generation values against the predicted values, with a regression line superimposed to indicate the degree of fit. The R^2 of 0.9212 reflected a strong linear relationship between the predicted and actual generation values, underscoring the model's great accuracy. The tight clustering of data points around the regression line further validated the model's consistency and reliability across the test set. This visual evidence is aligned with the quantitative metrics previously reported, including a MAE of 0.0324 and a RMSE of 0.0650, confirming the robust performance of the LSTM. The slight deviations from the regression line are minimal, suggesting effective generalization and minimal bias in the predictions.

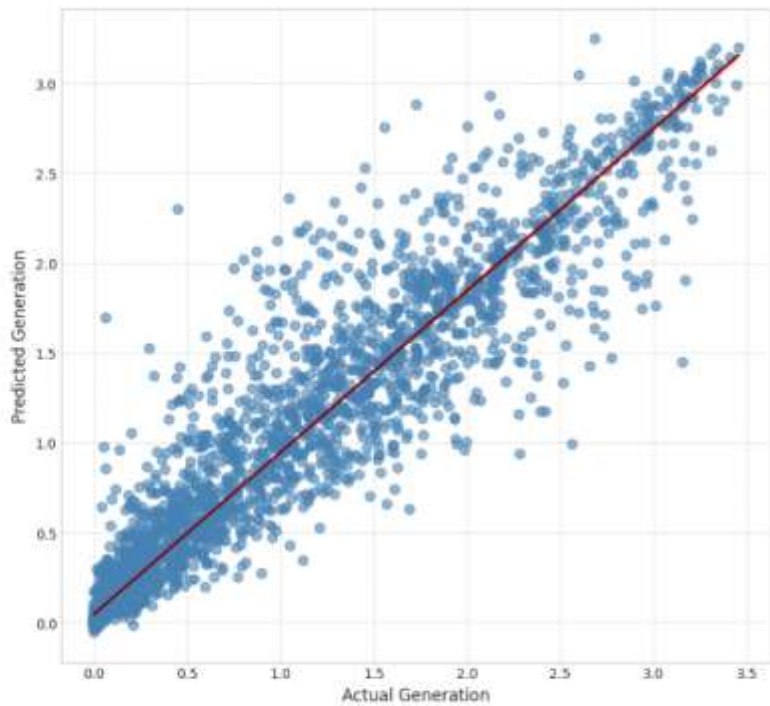

Figure 5. Comparison between the actual and predicted data through regression line ($R^2 = 0.9212$)

CONCLUSIONS

This study advanced short-term PV power forecasting by developing and validating a computationally efficient LSTM-based model tailored to meet the operational demands of renewable energy integration. Through meticulous data curation and rigorous feature selection, the model attained outstanding performance metrics a R^2 score of 0.9212 and a RMSE of 0.0650 on unseen data surpassing benchmark models such as CatBoost and GBR.

These results affirmed the model's ability to capture temporal dependencies in PV generation data while maintaining computational efficiency, which could render it ideal for real-time energy management applications. Nonetheless, limitations including reliance on high-quality input data and untested robustness under extreme weather conditions, warrant further investigation. Future research could enhance the model by incorporating probabilistic forecasting, lightweight attention mechanisms, or transfer learning to improve adaptability across diverse geographic and climatic contexts. Ultimately, this research work contributed a robust, practical tool to the domain of smart grid technologies, supporting the global transition to sustainable energy systems with enhanced forecasting precision and scalability.

Data and Code Availability

The dataset utilized in this study, along with a summarized version of the implementation code, is publicly available in the author's GitHub repository at:

https://github.com/SohaSamiii

The repository, entitled *"Simple LSTM-based Hourly Photovoltaic Power Forecasting"*, contains the preprocessed data files, model architecture, training pipeline, and evaluation scripts. This open-source release aims to promote transparency, reproducibility, and future research in the domain of solar energy forecasting using deep learning models.

REFERENCES

Abdullah, W., Elmasry, A., & Tolba, A. (2024). Hybrid attention-enhanced deep learning for accurate hourly energy consumption forecasting. *Information Sciences* *with Applications*, *3*, 74-83. https://doi.org/10.61356/j.iswa.2024.3314

Ahmed, I., Khan, M. A. U. H., Islam, M., Hasan, M. S., Jakir, T., Hossain, A., Abed, J., Hasanuzzaman, M., Shatyi, S. S., & Hasnain, K. N. (2025). Optimizing Solar Energy Production in the USA: Time-Series Analysis Using AI for Smart Energy Management. *Journal of Posth umanism*, *5*(6), 3396-3423. https://doi.org/10.63332/joph.v5i6.2457

Alexakos, A., Amaxilatis, D., & Zaroliagis, C. (2022). Photovoltaic energy production forecasting and operational analytics: a real-world study. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)* (pp. 439-444). IEEE.

An, Y., Dang, K., Shi, X., Jia, R., Zhang, K., & Huang, Q. (2021). A probabilistic ensemble prediction method for PV power in the nonstationary period. *Energies*, *14*(4), 859. https://doi.org/10.3390/en14040859

Araya, F., Long-Ha, D., Eddine, M. D., & Al Shakarchi, F. (2023, October). Optimal Energy Management System Using Probabilistic Day-ahead Forecasting. In *2023 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)* (pp. 1-5). IEEE.

Balal, A. T., Jafarabadi, Y. P. T., Demir, A. T., Igene, M. T., Giesselmann, M. T., & Bayne, S. T. (2023). Forecasting solar power generation utilizing machine learning models in Lubbock. *Emerging Science Journal*, *7*(4), 1052-1062. https://doi.org/10.28991/ESJ-2023-07-04-02

Bracale, A., Carpinelli, G., & De Falco, P. (2016). A probabilistic competitive ensemble method for short-term photovoltaic power forecasting. *IEEE Transactions on Sustainable Energy*, *8*(2), 551-560. https://doi.org/10.1109/TSTE.2016.2610523

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623

Di Leo, P., Ciocia, A., Malgaroli, G., & Spertino, F. (2025). Advancements and Challenges in Photovoltaic Power Forecasting: A Comprehensive Review. *Energies*, *18*(8), 2108. https://doi.org/10.3390/en18082108

Fan, Y., Ma, Z., Tang, W., Liang, J., & Xu, P. (2024). Using crested Porcupine optimizer algorithm and CNN-LSTM-Attention model combined with deep learning methods to enhance short-term power forecasting in PV generation. *Energies*, *17*(14), 3435. https://doi.org/10.3390/en17143435

Fraga-Hurtado, I., Gómez-Sarduy, J. R., García-Sánchez, Z., Hernández-Herrera, H., Silva-Ortega, J. I., & Reyes-Calvo, R. (2025). Advanced Multivariate Models Incorporating Non-Climatic Exogenous Variables for Very Short-Term Photovoltaic Power Forecasting. *Electricity*, *6*(2), 29. https://doi.org/10.3390/electricity6020029

Hapsari, G. I., Munadi, R., Erfianto, B., & Irawati, I. D. (2025). Feature Selection Using Pearson Correlation for Ultra-Wideband Ranging Classification. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *9*(2), 209-217. https://doi.org/10.29207/resti.v9i2.6281

Izdebski, W., & Kosiorek, K. (2023). Analysis and evaluation of the possibility of electricity production from small photovoltaic installations in Poland. *Energies*, *16*(2), 944. https://doi.org/10.3390/en16020944

Jailani, N. L. M., Dhanasegaran, J. K., Alkawsi, G., Alkahtani, A. A., Phing, C. C., Baashar, Y., Capretz, L. F., Al-Shetwi, A. Q., & Tiong, S. K. (2023). Investigating the power of LSTM-based models in solar energy forecasting. *Processes*, *11*(5), 1382.

Jakoplić, A., Franković, D., Havelka, J., & Bulat, H. (2023). Short-term photovoltaic power plant output forecasting using sky images and deep learning. *Energies*, *16*(14), 5428. https://doi.org/10.3390/en16145428

Liu, C., Li, M., Yu, Y., Wu, Z., Gong, H., & Cheng, F. (2022). A review of multitemporal and multispatial scales photovoltaic forecasting methods. *IEEE access*, *10*, 35073-35093. https://doi.org/10.1109/ACCESS.2022.3162206

Mazen, F. M. A., Shaker, Y., & Abul Seoud, R. A. (2023). Forecasting of solar power using GRU–temporal fusion transformer model and DILATE loss function. *Energies*, *16*(24), 8105. https://doi.org/10.3390/en16248105

Praveenraj, D. D. W., Madeswaran, A., Pastariya, R., Sharma, D., Abootharmahmoodshakir, K., & Dhablia, A. (2024). Machine learning integration for enhanced solar power generation forecasting. In *E3S Web of Conferences* (Vol. 540, p. 04007). EDP Sciences. https://doi.org/10.1051/e3sconf/202454004007

Rezvani, Z., Rezvani, F., & Arslan, S. (2022). Designing, Simulating and Technical Analysis of a 2 MW On-grid Photovoltaic System for Agricultural Applications. *Biomechanism and Bioenergy Research, 1*(2), 1-6. https://doi.org/10.22103/bbr.2022.20681.1036

Robeson, S. M., & Willmott, C. J. (2023). Decomposition of the mean absolute error (MAE) into systematic and unsystematic components. *PloS one, 18*(2), e0279774. https://doi.org/10.1371/journal.pone.0279774

Sabri, M., & El Hassouni, M. (2023). Photovoltaic power forecasting with a long short-term memory autoencoder networks. *Soft Computing, 27*(15), 10533-10553. https://doi.org/10.1007/s00500-023-08497-y

Salman, D., Direkoglu, C., Kusaf, M., & Fahrioglu, M. (2024). Hybrid deep learning models for time series forecasting of solar power. *Neural Computing and Applications, 36*(16), 9095-9112. https://doi.org/10.1007/s00521-024-09558-5

Sharma, J., Soni, S., Paliwal, P., Saboor, S., Chaurasiya, P. K., Sharifpur, M., Khalilpoor, N., & Afzal, A. (2022). A novel long term solar photovoltaic power forecasting approach using LSTM with Nadam optimizer: A case study of India. *Energy Science & Engineering, 10*(8), 2909-2929.

Sleiman, A., & Su, W. (2024). Combined K-means clustering with neural networks methods for PV short-term generation load forecasting in electric utilities. *Energies, 17*(6), 1433. https://doi.org/10.3390/en17061433

Stoliarov, O. (2024). Big Data technologies in the process of forecasting electricity generation from solar photovoltaic power plants. *Вісник Черкаського державного технологічного університету. Технічні науки, 29*(2). 79-92.https://doi.org/10.62660/bcstu/2.2024.79

Suyambu, M. R., Vishwakarma, P. K., & Shrivastava, V. (2024, December). Improving Energy Efficiency Through Machine Learning-Based Load Forecasting in power systems. In *2024 International Conference on Communication, Control, and Intelligent Systems (CCIS)* (pp. 1-6). IEEE. https://doi.org/10.1109/CCIS63231.2024.10931928

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research, 30*(1), 79-82. https://doi.org/10.3354/cr030079

Wu, Y., Gong, G., & Zhu, C. (2024, September). Electricity demand forecasting based on hybrid BiGRU-BiLSTM model prediction framework with dual time domain decomposition. In *2024 5th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)* (pp. 789-793). IEEE. https://doi.org/10.1109/ICBASE63199.2024.10761991

Yadav, S., Rajput, P., Balasubramanian, P., Liu, C., Li, F., & Zhang, P. (2025). Machine learning-driven prediction of biochar adsorption capacity for effective removal of Congo red dye. *Carbon Research, 4*(1), 11. https://doi.org/10.1007/s44246-024-00168-3

Zhao, Z., Zhai, M., Li, G., Gao, X., Song, W., Wang, X., Ren, H., Cui, Y., Qiao, Y., & Ren, J. (2023). Study on the prediction effect of a combined model of SARIMA and LSTM based on SSA for influenza in Shanxi Province, China. *BMC Infectious Diseases, 23*(1), 71. https://doi.org/10.1186/s12879-023-08025-1