International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Journal Européen des Systèmes Automatisés
Vol. 58, No. 11, November, 2025, pp. 2331-2336
Journal homepage: http://iieta.org/journals/jesa

Machine Learning Based Weather Prediction for Smart Battery Switching of PV System to Enhanced Reliability Optimization

Wahyu, Syafii*, Darwison

Department of Electrical Engineering, Andalas University, Padang 25163, Indonesia

Corresponding Author Email: syafii@eng.unand.ac.id

https://doi.org/10.18280/jesa.581111

Received: 2 August 2025
Revised: 9 September 2025
Accepted: 17 September 2025
Available online: 30 November 2025

Keywords:
solar energy, ESP32, LSTM, Internet of Things, weather prediction

ABSTRACT

The study developed an LSTM-based weather prediction system to improve the efficiency of off-grid solar photovoltaic (PV) systems. The system integrates BMKG historical data (2019-2024) and real-time monitoring data from IoT sensors to predict two main weather parameters: temperature and Global Horizontal Irradiance (GHI). LSTM models are trained to forecast weather conditions, and the results of these predictions are used to optimize energy use. Although using only 3 days of real-time data, the model demonstrates high accuracy, with MAPEs of 9.14% for GHI and 6.08% for temperature. The use of multi-source data enables the model to handle weather fluctuations better. Although there are some minor deviations due to rapid weather changes, the results of this study provide a solid basis for managing solar energy and optimizing energy storage in off-grid PV systems. The study also paves the way for the development of models that are more sensitive to local weather changes and the use of longer data to improve prediction accuracy.

1. INTRODUCTION

The global energy transition towards cleaner and more sustainable renewable energy sources is increasingly urgent to address the energy crisis. Solar energy is one of the important solutions to reduce dependence on fossil energy and overcome the deepening energy crisis [1-4]. However, the intermittent nature of solar energy and dependence on weather conditions pose a significant challenge, especially for off-grid systems that are not connected to the primary grid [5, 6]. These fluctuations in energy supply, caused by weather instability, directly impact the reliability of the electricity supply, which requires innovative solutions to manage these unstable energies [7-9]. Effective energy management is crucial to ensure that solar energy sources can be utilized optimally, despite rapid changes in weather conditions.

With advances in deep learning technology, the Long Short-Term Memory (LSTM) model has proven to be effective in processing time series data, including weather forecasts, due to its ability to recall important information from sequential data [10-12]. The LSTM model is particularly suitable for handling weather instability due to its ability to predict long-term fluctuations and provide high accuracy in predicting complex weather patterns. In recent years, LSTMs have been widely used to improve the efficiency and reliability of solar energy systems, including off-grid PV systems that are highly dependent on weather conditions [13-15]. However, while the application of LSTM is promising, many studies rely on a single data point from weather stations that are often not accurate enough in describing the microclimate conditions at the location of solar panels. In addition, many studies have focused solely on the accuracy of the model without providing practical solutions for managing solar Energy affected by weather fluctuations.

To address this research gap, we propose two clear scientific contributions. First, we integrate multi-source data by combining BMKG historical data (2019-2024) and real-time data obtained through IoT sensors on the Arduino Cloud to improve the accuracy of weather forecasts. The incorporation of BMKG's historical data, which provides long-term context and real-time data that is more responsive to weather changes, allows the LSTM model to predict weather conditions more accurately and promptly. Second, we developed an optimal adaptive control system that leverages weather prediction results to manage backup batteries automatically [16, 17], thereby reducing the impact of intermittent energy on off-grid PV systems. By integrating historical and real-time data, the system is expected to optimize energy storage and ensure a stable energy supply, even in the face of rapidly changing weather conditions.

Thus, this research aims not only to enhance the accuracy of weather predictions but also to offer practical solutions that improve the efficiency and reliability of solar energy systems. The developed system is expected to provide a more adaptive solution in managing solar energy based on changing weather conditions, by optimizing energy use through accurate weather predictions. This research also paves the way for the development of more sustainable systems by utilizing multiple data sources to enhance the accuracy of LSTM models in predicting weather fluctuations.

2. METHODOLOGY

The study proposes a system that integrates IoT and machine learning technologies to optimize solar energy management in off-grid systems. The methodology employed includes real-time data collection, predictive model development, and energy management processes informed by weather predictions. This entire process is designed to ensure the reliability of the energy supply and efficient energy management based on predicted weather conditions.

2.1 Block diagram of a weather prediction system using LSTM for off-grid solar energy

Figure 1 illustrates a block diagram of the proposed system's overall architecture. Real-time weather data, including solar radiation intensity and temperature (collected using a Pyranometer and DHT11 sensors), is gathered by the ESP32 microcontroller. This data is then sent to a cloud server over an internet connection. At the same time, historical data from BMKG (2019-2024) is also integrated to enrich the dataset. The merging of these two data sources enables LSTM models to enhance the accuracy of weather predictions by leveraging historical data that provides long-term context and real-time data that facilitates a faster response to weather changes.

2.2 Weather prediction flow diagram and solar energy management process using LSTM

The flowchart illustrates the detailed steps involved in data collection, model training, and energy management, all of which are based on weather predictions. Figure 2 provides a detailed description of the system's workflow. The process begins with the collection of weather data from the Pyranometer sensor and the DHT11 sensor, as well as the BMKG server. The data is then normalized to ensure scale consistency before being fed into the LSTM model. Once the LSTM model is trained using historical and real-time data, it generates weather predictions that are used for energy analysis.

A significant change here is that energy management no longer involves physical relays, but instead is based on weather predictions generated by LSTM models. For example, if the weather prediction shows sunny weather, the LSTM model indicates that the energy from the solar panels can be used optimally. Conversely, if the prediction indicates bad weather, the LSTM model suggests that Energy from the backup battery should be used to maintain the sustainability of the energy supply. This process will continue to take place adaptively, with models that update weather predictions and adjust energy usage according to changing weather conditions.

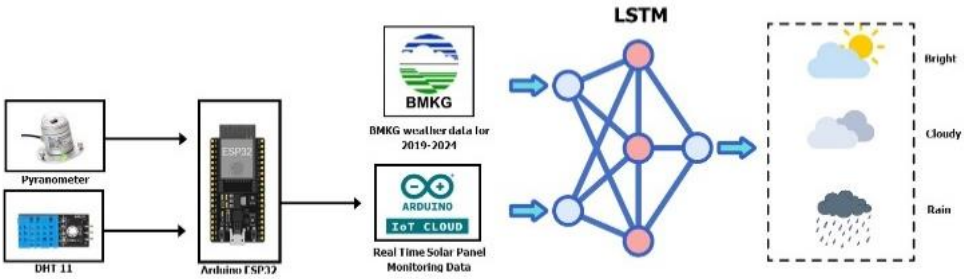

Figure 1. Weather prediction block diagram

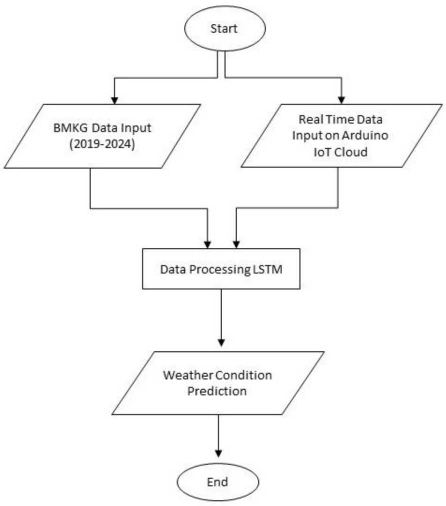

Figure 2. Weather prediction flowchart with LSTM

2.3 Data collection

Data collection is a critical initial stage in this study, where two main types of data are used, namely BMKG historical data and real-time IoT data. BMKG data includes temperature and Global Horizontal Irradiance (GHI) taken from BMKG weather stations for the period 2019 to 2024. This historical data is used to provide long-term weather context that is critical in studying overall weather patterns. In addition, the IoT sensors used in this study are a Pyranometer (to measure solar radiation) and a DHT11 (to measure temperature and humidity). The data from this sensor is stored in an Excel file for further processing.

BMKG historical data and IoT sensor data are utilized to train the LSTM model, aiming to produce accurate weather predictions. The unification of these two data sources enables LSTM models to examine long-term weather patterns and address the rapid weather changes that often occur in real-time data. This data is then processed using Visual Studio Code with the Python programming language for further processing.

2.4 Data processing

The dataset used in this study consists of historical meteorological data (2019–2024) and real-time measurements of solar radiation and temperature collected from August 25 to 28, 2025. For model training and testing, chronological separation is applied. Specifically, data from August 25–27, 2025, was used for training, while data from August 28, 2025, served as a test set. This design ensures that the model learns from past observations and is evaluated on invisible future data, which better reflects real-world forecasting scenarios.

To prevent information leakage, normalization is performed using statistics calculated solely from the training set, as per Eq. (1).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Once the data is normalized, the BMKG historical data and real-time data from IoT are combined in a single dataset, which is then used as input for the LSTM model. This incorporation enables the model to analyze long-term weather patterns using BMKG data and make informed decisions based on real-time weather data from IoT sensors. All of this data processing is done using Visual Studio Code.

2.5 LSTM model development

LSTM is a type of Recurrent Neural Network (RNN) used to process complex time series data, such as weather data. In this study, LSTM was employed due to its ability to analyze long-term relationships in sequential data, which is crucial for predicting weather fluctuations. LSTM works by controlling the flow of information through three main gates: the input gate, the forget gate, and the output gate [18].

$$f_t = \sigma(W_f \,.\, [h_{t-1}, x_t] + b_f) \quad (2)$$

Gate forgetfulness controls information that needs to be forgotten based on previous hidden inputs and statuses.

$$i_t = \sigma(W_i \,.\, [h_{t-1}, x_t] + b_i) \quad (3)$$

The input gate determines how much new information will be stored in the cell's state.

$$o_t = \sigma\, (W_o \,.\, [h_{t-1}, x_t] +)b_o \quad (4)$$

$$h_t = o_t \,.\, tanh.\,(\,)C_t \quad (5)$$

The output gate specifies the information to be ejected from the cell.

$$\hat{C}_t = \tanh\, (W_c . [h_{t-1}, x_t] + b_c + C_t = f_t \,.\, C_{t-1} i_t \hat{C}_t \quad (6)$$

The LSTM layer in this model comprises 64 hidden units, which are responsible for processing sequential data (time-order-dependent data) and capturing long-term dependencies in weather data. LSTMs are designed to address the vanishing gradient problem that often occurs in standard RNNs, allowing them to process information over long time series more efficiently. The tanh (hyperbolic tangent) activation function is used in the LSTM layer to control the output values within the range of -1 to 1, which enables the model to handle both large and small values effectively, as well as assist in selecting relevant information to retain or discard during the training process.

After the LSTM layer, a Dense layer with 32 units is present, followed by a ReLU (Rectified Linear Unit) activation function. ReLU is used to introduce non-linearity to the model, allowing LSTM to capture more complex relationships between inputs and outputs. The ReLU function is highly effective in accelerating convergence and mitigating issues that often arise with sigmoid or tanh activation functions, such as vanishing gradients.

This model is trained using the Adam optimizer, which is an adaptive optimization algorithm that automatically adjusts the learning rate during training. This helps the model to evolve faster and converge to the optimal solution in less time. The loss function used is the Mean Squared Error (MSE), which measures the average of the squares of the differences between the predicted value and the actual value, and is used to minimize prediction errors. To improve performance and prevent overfitting, the model is trained with a batch size of 32, which divides the data into small groups to update the model weights more stably. A total of 120 epochs was used to ensure that the model was well-trained and well-generalized against data that had never been seen before.

During the training, real-time data from IoT (Pyranometer and DHT11) and BMKG historical data were used to train the model in predicting two main weather parameters: GHI and temperature. The combination of BMKG's historical data, which provides macro-level information, and real-time IoT data, which is more responsive to rapid weather changes, enables LSTM to study long-term weather patterns better and address rapid weather fluctuations.

The LSTM training process is carried out using Visual Studio Code as the IDE and Python as the programming language. The TensorFlow and Keras libraries are used for implementing this LSTM model, as both offer the convenience of building and training deep learning models with various types of networks, including LSTM.

2.6 Model performance evaluation

The performance of the predictive model was evaluated using standard accuracy metrics, namely Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) [19, 20]. These three metrics provide an overview of how closely the predicted value aligns with the actual value.

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |Prediksi\ (t) - Aktual\ (t)| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N} |Prediksi\ (t) - Aktual\ (t)|} \quad (8)$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N} \left| \frac{Aktual(t) - Prediksi\ (t)}{Aktual\ (t)} \right| \times 100\% \quad (9)$$

These metrics are used to measure how closely the model's predictions align with the actual value and to ensure that the model can properly generalize to previously unseen data. MAPE, for example, provides an overview of the percentage of error between model predictions and actual data, which is very useful for evaluating model performance in weather predictions.

Although direct comparisons with ARIMA were not conducted in this study due to time and data limitations, previous research has consistently demonstrated that LSTM-based models outperform traditional statistical approaches such as ARIMA in solar radiation forecasting. For instance, [21] showed that a hybrid SARIMA–LSTM model significantly reduced forecasting errors compared to standalone ARIMA, indicating the superiority of deep learning-based methods for solar irradiance prediction.

Therefore, although direct comparisons were not made, evidence from previous studies supports the selection of LSTMs as a more effective model for weather prediction and solar irradiation.

3. RESULTS AND DISCUSSION

The study successfully developed an LSTM-based weather forecasting system to enhance the efficiency of solar energy systems, particularly off-grid solar panel systems that heavily rely on weather fluctuations. LSTM models are used to predict two main weather parameters: temperature and GHI, both of which have a direct impact on solar energy production. The results of the experiment demonstrated that the LSTM could predict temperature and GHI with a high degree of accuracy. However, minor deviations occurred at certain times, as illustrated in Figure 3.

The temperature prediction showed satisfactory results, with the LSTM model successfully capturing daily temperature fluctuations. The data used for temperature prediction comes from two primary sources: historical BMKG data (2019–2024) and real-time monitoring data collected over three consecutive days. Combining BMKG's historical data with real-time monitoring data allows the model to study long-term weather patterns and predict temperature changes more accurately. This proved to be effective in reducing prediction errors, which is reflected in the low MAPE value and high accuracy of the temperature prediction results (6.08%), as shown in Table 1.

In contrast, the GHI prediction only utilizes 3-day real-time data, as the BMKG data does not provide information on solar radiation. Nevertheless, the model still managed to predict the pattern well, despite a moderate negative bias between the prediction and the actual data, where the measured average irradiation was about 998 W/m² and the average prediction was 926 W/m², resulting in a difference of about −72 W/m². Figure 4 shows the results of the GHI prediction, where the LSTM model can capture the pattern of daily fluctuations stably. However, the model is slow to lower GHI values after the peak, which is caused by the effects of rapid weather, such as the appearance of clouds that cannot be fully captured in the training data. The slight difference in GHI peaks (about 37 W/m² at 12:28 (measured) vs. 12:36 (predicted)) also reflects how faster-than-predicted fluctuations in solar radiation affect model performance.

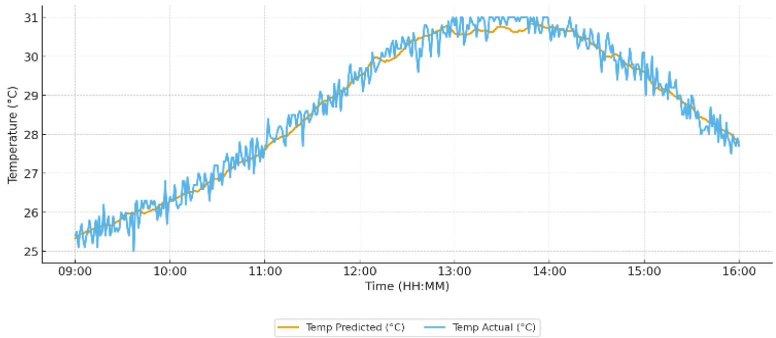

Figure 3. Temperature prediction and actual data graph

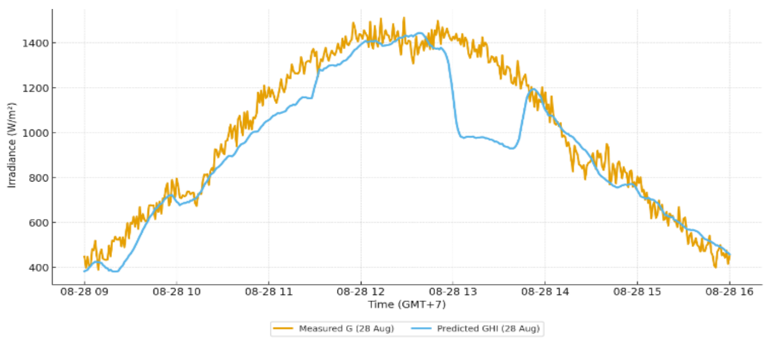

Figure 4. Solar radiation prediction graph and actual data

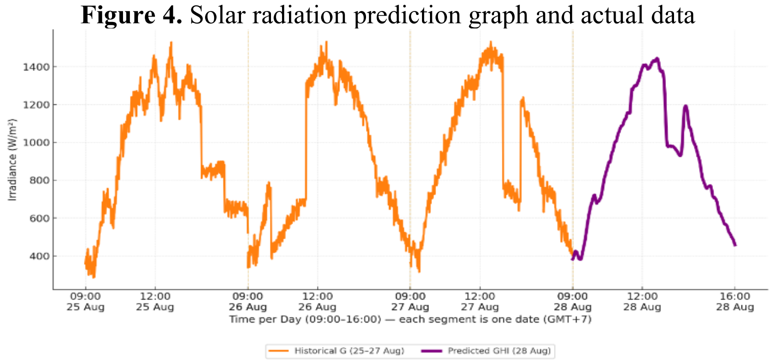

Figure 5. Sun radiation chart: Day and 1 day prediction

Table 1. Forecast error metrics for sun irradiation and temperature

Parameters	MAE	RMSE	MAP
Irradiance	94.14 W/m²	140.18 W/m²	9.14
Temperature	5.12°C	7.85°C	6.08

Based on the results of the weather prediction using the LSTM model and comparing it with the actual data on August 28, 2025, the graph shown in Figure 4 shows a stable pattern of solar radiation with a reasonably consistent peak at around 12:30. The MAPE measured at 9.14% shows that the LSTM model can predict weather conditions with high accuracy. With a low MAPE value, it can be ensured that the weather conditions on that day are sunny. The temperature on that date, as shown in Figure 3, exhibits a consistent temperature pattern throughout the day, which also supports the decision that the weather on August 28, 2025, was sunny. Thus, the weather forecast for August 28, 2025, can be categorized as sunny throughout the day, with a slight change to cloudy in the afternoon.

Figure 5 shows a graph of solar radiation predictions for a day, along with a comparison to the actual measured data. This graph shows that the LSTM can provide stable daily predictions, although there is a decrease in accuracy at certain times when the weather changes rapidly, as seen in the slight differences at the top of GHI. More diverse and extended data can improve model accuracy, and the use of multi-source data allows models to capture more complex weather patterns, such as those seen in predictions of solar radiation.

It is essential to note that collecting increasingly varied data has a significant impact on improving model accuracy. As the amount of data increases, LSTM models are becoming increasingly capable of capturing more complex weather patterns, particularly when historical and real-time data are combined for temperature prediction. It offers a more effective solution for mitigating weather fluctuations and energy intermittency in off-grid solar panel systems. Although solar radiation data is not available in the BMKG dataset, the use of real-time data for three consecutive days has demonstrated that the LSTM model is highly effective in predicting daily fluctuations in solar radiation and providing accurate predictions.

Overall, LSTM has proven helpful in improving weather prediction in the context of solar Energy, with stable results and high accuracy as previous research [21]. Although there are minor deviations in the prediction, especially in the period after the peak of GHI, the model still provides valuable information for managing energy intermittency caused by weather fluctuations. The use of multi-source data enables the LSTM model to analyze long-term weather patterns from BMKG data, providing historical context and facilitating faster decisions based on real-time IoT data. The combination of these two types of data allows the model to capture more complex weather fluctuations, as well as optimize energy management in off-grid PV systems.

Although LSTM can predict temperature and GHI with high accuracy, there are minor deviations at times, especially after the peak of GHI. This is due to the influence of fast weather, such as cloud shifts that cannot be fully predicted by models trained with only limited data (3 days). To refine the model, it is recommended to use longer data and augmented data to help the model better capture faster weather fluctuations.

4. CONCLUSION

This research has successfully developed an LSTM-based weather prediction system to enhance the efficiency of solar energy systems, particularly in off-grid PV systems that are highly dependent on weather fluctuations. The proposed LSTM model can predict temperature and GHI with high accuracy, as reflected in MAPE values of 9.14% for GHI and 6.08% for temperature. Although the 3-day data is used for training and testing, the model performs well. However, the limitations of the data affect the model's generalization ability in the face of faster weather changes.

This study demonstrates that by combining BMKG's historical data with real-time data from IoT sensors, the LSTM model can provide more accurate weather predictions, which aid in the automatic and efficient management of solar Energy. Nonetheless, further improvements are needed to improve the model's sensitivity to rapid local weather changes, such as sudden changes caused by clouds or wind.

One of the key limitations in this study is the use of real-time 3-day data for training and testing, which limits the model's ability to generalize under more varied weather conditions. Therefore, future research will focus on improving datasets, including the extended use of data and the application of augmented data. Additionally, Transformer-based models and multi-target forecasting (including temperature, humidity, and wind) can be explored further to enhance the resilience and generalization of weather prediction systems.

ACKNOWLEDGMENT

This research is supported by Andalas University under the PTM Research Scheme, with contract number 183/UN16.19/PT.01.03/PTM/2025.

REFERENCES

[1] Dhaked, D.K., Dadhich, S., Birla, D. (2023). Power output forecasting of solar photovoltaic plant using LSTM. Green Energy and Intelligent Transportation, 2(5): 100113. https://doi.org/10.1016/j.geits.2023.100113

[2] Beigi, M., Beigi Harchegani, H., Torki, M., Kaveh, M., Szymanek, M., Khalife, E., Dziwulski, J. (2022). Forecasting of power output of a PVPS based on meteorological data using RNN approaches. Sustainability, 14(5): 3104. https://doi.org/10.3390/su14053104

[3] Campos, F.D., Sousa, T.C., Barbosa, R.S. (2024). Short-term forecast of photovoltaic solar energy production using LSTM. Energies, 17(11): 2582. https://doi.org/10.3390/en17112582

[4] Satria, H., Syafii, S. (2018). Sistem monitoring online dan analisa performansi PLTS rooftop terhubung ke grid PLN. Jurnal Rekayasa Elektrika, 14(2): 136-144. https://doi.org/10.17529/jre.v14i2.11141

[5] Sabareesh, S.U., Aravind, K.S.N., Chowdary, K.B., Syama, S., VS, K.D. (2023). LSTM based 24 hours ahead forecasting of solar PV system for standalone household system. Procedia Computer Science, 218: 1304-1313. https://doi.org/10.1016/j.procs.2023.01.109

[6] Syafii, S., Krismadinata, K., Muladi, M., Agung, T.K.,

Sandri, D.A. (2024). Simple photovoltaic electric vehicles charging management system considering sun availability time. Journal of Sustainable Development of Energy, Water and Environmental Systems, 12(1): 1110476. https://doi.org/10.13044/j.sdewes.d11.0476
[7] Mcingani, I., Meyer, E.L., Overen, O.K. (2024). The impact of ambient weather conditions and energy usage patterns on the performance of a domestic off-grid photovoltaic system. Energies, 17(19): 5013. https://doi.org/10.3390/en17195013
[8] Khan, U.A., Khan, N.M., Zafar, M.H. (2023). Resource efficient PV power forecasting: Transductive transfer learning based hybrid deep learning model for smart grid in Industry 5.0. Energy Conversion and Management: X, 20: 100486. https://doi.org/10.1016/j.ecmx.2023.100486
[9] Yu, Y., Cao, J., Zhu, J. (2019). LSTM short-term solar irradiation forecast under complex weather conditions. IEEE Access, 7: 145651-145666. https://doi.org/10.1109/ACCESS.2019.2946057
[10] Ozdemir, T., Taher, F., Ayinde, B.O., Zurada, J.M., Tuzun Ozmen, O. (2022). Comparison of feedforward perceptron network with LSTM for solar cell radiation prediction. Applied Sciences, 12(9): 4463. https://doi.org/10.3390/app12094463
[11] Kuo, W.C., Chen, C.H., Hua, S.H., Wang, C.C. (2022). Assessment of different deep learning methods of power generation forecasting for solar PV system. Applied Sciences, 12(15):7529. https://doi.org/10.3390/app12157529
[12] Konstantinou, M., Peratikou, S., Charalambides, A.G. (2021). Solar photovoltaic forecasting of power output using LSTM networks. Atmosphere, 12(1): 124. https://doi.org/10.3390/atmos12010124
[13] Nepal, R.K., Khanal, B., Ghimire, V., Neupane, K., et al. (2024). Solar power predictions using satellite data in different parts of Nepal. http://arxiv.org/abs/2406.11877. https://doi.org/10.48550/arXiv.2406.11877
[14] Alam, F., Islam, M., Deb, A., Hossain, S.S. (2024). Comparison of deep learning models for weather forecasting in different climate zones. Journal of Computer Science and Engineering (JCSE), 5(1): 12-19. https://doi.org/10.36596/jcse.v5i1.729
[15] Qasim, O.A., Alhayani, M.Y., Noori, M.S. (2025). Deep learning-based weather prediction: A focused case study in the City of Mosul. Asian Journal of Research in Computer Science, 18(4): 176-186. https://doi.org/10.9734/ajrcos/2025/v18i4614
[16] Syafii, S., Rusydi, M.I., Son, L., Zikri, I. (2020). Web-based net energy meter for grid connected PV system. TEM Journal, 9(1): 37-41.
[17] Wu, X., Li, S., Gan, S., Hou, C. (2022). An adaptive energy optimization method of hybrid battery-supercapacitor storage system for uncertain demand. Energies, 15(5): 1765. https://doi.org/10.3390/en15051765
[18] Sarmas, E., Spiliotis, E., Stamatopoulos, E., Marinakis, V. (2023). Short-term photovoltaic power forecasting using meta-learning and independent LSTM models of numerical weather prediction. Renewable Energy, 212: 1-15. https://doi.org/10.1016/j.renene.2023.07.112
[19] Sharma, J., Soni, S., Paliwal, P., Saboor, S., Chaurasiya, P.K., Sharifpur, M., Khalilpoor, N., Afzal, A. (2022). A new long-term photovoltaic solar forecasting approach using LSTMs with Nadam optimizers: An Indian case study. Energy Science & Engineering, 10(8): 2909-2929. https://doi.org/10.1002/ese3.1178
[20] Syafii, S., Izrillah, I.N., Aulia, A., Rusydi, M.I. (2025). An internet of things-based weather system for short-term solar and wind power forecasting using double moving average. Bulletin of Electrical Engineering and Informatics, 14(5): 3309-3317. https://doi.org/10.11591/eei.v14i5.9852
[21] Yesil, F.N., Serttas, T.N. (2025). A novel approach to solar radiation forecasting using a SARIMA-LSTM hybrid model. Neural Computing and Applications, 37: 20069-20103. https://doi.org/10.1007/s00521-025-11442-9

NOMENCLATURE

$T(t)$	Temperature at time t
$H(t)$	Humidity in time
$W(t)$	Wind speed at time t
X_{norm}	Normalized data values
t	Time
$t-1$	Previous time steps
X	Actual data value
X_{min}	Data threshold
X_{max}	Maximum data value
f_t	Forget the gate
i_t	Input gateway
o_t	Output gateway
C_t	Cell status
h_{t-1}	Hidden status of previous time steps
x_t	Input on the current time step
MAE	Average Absolute Error
RMSE	Root Mean Square Error
MAPE	Average Absolute Error Percentage

Greek symbols

σ	Sigmoid function
Tanh	Hyperbolic tangent activation function