JST: Smart Systems and Devices
Volume 35, Issue 2, May 2025, 054-061

A Comparative Study on the Operational Effectiveness of Machine Learning Models in Solar Power Forecasting

Manh-Hai Pham[1*], Nguyen Tuan Anh[1], Vu Minh Phap[2], Pham Van Duy[2], Le Thanh Doanh[1], Vu Thi Anh Tho[1], Nguyen Dang Toan[1]

[1]Electric Power University, Ha Noi, Vietnam
[2] Institute of Science and Technology for Energy and Environment, Vietnam Academy of Science and Technology, Ha Noi, Vietnam
[*]Corresponding author email: haipm@epu.edu.vn

Abstract

Accurate solar power forecasting is crucial for optimizing grid operations and balancing energy supply and demand. Due to the high variability of solar radiation, advanced machine learning methods are needed to enhance forecasting accuracy. This study compares three models: Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and a single-hidden-layer Bidirectional Gated Recurrent Unit (BiGRU). XGBoost and LightGBM are decision tree-based boosting models known for fast training and high accuracy, while BiGRU is a recurrent neural network designed for time-series data but prone to overfitting. Experimental results show that XGBoost and LightGBM train significantly faster and achieve lower errors (Normalized Mean Absolute Percentage Error-NMAPE is lower than 5%), demonstrating superior generalization. In contrast, BiGRU exhibits overfitting with NMAPE equal to 23.986% and Root Mean Squared Error (RMSE) equal to 18,763.12 kW on June 30, 2021. Notably, on December 31, 2021, XGBoost and LightGBM closely followed actual power generation trends, whereas BiGRU struggled to capture variations, further indicating its generalization issues. The findings highlight XGBoost and LightGBM as more suitable models for solar power forecasting, providing valuable insights for researchers and engineers in power grid management.

Keywords: Solar power forecasting, XGBoost, LightGBM, BiGRU, overfitting, execution time

1. Introduction

Solar power generation forecasting is essential for managing and operating renewable energy systems [1]. As the share of renewable energy sources in the power supply structure increases, ensuring stability and optimizing energy distribution become critical requirements [2]. Despite being a clean and abundant energy source, solar power is highly variable, as it depends on weather factors such as solar radiation levels, temperature, humidity, and seasonal changes [3]. This instability can cause imbalances between electricity supply and demand, reduce power grid efficiency, and increase the risk of local power shortages. Therefore, an accurate solar power forecasting system can help system operators make informed decisions regarding resource allocation, reduce operational costs, and enhance the reliability of the power system.

Traditional forecasting methods, such as linear regression and ARIMA models, are simple and easy to implement but are often ineffective when applied to systems with high variability and non-linearity. In recent years, modern machine learning algorithms have demonstrated superior forecasting capabilities due to their ability to learn and extract complex relationships between input features. Among these, decision tree-based boosting techniques such as Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) [4], as well as deep learning models such as Bidirectional Gated Recurrent Unit (BiGRU), are expected to significantly improve solar power forecasting accuracy.

This study selects and compares three forecasting models: XGBoost, LightGBM, and BiGRU (1 hidden layer). XGBoost and LightGBM are boosting models based on decision trees, known for their efficient data processing, fast training speed, and good generalization ability [4]. Meanwhile, BiGRU is a variant of the Gated Recurrent Unit (GRU), which can better capture time-series relationships in data but requires more computational resources and is susceptible to overfitting if not properly optimized.

Recent studies have highlighted the effectiveness of XGBoost and LightGBM in solar power forecasting, particularly due to their fast training speed and high accuracy. Unlike conventional machine

ISSN 2734-9373
https://doi.org/10.51316/jst.182.ssad.2025.35.2.7
Received: Feb 28, 2024; revised: Apr 18, 2025
accepted: Apr 18, 2025

learning models, XGBoost optimizes training by employing parallel computation and regularization techniques, reducing overfitting and improving generalization. Similarly, LightGBM introduces histogram-based feature selection and Gradient-based One-Side Sampling (GOSS), making it even faster than XGBoost while maintaining comparable accuracy [5]. These characteristics make both models well-suited for renewable energy forecasting, where rapid computation and adaptability to fluctuating weather conditions are crucial.

While XGBoost and LightGBM provide high accuracy with fast processing speeds, deep learning models like BiGRU take a different approach by capturing sequential dependencies in time-series data. BiGRU utilizes a bidirectional structure, allowing it to learn both past and future temporal relationships, However, this design also significantly increases the number of trainable parameters, leading to higher memory consumption and greater sensitivity to noise. BiGRU is often prone to overfitting, particularly when dealing with highly variable datasets. The challenge lies in balancing short-term fluctuations with long-term trends, where excessive reliance on historical data can introduce noise and reduce the model's generalization ability [6]. This limitation suggests that BiGRU requires careful optimization, including appropriate data size selection and regularization techniques, to be effectively applied in dynamic forecasting.

Given these observations, this study aims to compare XGBoost, LightGBM, and BiGRU in the context of solar power forecasting, evaluating their performance in terms of forecast accuracy, training time, and generalization ability. The findings will provide insights into the suitability of machine learning and deep learning models for renewable energy forecasting, helping researchers and industry professionals select the most efficient and reliable forecasting approach.

The main contributions of this paper include verification of the execution time of decision tree algorithms (XGBoost, LightGBM) and the overfitting limitations of RNN (Recurrent Neural Networks) algorithms (BiGRU). The remainder of this paper is organized as follows: Section 2 provides the research methodology of research with selected algorithms; Section 3 presents the experimental results, analysis and discussion; and finally, Section 4 concludes the paper with a summary of the key points and suggestions for future research.

2. Methodology and Forecasting Models

2.1. XGBoost and LightGBM Model

XGBoost and LightGBM [5, 7] are built on the Gradient Boosting principle, employing an iterative learning approach where decision trees are trained sequentially to minimize residual errors from previous iterations. Specifically, each new tree in the model learns from the residual errors of the preceding model to optimize overall accuracy. This process continues until an optimal number of trees is reached or until the error reduction stabilizes. Finally, predictions from all trees are aggregated using weighted averaging or majority voting to produce the final output. The use of XGBoost and LightGBM improves training speed and generalization ability compared to traditional Gradient Boosting, thanks to improvements such as memory optimization, tree pruning techniques, histogram-based splitting, and more efficient support for large data processing. The workflow of XGBoost and LightGBM based on Gradient Boosting is shown in Fig. 1 below.

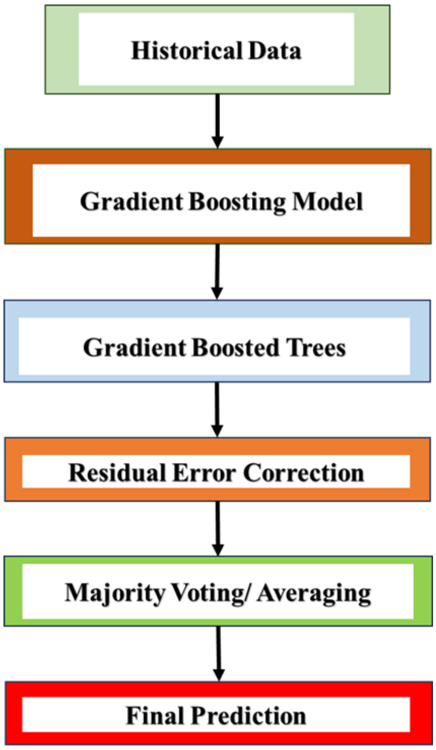

Fig. 1. The working process of XGBoost and LightGBM is based on Gradient Boosting

Both XGBoost and LightGBM offer advantages over traditional Gradient Boosting due to their enhancements as following:

a) XGBoost optimizes training with Gradient Descent Boosted Trees, where each tree is updated based on residual errors [7]:

- It uses shrinkage (learning rate) to prevent overfitting.
- It employs column sampling to reduce correlation between trees.
- It supports pruning to eliminate unimportant branches.

b) LightGBM improves Gradient Boosting by [5] following properties:

- GOSS (Gradient-based One-Side Sampling): Prioritizing samples with higher gradient values for training, increasing efficiency without losing accuracy.
- Histogram-based Splitting: Faster branch splitting compared to scanning the entire dataset.
- Leaf-wise Growth Strategy: Growing trees by optimizing the most beneficial leaf rather than level-wise splitting like XGBoost, improving training speed.

2.2. BiGRU Model

BiGRU [8] is a variant of RNNs designed for time-series data processing by capturing information in both directions: from past to present and vice versa. This model uses gating mechanisms to retain relevant information and discard unnecessary data, enhancing memory retention compared to traditional RNNs. In the solar power forecasting problem, BiGRU can effectively capture time trends and complex relationships between input factors. The choice of a BiGRU model with a single hidden layer is based on a balance between training time and accuracy. Specifically, a single-layer BiGRU typically trains faster than a two-layer BiGRU because it has fewer parameters to optimize [9]. This helps reduce computational resources and processing time, which is especially important when working with large datasets or real-time applications.

At the same time, a single hidden layer still allows the model to capture important features in time-series data, ensuring the necessary accuracy for solar power forecasting. However, it is important to note that an overly complex model can lead to overfitting, while a model that is too simple may not have enough capacity to learn complex data patterns. Therefore, choosing a single-layer BiGRU is considered a reasonable solution, ensuring both efficiency and accuracy in forecasting. The structure of the single-layer BiGRU model is presented in Fig. 2 below [10].

The 1-hidden-layer BiGRU model combines two GRUs in both forward and reverse directions. Each GRU consists of update gates and reset gates, which are defined through specific mathematical formulas as follows [11]:

- The reset gate (r_t) is calculated as follows:

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \tag{1}$$

- The update gate (z_t) is calculated as follows:

$$z_t = \sigma(W_{xt}x_t + W_{hz}h_{t-1} + b_z) \tag{2}$$

- The candidate memory content ($\tilde{h}_t$) is updated as:

$$\tilde{h}_t = \tanh(W_{xh}x_t + W[r_j \odot h_{t-1}]) \tag{3}$$

- The hidden state at time step t (h_t) is updated as:

$$h_t = (1\text{-}z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{4}$$

In these fomulations: σ is sigmoid function, tanh is hyperbolic tangent function, W_{xr}, W_{hr}, W_{hz}, W_{xt}, W_{xh} are weight matrices corresponding to the input, hidden state, and the reset and update gates, W is a temporary weight after element-wise multiplication, b_r, b_z are bias vectors for the reset and update gates, x_t is the input feature vector at time step t, h_{t-1} is the hidden state from the previous time step, h_t is the hidden state at time step t, and ⊙ represents the Hadamard product (element-wise multiplication).

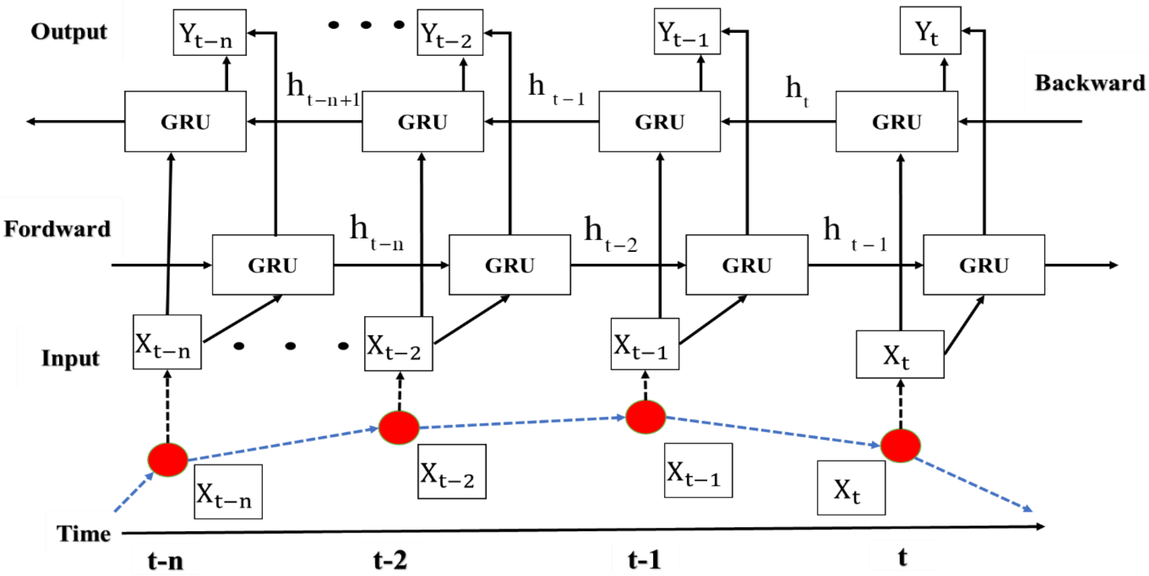

Fig. 2. Structure of BiGRU model (single hidden layer)

The BiGRU architecture used in this study consists of one hidden layer with 64 units in both forward and backward directions, resulting in a total of 128 hidden states. The model uses the 'tanh' activation function, the 'adam' optimizer, and the mean squared error (MSE) as the loss function, dropout rate of 0.1. The model was trained for 30 epochs with a batch size of 32.

2.3. Algorithm Diagram Comparing Three Models XGBoost, LightGBM and BiGRU in Forecasting Solar Power Generation Capacity

The diagram below illustrates the solar power generation forecasting process, encompassing steps from collecting and preprocessing historical weather and power data, to training and evaluating forecasting models such as XGBoost, LightGBM, and BiGRU. Subsequently, forecasted weather data is applied to estimate power generation, and finally, the forecast results are analyzed to assess the performance of these models.

Steps to perform the process:

Step 1: Collect Historical Data

The historical data was sampled every 5-minute. This high resolution frequency was selected to better capture short-term fluctuations in solar power generation, while maintaining acceptable computational efficiency after aggregation and normalization steps.

Historical weather and solar power output data were gathered from January 1, 2022, to December 31, 2022, for subsequent model training and evaluation. This dataset includes factors affecting power generation, such as sunlight intensity, ambient temperature, panel temperature, and the actual electricity produced by the system.

Step 2: Preprocess the Data

Before building models, ensure the data is clean and ready:

- Handle Missing Data: Fill in missing values using methods like interpolation or by using the average value.
- Remove Outliers: Eliminate unusual data points using the Interquartile Range (IQR) method, which helps identify and remove values that are significantly different from others.
- Normalize Data: Scale the data to a standard range using MinMaxScaler, which adjusts all values to be between 0 and 1.
- Split the Data: Divide the dataset into two parts: 70% for training the model and 30% for testing and evaluating its performance.

Step 3: Train Forecasting Models

Choose three models to predict solar power output: LightGBM, XGBoost, and BiGRU. Training and fine-tuning of each model are conducted using the preprocessed data to optimize their key parameters. Each model was assessed by analyzing its training time and prediction performance on the test data, aiming to ensure its ability to generalize to unseen data.

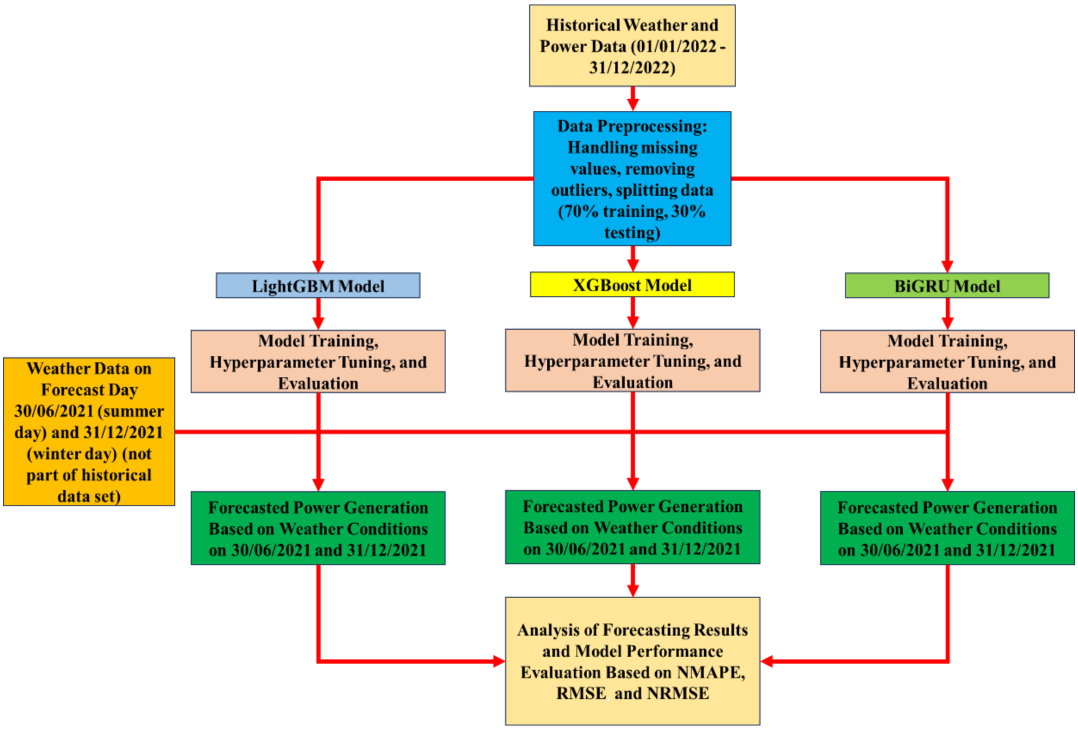

Fig. 3. Solar power generation forecasting process

Step 4: Forecast Using Specific Day's Weather

After training, we evaluate the forecasting models under different seasonal conditions using weather data on June 30, 2021 (summer), and December 31, 2021 (winter). For each of these days, the input consists of 288 weather data points, corresponding to 5-minute intervals throughout the 24 hours. These dates were chosen to represent peak solar generation in summer and reduced solar generation in winter, allowing for a comprehensive analysis of model performance across varying weather conditions.

Step 5: Generate Predictions

Each model (LightGBM, XGBoost, BiGRU) makes its own prediction of power generation based on the June 30, 2021, and December 31, 2021 weather data. The time taken for each prediction and the results were recorded to compare the accuracy.

Step 6: Analyze and Evaluate Results

The predictions from all three models were compared to assess their performance. Use metrics like Normalized Mean Absolute Percentage Error (NMAPE), Root Mean Squared Error (RMSE), and Normalized Root Mean Squared Error (NRMSE). This comparison helps determine which model is the most accurate and suitable for forecasting solar power generation.

2.4. Errors Used to Evaluate Model Performance

To assess the performance of solar power forecasting models, three key metrics are used: NMAPE, RMSE, and NRMSE.

NMAPE is chosen because it represents the average percentage error relative to actual values. It helps evaluate model accuracy without being affected by the scale of the data [12].

$$NMAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_{nominal}} \tag{5}$$

where: $\hat{y}_i$ is the predicted power (kW), y_i is the actual power (kW), n is the number of data points, $y_{nominal}$ is the total installed capacity of the plant (kW).

RMSE measures the average squared difference between actual and predicted values, providing insight into how much the predicted values deviate from reality. This metric is particularly useful for identifying large forecasting errors since RMSE gives greater weight to larger errors compared to MAE (Mean Absolute Error) .

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{6}$$

where: $\hat{y}_i$ is the predicted power *(kW)*, y_i is the actual power *(kW)*, n is the number of data points.

NRMSE is the normalized version of RMSE, which adjusts the squared error difference between actual and predicted values to a standard range [12].

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}}{y_{nominal}} \times 100 \tag{7}$$

where: $\hat{y}_i$ is the predicted power (kW), y_i is the actual power (kW), n is the number of data points, $y_{nominal}$ is the total installed capacity of the plant (kW)

Combining these metrics ensures a comprehensive evaluation of model performance: NMAPE reflects overall accuracy, while RMSE, NRMSE helps detect large errors, allowing for the identification of the most suitable model for real-world applications.

2.5. Parameters of the Forecasting Models

Parameters of the XGBoost Model

Table 1 presents the key parameters used in training XGBoost model. These parameters were selected based on empirical studies to balance model complexity, prevent overfitting, and optimize predictive performance.

Table 1. Parameters of the XGBoost Model

learning_rate	0.1
max_depth	8
n_estimators	150
Min_child_weight	1
Gamma	0
subsample	1
colsample_bytree	1

Parameters of the LightGBM Model

Table 2 presents the key parameters used in training LightGBM model. The parameter choices aim to improve computational efficiency while maintaining high accuracy and strong generalization capability.

Table 2. Parameters of the LightGBM Model

Learning_rate	0.1
Max_depth	8
Num_leaves	31
Min_child_samples	20
reg_alpha	0
reg_lambda	0
Colsample_bytree	1

Parameters of the BiGRU Model

Table 3 presents the key parameters used in training BiGRU model. The selected parameters are configured to capture temporal dependencies in the data while balancing model complexity during training.

Table 3: Parameters of the BiGRU Model

Optimization algorithm	adam
Activation function	tanh
Loss function	MSE
Epochs	30
Batch sizes	64

3. Results and Discussion

To assess the accuracy and performance of solar power generation forecasting models, Table 4 presents the forecasting results of XGBoost, LightGBM, and BiGRU on the test set, including error metrics and execution time.

The results in Table 4 highlight significant differences in execution time and forecast accuracy among XGBoost, LightGBM, and BiGRU for solar power generation forecasting. Among them, LightGBM performs the best, with the fastest training time (0.705 seconds) and the lowest error rates (NMAPE equal to 1.004%, RMSE equal to 1373.27 kW, NRMSE equal to 2.77%), proving that it is both accurate and stable. XGBoost also gives good results, with NMAPE equal to 1.012% and RMSE equal to 1397.28 kW, which are slightly higher than LightGBM, but it takes more time to train (2.15 seconds). On the other hand, BiGRU has the worst performance, taking a very long time to train (819.32 seconds, over 380 times longer than XGBoost and more than 1160 times longer than LightGBM). It also has much higher errors (NMAPE equal to 1.685%, RMSE equal to 2253.68 kW, NRMSE equal to 4.55%), showing that it does not generalize well and may be overfitting.

To analyze the performance of the forecasting models under different seasonal conditions, we evaluate their predictions on two distinct weather datasets: June 30, 2021 (summer), and December 31, 2021 (winter). Table 5 presents the forecast results of the XGBoost, LightGBM, and BiGRU models on these datasets, allowing for a comparative assessment of their accuracy and robustness across varying solar radiation levels.

The forecasting results of XGBoost, LightGBM, and BiGRU for solar power generation on June 30, 2021, as shown in Table 5, reveal clear differences in execution time and accuracy. Among the three models, XGBoost has the highest accuracy, with NMAPE equal to 1.964%, RMSE equal to 1804.775 kW, and NRMSE equal to 3.646%, showing stable predictions and the lowest error rates. LightGBM has the fastest processing speed (0.705 seconds), achieving NMAPE equal to 2.051%, RMSE equal to 1714.571 kW, and NRMSE equal to 3.463%. Its accuracy is slightly lower than that of XGBoost, but it still provides highly reliable predictions. On the other hand, BiGRU performs the worst, with a very high NMAPE (23.986%), RMSE equal to 18713.62 kW, NRMSE equal to 37.805%, and an execution time of 33.406 seconds, which is much longer than the two Gradient Boosting models (XGBoost and LightGBM). These results indicate that BiGRU fails to provide accurate forecasts and is likely affected by overfitting.

Table 4. Forecast results on the test set of three models XGBoost, LightGBM and BiGRU

Models	Training time (s)	NMAPE (%)	RMSE (kW)	NRMSE (%)
XGBoost	2.15	1.012	1397.28	2.82
LightGBM	0.705	1.004	1373.27	2.77
BiGRU	819.32	1.685	2253.68	4.55

Table 5. Forecast results on the test set of three models XGBoost, LightGBM and BiGRU on the weather dataset of June 30, 2021 and December 31, 2021

Day	Models	Forecast execution time (s)	NMAPE (%)	RMSE (kW)	NRMSE (%)
June 30, 2021	**XGBoost**	1.212	1.964	1804.775	3.646
	LightGBM	0.705	2.051	1714.571	3.463
	BiGRU	33.406	23.986	18713.62	37.805
December 31, 2021	**XGBoost**	1.112	0.64	506.301	1.022
	LightGBM	0.716	0.85	659.16	1.331
	BiGRU	32.641	1.023	881.189	1.78

Also from Table 5, the forecast for December 31, 2021 with the XGBoost model achieved the best performance on December 31, 2021, with the lowest error (NMAPE equal to 0.64%, RMSE equal to 506.301 kW) and the fastest execution time (1.112 seconds). LightGBM also performed well, although its error was slightly higher (NMAPE equal to 0.85%, RMSE equal to 659.16 kW). BiGRU had a similar level of error (NMAPE equal to 1.023%, RMSE equawwwwwwwwwww2l to 881.189 kW), but its execution time (32.641 seconds) was significantly longer.

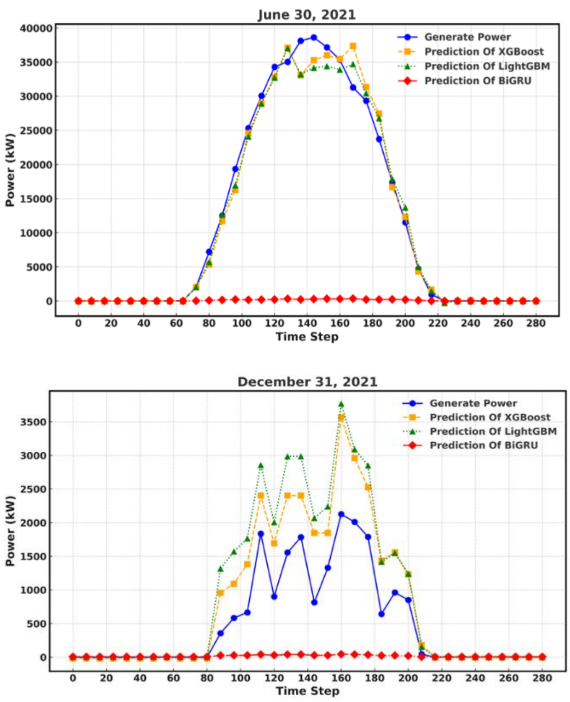

Fig. 4. Forecasted with Actual Power Generation (●) for XGBoost (■), LightGBM (▲), and BiGRU (◆) on June 30, 2021 and December 31, 2021

From Fig. 4, we can see that the forecast curves of LightGBM (▲-upward-pointing triangles) and XGBoost (■-squares) closely follow the actual power curve (●-circles), especially during the rising and falling power phases, showing that both models have good forecasting ability. XGBoost has a smoother curve compared to LightGBM, while LightGBM reacts more sensitively to small variations. This suggests that XGBoost may generalize better in some cases. The forecasted curve of BiGRU (◆-diamonds) is almost flat at zero, failing to reflect the actual power output, which indicates that the model does not provide accurate predictions. This could be a sign of overfitting issues. LightGBM shows large fluctuations at peak power levels, possibly because the model overreacts to the training data. XGBoost is smoother but still has some small errors in high-power regions. BiGRU fails to capture trends, making it less practical for solar power forecasting.

Fig. 4 also compares the power generation forecasts of XGBoost, LightGBM, and BiGRU on December 31, 2021. XGBoost and LightGBM closely follow the actual trend, showing their ability to capture fluctuations in power generation. However, LightGBM exhibits slightly higher variability. In contrast, BiGRU fails to follow the trend, producing almost flat predictions, which suggests poor adaptability to real changes in power generation and the risk of overfitting. Although BiGRU shows signs of overfitting, this is not reflected in error metrics like RMSE or NMAPE. This is because these metrics mainly measure the average difference between actual and predicted values but do not assess how well the model captures trends. BiGRU may have memorized patterns from the training data too much, leading to forecasts that lack variation and appear flat. This makes the average error seem low, even though the model does not correctly reflect real power fluctuations. Additionally, when actual power generation is low, the absolute error can be small, reducing the impact of overfitting on metrics like NMAPE. However, by observing the chart, it is clear that BiGRU does not respond quickly enough to changes, resulting in inaccurate forecasts during periods of rapid power fluctuations.

4. Conclusion

The comparison of XGBoost, LightGBM, and BiGRU in solar power generation forecasting highlights significant differences in execution time, accuracy, and generalization ability. XGBoost and LightGBM both achieve low errors in NMAPE, RMSE, and NRMSE, proving their accuracy and stability in forecasting. LightGBM consistently outperforms in speed, completing predictions faster than other models, while XGBoost provides slightly higher accuracy but requires more computation time.

In contrast, BiGRU performs the worst, with significantly higher errors and an extremely long execution time, requiring far more computing resources without delivering the expected accuracy. The model also struggles with generalization, as seen from its high NMAPE (23.986%) and RMSE (18763.12 kW) when tested on independent data on June 30, 2021, reinforcing concerns about overfitting. When forecasting on December 31, 2021, although BiGRU shows signs of overfitting, this is not fully reflected in RMSE and NMAPE, as these metrics focus on average error rather than trend accuracy. BiGRU's predictions are overly smooth, failing to capture real power fluctuations. This makes the error appear low despite poor adaptability to rapid changes. When power generation is low, absolute errors are also small, further masking overfitting. However, the comparison chart reveals BiGRU's limitations in tracking variations in power output.

Additionally, XGBoost and LightGBM significantly outperform BiGRU in both training and forecasting efficiency. LightGBM achieves the lowest error NMAPE, and RMSE is followed closely by XGBoost. Meanwhile, BiGRU requires over 380 times longer than XGBoost and 1160 times longer than LightGBM to train, yet fails to improve accuracy. This reaffirms the comparative advantage of XGBoost and LightGBM in both accuracy and robustness.

In summary, XGBoost and LightGBM are both strong candidates for solar power forecasting, offering fast processing and high accuracy. Among them, LightGBM stands out as the best choice due to its superior speed and slightly better accuracy. Conversely, BiGRU demonstrates limitations in this case, with its long execution time, higher forecasting errors, and signs of overfitting. However, its performance might improve with further optimization, regularization techniques, or different datasets, suggesting that its applicability depends on specific conditions.

Reference

[1] I. Jebli, F. Z. Belouadha, M. I. Kabbaj, and A. Tilioua, Deep learning based models for solar energy prediction, Advances in Science, Technology and Engineering Systems Journal, vol. 6, iss. 1, pp. 349-355, 2021. https://doi.org/10.25046/aj060140

[2] X. Guo, Y. Zhan, D. Zheng, L. Li, and Q. Qi, Research on short-term forecasting method of photovoltaic power generation based on clustering SO-GRU method, Energy Reports, vol. 9, sup. 10, pp. 786-793, Oct. 2023. https://doi.org/10.1016/j.egyr.2023.05.208

[3] Y. Peng, S. Wang, W. Chen, J. Ma, C. Wang, and J. Chen, LightGBM-integrated PV power prediction based on multi-resolution similarity, Processes, vol. 11, iss. 4, Apr. 2023. https://doi.org/10.3390/pr11041141

[4] J. Gaboitaolelwe, A. M. Zungeru, A. Yahya, C. K. Lebekwe, D. N. Vinod, and A. O. Salau, Machine learning based solar photovoltaic power forecasting: a review and comparison, IEEE Access, vol. 11, pp. 40820-40845, Apr. 2023. https://doi.org/10.1109/ACCESS.2023.3270041

[5] G. Ke and I.-Y. Meng, QiFinely, ThomasWang, TaifengChen, WeiMa, WeidongYe, QiweiTLiu, LightGBM: a highly efficient gradient boosting decision tree, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, Dec. 2017.

[6] G. Xiong, J. Zhang, X. Fu, J. Chen, and A. W. Mohamed, Seasonal short-term photovoltaic power prediction based on GSK–BiGRU–XGboost considering correlation of meteorological factors, Journal of Big Data, vol. 11, art no. 164, Nov. 2024. https://doi.org/10.1186/s40537-024-01037-x

[7] 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, vol. 42, no. 8, pp. 785-794, Aug. 2016. https://doi.org/10.1145/2939672.2939785

[8] L. Kumar and S. Sarkar, Neural search: learning query and product representations in fashion e-commerce, vol. 1, no. 1, Association for Computing Machinery, 2021.

[9] T. A. Nguyen, M. H. Pham, M. P. Vu, D. T. Nguyen, and A. T. Vu, Solar power generation prediction using BiGRU model, Journal of Science and Technology in Energy, Electric Power University, vol. 34, pp. 1-14, 2024.

[10] S. Goswami, S. Malakar, B. Ganguli, and A. Chakrabarti, A Novel transfer learning-based short-term solar forecasting approach for India, Neural Computing and Applications, vol. 34, pp. 16829-16843, Jun. 2022. https://doi.org/10.1007/s00521-022-07328-9

[11] M. Sabri and M. El Hassouni, A novel deep learning approach for short term photovoltaic power forecasting based on GRU-CNN model, E3S Web of Conference, vol. 336, art no. 00064, Jan. 2022. https://doi.org/10.1051/e3sconf/202233600064

[12] Q. T. Phan, Y. K. Wu, and Q. D. Phan, Short-term solar power forecasting using XGBoost with numerical weather prediction, 2021 IEEE International Future Energy Electronnics Conference IFEEC, Taipei, Taiwan, 16-19 Nov. 2021, pp. 1-2. https://doi.org/10.1109/IFEEC53238.2021.9661874