Application of Artificial Intelligence in Solar Power Generation Forecasting

Keyu Wang

{ wangkeyu2333@outlook.com }

Microelectronics, Wuhan Textile University, Wuhan 430200, China

Abstract. With the world going through the energy paradigm shift to clean and low-carbon energy sources, the development of solar power, as one of the cornerstone renewable energy sources, shows grid stability and energy optimization challenges that involve the intermittency and unpredictability of solar power output. The specifics of solar power generation prediction have turned out to be the key to the solution of this problem. This paper will discuss the use of artificial intelligence in solar power generation prediction, covering the core technologies and the viable solutions in the given field. The studies involve deep learning applications, machine learning, and hybrid model applications. The problems of the current research include the lack of data quality and standardization, the challenge of balancing model complexity and real-time performance, and the lack of interpretability. It is desirable to advance in the future in the direction of multimodal data fusion and such directions. This paper represents a systematic review of the implementation results of artificial intelligence technology in solar energy forecasting, and it has tremendous implications in the optimization of energy distribution and the stability of the grid.

Keywords: Solar Power Generation Forecast, Artificial Intelligence, Deep Learning, Machine Learning

1 Introduction

With an overall global setting of global production of duo carbon environment which has been propelling the energy paradigm shift to low-carbon and clean energy production, energy generation of solar photovoltaic type has become a fundamental part of the renewable energy industry. The advantages associated with this change include zero carbon emissions and its ubiquitous presence of resources. Nevertheless, solar power generation is dynamically controlled by the variables of solar power, ambient temperature, and the variations in cloud cover, which have high levels of intermittency and variability. This attribute is a problem in the distribution of large-scale solar energy to the grid, such as the possibility of grid frequency instability, unnecessary redundancy of the reserve, and nightingaling solar power. These challenges are gross problems to the stable functioning of power systems and optimization of the energy dispatch, thus, the development of the high-precision solar power generation

ICMEEA 2026, April 24-26, Singapore, Singapore

DOI 10.4108/eai.24-4-2026.2364848

forecasting has become an outstanding technological solution of the solar intermittency issue and increasing the rates of renewable energy usage. The question of how to beat around the constraints of the traditional forecasting methods using technological innovation has become a topic of significant research interest in the energy industry.

Over the past few years, the artificial intelligence technology has steadily substituted the traditional physical and statistical modeling in solar power generation forecasting because of its singular strengths of nonlinear, high-dimensional time-series data, becoming a research hotspot in this area.

Since the effective use of the traditional machine learning algorithms in specific cases that need limited data [1], to achievements in deep learning technologies, including the CNN-LSTM hybrid architecture based on infrared image processing and analysis [2], the LSTM time series analysis model of hybrid photovoltaic equipment and improvements in prediction accuracy with hybrid models and multi-source data fusion strategies [3,4], artificial intelligence has brought a complete system of predictions technology, covering ultra-short-term, short time, and medium-long-term times. At the same time, the data preprocessing methodology also undergoes optimizations [2] accompanied by the transfer learning introduction, which offers successful ways to handle data quality flaws and to allow cross-regional forecasting. Prediction accuracy and robust prediction has been further enhanced by innovative models like integrated LSTM [5] and bidirectional LSTM that have optimized attention mechanisms [6]. Furthermore, the benefits of Transformer models in managing long-term time series data [7], the fact that better LSTM-based photovoltaic power forecasting models are able to collect nonlinear features [8], the two-stage forecasting system utilizing multi-model comparison and scenario adaptation are all add benefits to the technical application framework of artificial intelligence involved in this domain.

In this paper, the systematic review of the state of art of artificial intelligence in solar power generation forecasting will be performed. The study encompasses the technical principles and performance of deep learning, CNN-LSTM hybrid models, Transformers, the traditional machine learning and hybrid models. It examines how preprocessing of the data and multi-source data fusion can be used in maximising the accuracy of the prediction, and also discusses whether the transfer learning and model optimization strategies are useful in multi-scenario prediction. In this paper, the authors will attempt to combine research available in an effort to develop an applications framework AI technology in solar power generation forecasting. It explains the benefits and situations of use of various technical methods and provides a clear direction on future research and practice of engineering. This structure will contribute to stable operation of the grids and optimization of the energy system in the infiltration of photovoltaic grids under high penetration, which appears to be technologically favorable in the realization of carbon neutrality as well as carbon capping of the world energy system.

2 Relevant Theoretical Framework

2.1 Long Short-Term Memory Networks and Their Variants

The LSTM technique successfully creates a solution to the vanishing gradient problem in conventional recurrent neural networks by design, with the introduction of the forgetting gates, ingredients of input gates, and output gates. It is capable of dependency relationships in long-term sequence data, as well as it is effective in the vanishing gradient problem in traditional recurrent neural networks. It is able to deal with dependencies within long-term time series data, and is used extensively in solar irradiance and power projections. Febriansyah et al. used an LSTM network to make future-day irradiance and temperature forecasts of a hybrid photovoltaic system in Temajuk Village in Indonesia. All RMSE, MAE, and MAPE values were below 5% and this corresponded to the standard of prediction of a high level of accuracy. Ban et al. trained an LSTM model with data from a photovoltaic power station in Changzhou with an R^2 of 0.9847 and an MSE of 6.6374×10^3 on the test set. This greatly surpassed the traditional linear models, proving the fact that this model was able to reflect the nonlinear attributes of photovoltaic power output.

The bidirectional LSTMs are used to improve the ability to learn complex patterns because the sequences of data are processed in both forward and backward directions, and use past and future information at the same time. Proposed attention-based bidirectional LSTM model optimized by using the manta ray algorithmic optimization function has standard prediction work of MAE=0.15, RMSE=0.21, and R^2=0.97on the Rajasthan meteorological data in India. It is more than 40 percent lower than the error of traditional LSTMs [6].

To make it ever more accurate, researchers suggested an integrated strategy and method of hyperparameter optimization. Wang et al. NGO-LSTM model applies the Northern Goshawk Optimization algorithm to the hyperparameters of LSTM, including the number of neurons in the hidden layers and the learning rate. It was also implemented on the Huanggang photovoltaic power station data, where it was observed that it reduced MAE and MSE by 49.31% and 80.31% respectively compared to the operation of the traditional LSTM and with this it appraised the effect of randomness of manual parameter optimization.

2.2 Convolutional Neural Networks (CNN) and Hybrid CNN-LSTM Models

CNN is good at retrieving spatial features, whereas LSTM is good at handling the time dependencies. The combined hybrid model of both approaches shows excellent outcomes in image-based forecasts.

Rahman et al. overcame the constraints of infrared sky images that had low resolutions by introducing a CNN-LSTM hybrid network: Bicubic interpolation was used to improve the resolution of IR images. After that, they executed single-channel grayscale images with the JET color map of OpenCV to transform them into three-channel RGB images to enhance the ability to visualize cloud cover and variations in temperature. Subsequently, they employed a CNN regression model for “instantaneous prediction” of irradiance, followed by short-term temporal forecasting via an LSTM model. Experimental results demonstrate that this framework reduces the RMSE from 47.24 W/m^2 to 33.51 W/m^2, proving the critical role of image enhancement and CNN-LSTM fusion in improving prediction accuracy [2].

2.3 Transformer Model

Transformers are self-attention-based models that process parallel time-series data, effectively modeling long-range interactions. They have increasingly been used in the last few years in solar power prediction. Talukdar et al. compared LSTMs and Transformers and discovered that the higher computational cost of Transformers is measured when using long-term sequence data, but Transformers outperform LSTMs by an increase of 5 to 8 percent in the performance metric, R 2. Transformers are especially beneficial in high complexity temporal variations in extreme weather patterns and other instances, such as heavy cloud cover or torrential rain, which minimizes prediction errors by over 15%.

2.4 Traditional Machine Learning Techniques

Traditional machine learning models, though structurally simple, remain practical in scenarios with limited data or high real-time requirements. These primarily include gradient-boosted trees, random forests, and support vector machines.

XGBoost is based on gradient boosting, combining a variety of decision trees, and is resistant to overfitting and is also quick to train. An XGBoost model was built in the paper of Talukdar et al. with the Kaggle photovoltaic dataset. The analysis of the importance of features showed that clear-sky global horizontal irradiance carries the most weight on the power output. The final RMSE of the model was 0.3733; thus, the actual values were more than 90 percent of the predicted values. It has the capability of real-time deployment in smart grids [1].

Random Forest reduces variance through Bootstrap sampling and random feature selection, making it suitable for scenarios with multiple features and low volatility. Thongwily et al. observed in their summer forecasting of Thai photovoltaic systems that RF demonstrated superior learning efficiency for stable data, achieving the lowest RMSE, MAE, and MAPE among all models. With a training time of only 12 seconds—significantly shorter than LSTM and Bi-LSTM—it is well-suited for scenarios with stable summer irradiance [9,10].

2.5 Hybrid Models and Multi-Technology Integration

The hybrid models address the weaknesses of the individual models by incorporating the advantages of the various AI technologies. These are commonly referred to as physical models + AI models and typical machine learning + deep learning.

The model created by Aneesh et al. combines both physical equations to compute solar irradiance with the results of LSTM prediction by having physical models to narrow the reality of the AI model output. This method increases the accuracy of prediction by 20% in low-density data areas [4].

3 Technical Scenario Exploration

3.1 Time-Scale-Based Prediction Scenarios

The primary applications of ultra-short-term forecasting are real-time management of photovoltaic power plants, grid frequency regulation, and real-time response to sudden variations that require high real-time performance and dynamic tracking. In this case, the CNN-LSTM hybrid model, which is based on image data, shows excellent results. It can capture the change in irradiance due to the rapid change of clouds by processing the infrared sky image in real time and giving the prediction of the irradiance instantaneously and this can be used by real-time dispatching of a photovoltaic power plant. At the same time, the traditional machine learning models are lightweight, i.e. they can be trained and predict rapidly, which can be used in edge computing applications with extreme time response needs, i.e. commanding milliseconds to respond to power grid events. This type of situation is typical in intraday control of large-scale PV plants in power smoothing, real-time management of energy storage system charging and discharging, and real-time balancing of PV output and load in microgrids.

Short-term forecasting is the most widely applied scenario in solar power generation prediction, primarily used for day-ahead power system scheduling, unit combination optimization, and daytime balancing of PV output and load. It requires maintaining accuracy while ensuring model versatility. LSTM and its variants represent the mainstream technical choice for this scenario. By utilizing historical time-series data such as irradiance, temperature, and humidity, they can accurately predict PV output for the following day or specific time slots within a day. Adaptation to seasonal climate characteristics is also achievable. Random forest models emerge as the preferred option due to their high efficiency and superior accuracy. During the rainy season, where the weather is highly variable and output fluctuates significantly, gated recurrent units (GRU) demonstrate superior predictive performance owing to their strong adaptability to dynamic changes. Winter days feature shorter daylight hours and complex influencing factors. Bi-LSTM effectively enhances prediction accuracy by capturing bidirectional temporal correlations. Furthermore, two-stage forecasting methods—first predicting key influencing factors, then using these predictions for power forecasting—further improve the reliability of short-term forecasts, providing critical support for grid daytime scheduling planning.

Medium-to-long-term forecasting is primarily used for photovoltaic power plant planning and design, grid investment decisions, energy policy formulation, and seasonal power supply-demand balancing. It requires integrating long-term meteorological trends with historical data, emphasizing forecast stability and trend control. In this scenario, hybrid models demonstrate significant advantages. By integrating satellite annual irradiance data, regional climate characteristics, and photovoltaic module physical parameters, they enable monthly, quarterly, or even annual PV output forecasting. Simultaneously, transfer learning technology addresses the scarcity of historical data in underdeveloped regions. By leveraging model knowledge from data-rich areas and fine-tuning models with local limited data, it enables cross-regional medium-to-long-term forecasting. The results from this scenario provide crucial support for photovoltaic power plant site selection, capacity determination, and grid transmission line planning, thereby contributing to the long-term optimization of energy system layouts.

3.2 Application-Based Prediction Scenarios

Large-scale ground-mounted photovoltaic power plants feature substantial installed capacity and extensive output influence, demanding exceptionally high prediction accuracy and stability. This requires comprehensive consideration of multiple factors, including cloud conditions, temperature, and wind speed. The CNN-LSTM hybrid model processes infrared sky imagery to precisely capture irradiance variations caused by cloud movement, providing spatial feature support for power plant output forecasting. The LSTM model, optimized using meta-heuristic algorithms, enhances robustness through hyperparameter tuning. This enables adaptation to output fluctuations across different seasons and weather conditions, mitigating the impact of extreme weather on prediction accuracy. Furthermore, multi-source data fusion enhances prediction comprehensiveness, providing technical assurance for large-scale PV power plants to participate in grid dispatch and avoid curtailment.

Distributed photovoltaic systems are characterized by their small scale and dispersed deployment, often closely integrated with user loads. Forecasting requires balancing flexibility with real-time responsiveness while being significantly influenced by local environmental conditions. Traditional machine learning models, characterized by rapid training and low data requirements, are well-suited for swift deployment in distributed scenarios. For floating PV systems, the Bi-LSTM model can enhance prediction accuracy by integrating surface meteorological features with time-series data. Concurrently, the combination of edge computing technology and lightweight models enables local forecasting and real-time control for distributed PV systems. This approach reduces data transmission delays, optimizes the matching between PV output and user load, and improves the utilization rate of distributed energy resources.

Hybrid photovoltaic systems require coordinated output from multiple energy sources. Forecasting must account for the complementarity and interdependencies among different energy sources, demanding high multivariate processing capabilities from the model. The LSTM time-series analysis model can simultaneously process multidimensional time-series information such as PV output, energy storage status, and meteorological data, providing predictive support for system coordinated scheduling. The integrated model combines the strengths of different models to enhance prediction robustness in multi-energy scenarios, helping hybrid systems optimize energy allocation, reduce reliance on traditional energy sources, and improve system operational efficiency and economic performance.

4 Limitations and Shortcomings

4.1 Data-Level Constraints

The available literature mostly utilizes datasets of photovoltaic power plants of particular regions. Collection concerns: Data gathering is prone to sensor malfunctions, data gaps, and severe weather-related disturbances, as well as interruptions in the transmission, resulting in many outliers. In other instances, others have long data lapses as a result of lightning or machine faults. Though the gaps may be filled by interpolation, interpolation entails an error. Abnormal data points (non-zero output at night or irradiance that is beyond the equipment capability) have the negative quality to impact model training results in an unfavorable way when not addressed correctly. Moreover, there are some drastic variations in the data formats and observation frequency of various regions and power plants, and the lack of standardized collection and storage criteria. This leads to a lack of compatibility of data between regions and sceneries, which impedes the model transfer and validation between regions and sceneries.

The available datasets are typically concentrated in selected regions of climatic areas, and the magnitude of special climatic areas like high latitudes, plateaus, and deserts is unrepresented. This creates poor model generalization processes in these fields. Moreover, the data of photovoltaic output during extreme weather conditions is limited, and as such, models find it hard to learn the special patterns during such periods. As a result of this, the errors made in prediction rise considerably in severe weather. At the same time, the multi-source data is not properly integrated. Other research does not maximize multiple types of data because the research depends on one kind only and does not utilize multidimensional data to contribute to the optimal prediction performance.

4.2 Model-Level Constraints

Deep learning models have high accuracy in prediction and are characterized by complicated architecture and large numbers of parameters, which make them slow to train and expensive to run. All these models have a black box nature with the prediction processes and the logic of decision-making that can hardly be intuitively visualized. Models that are in use are oftentimes trained and modeled to work under a set of circumstances, which causes great losses of accuracy when attempting to predict other situations. Moreover, the models are data-intensive to changes in data distribution. Where there is a significant disparity between the distributions of operational data and training data, it is likely that the model will experience a distribution drift, which will result in poor prediction.

4.3 Application-Level Restrictions

The current studies are mostly restricted to the laboratory phase, and model performance is established on the condition of using concrete datasets and does not take into consideration all the factors involved in real-life engineering conditions.

5 Recommendations for Future Research

5.1 Data-Level Optimization Recommendations

Develop uniform requirements on the similarity of the collection, storage, and preprocessing of forecast data on solar power generation by defining data structures, observation frequency, definition of data metrics to increase interviewee data compatibility and utility. At the same time, create national and regional photovoltaic data sharing systems to combine historical and real-time data of various power plants and regions. This will allow the sharing of open data, which will be of much assistance in cross-regional and cross-scenario modelling studies. The platforms must have data quality review systems, which include automatic functionality together with manual validation to verify data integrity and accuracy with reduced effects of anomalies and missing values.

Enhance photovoltaic data collection within areas of special climatic change like plateaus, high latitudes, and deserts. To increase the level of generalization of the model, it is necessary to create various datasets that represent the conditions of various types of climates and photovoltaic systems. In an attempt to solve the problem of the lack of extreme weather data, physical modeling and data augmentation techniques simulate extreme weather conditions to enable real-world data gaps to be covered. At the same time, increase studies on the multi-source data fusion technologies in order to integrate the multi-dimensional data in its entirety, i.e. the satellite data, GIS data, IoT sensor data and the PV module operational data to develop a holistic system of data features. This will enhance the flexibility of the model to handle complicated situations.

5.2 Innovation Recommendations at the Model Level

In order to solve the trade-off between model complexity and real-time performance, studies regarding approaches to compressing and accelerating models are in progress. The technique to cut down computational complexity and the number of parameters in deep learning models and preserve prediction accuracy includes parameter pruning, quantization, and knowledge distillation techniques. To illustrate, more intricate Bi-LSTM algorithms can be transformed into much smaller, light-weight ones that can be deployed to edge computing platforms to fulfill real-time predictions on the distributed photovoltaic systems. At the same time, we also experiment with the efficient model architecture. The combination of time and space properties of the solar power generation data allows us to come up with specialized models of a simplified structure and high computation rate, with a balance between accuracy and real-time consideration.

Conduct research on the application of explainable artificial intelligence in solar power generation forecasting. Utilize methods such as feature importance analysis, attention visualization, and model deconstruction to reveal the decision-making logic and predictive mechanisms of the model. Investigate model adaptation and transfer learning techniques to enhance the model's adaptability across diverse scenarios. Employ domain-adaptive networks and meta-learning approaches to enable the model to rapidly adapt to new regions, climates, and photovoltaic system data sources, thereby reducing reliance on extensive local datasets.

5.3 Application-Level Development Recommendations

Strengthen research on integrating artificial intelligence forecasting technology with photovoltaic power plant monitoring systems and grid dispatch platforms. Develop standardized interfaces and data exchange protocols to achieve seamless integration of models with existing systems. Design customized solutions for different application scenarios (large-scale power plants, distributed systems, hybrid systems). Conduct integrated research on combining AI forecasting technology with PV system scheduling, energy storage system control, and grid-cooperative optimization. Build an integrated "forecasting-scheduling-control" solution.

6 Conclusion

Artificial intelligence technology has become the core method for enhancing the accuracy of solar power generation forecasting. From the efficient practicality of traditional machine learning to the complex pattern capture of deep learning, and further to the innovative breakthroughs in hybrid models and multi-source data fusion, it has formed a technical system covering all time scales and diverse application scenarios. LSTM and its variants excel in handling temporal dependencies, while CNN-LSTM hybrid models demonstrate significant advantages in image-data-driven forecasting. Traditional machine learning models remain irreplaceable in scenarios demanding real-time performance and sparse data. Data preprocessing and multi-source data fusion further enhance model robustness and generalization, with transfer learning providing an effective pathway for cross-regional applications. However, current research still faces limitations such as insufficient data quality and standardization, challenges in balancing model complexity and real-time performance, lack of interpretability, and difficulties in engineering implementation. These issues constrain the large-scale deployment of AI technologies in solar power generation forecasting. Future efforts should focus on establishing unified data standards and shared platforms, developing lightweight and interpretable models, and advancing technology engineering and cost reduction to overcome existing bottlenecks, thereby further enhancing the performance and practicality of AI forecasting technologies.

As the global energy transition deepens, high-precision solar power generation forecasting will become a critical enabler for achieving high penetration of photovoltaic grid integration. Continuous innovation and application of artificial intelligence technologies in this field will effectively address the challenge of solar intermittency, optimize energy allocation, and ensure stable grid operation.

References

[1] Talukdar, A., Panda, B., Hota, S., Maharana, M.K., Nayeemuddin, M., Panda, S.: Intelligent Solar Power Forecast with Machine Learning. In: 2024 6th International Symposium on Advanced Electrical and Communication Technologies (ISAECT) (2024)

[2] Nijhum, I.R., Hossain, M.Y., Akinola, T., Rakib, M.M.H., Dev, S.: A Hybrid CNN-LSTM Framework and Infrared Image Processing for Solar Irradiance Forecasting. In: 2024 17th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) (2024)

[3] Febriansyah, Y.I., Barus, D.H., Brunner, I.M.I.M.: Enhancing of Solar Power Forecasting in Hybrid PV Systems: LSTM-Based Time Series Analysis for Solar Irradiance and Temperature Prediction. In: 2024 International Conference on Technology and Policy in Energy and Electric Power (ICTPEP) (2024)

[4] A, M., U, N., Shendige, S.N., A, A.: Enhancing Solar Power Forecasting Through Hybrid Machine Learning Models, Satellite Data Integration, and Transfer Learning Techniques. In: 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (2025)

[5] Daebes, S., Darwish, M., Sing, C.: Machine Learning of Solar Energy Forecasting using Ensemble LSTM Method. In: 2024 59th International Universities Power Engineering Conference (UPEC) (2024)

[6] Prasad, R., Jarin, T., Jose, J.K., Arthi, M., Mewada, H., Ramu, T.B.: Solar Energy Forecasting using Optimized Attention-based Bidirectional Long Short-Term Memory. In: 2024 6th International Symposium on Advanced Electrical and Communication Technologies (ISAECT) (2024)

[7] Talukdar, A., Hota, S., Panda, B., Maharana, M.K., Ravikumar, N.T., Sundar, L.S., Arthi, M., Dinakarraj, C.K.: Performance Analysis of Different Machine Learning Techniques in Solar Energy Forecasting. In: 2025 11th International Conference on Communication and Signal Processing (ICCSP) (2025)

[8] Ban, Y., Sun, Z.: Photovoltaic Power Prediction Based on Improved Long Short-Term Memory Network. In: 2024 6th International Academic Exchange Conference on Science and Technology Innovation (IAECST) (2024)

[9] Wang, Y., Zhang, D., Li, M., Zhang, X.: Study on Power Prediction of Photovoltaic Power Generation Based on NGO-LSTM. In: 2024 8th International Conference on Electrical, Mechanical and Computer Engineering (ICEMCE) (2024)

[10] Thongwily, W., Junhuathon, N., Chayakulkheeree, K.: Two-Stage Data-Driven Modeling for Day-Ahead Solar Power Forecasting: A Case Study of the SUT Campus Floating Solar System. In: 2025 13th International Electrical Engineering Congress (iEECON) (2025).