INDOCHINA APPLIED SCIENCES

M. Lei et al. / Indochin. Appl. Sci. 14(3) (2025) 262329

Research Article

Solar Power Generation Prediction Using LSTM Deep Learning Algorithm in Ningxia Province, China

Mingze Lei[a], Tao Chen[b], Caixia Yang[b], Yao Xiao[b], Jianhui Luo[b], Buncha Wattana[a,*]

[a] *Faculty of Engineering, Mahasarakham University, Maha Sarakham, 44150, Thailand.*
[b] *College of Electrical Engineering, Hunan Mechanical & Electrical Polytechnic, Changsha, 410073, Hunan, China.*

***Corresponding authors** : *buncha.w@msu.ac.th*

https://doi.org/10.55674/ias.v14i3.262329

Received: 26 May 2025 ; **Revised:** 22 June 2025 ; **Accepted:** 07 July 2025 ; **Available online:** 01 August 2025

Abstract

The rapid expansion of photovoltaic (PV) power generation faces significant challenges due to the intermittent and stochastic characteristics of solar energy, which affect grid stability and energy management. Accurate forecasting of PV power output is crucial for optimizing grid operations and supporting the transition to clean energy. This paper proposes a deep learning approach based on Long Short-Term Memory (LSTM) networks to predict PV power generation in Ningxia Province, China. The model leverages historical power and meteorological data, which undergo comprehensive preprocessing, including outlier removal, normalization, and feature correlation analysis. The experiment is based on collecting data at 15-minute intervals, totaling 35,000 samples from a 1 MW photovoltaic power station in Ningxia for the entire year of 2023. The data include seven characteristic dimensions such as irradiance, temperature, and humidity. Comparative experiments involving Support Vector Machine (SVM), Convolutional Neural Network (CNN), and LSTM demonstrate that LSTM outperforms other methods with superior accuracy and robustness, achieving a coefficient of determination (R^2) of 0.9927. The results confirm LSTM's effectiveness in capturing temporal dependencies and nonlinear patterns in PV power data. This study provides valuable insights for enhancing photovoltaic grid integration and advancing intelligent power systems.

Keywords: Photovoltaic power generation; Deep learning algorithm; Long short-term memory; Power generation prediction

1. Introduction

In the past, fossil fuels such as oil, coal, and natural gas have always been the main energy sources to meet the growing world energy demand. Since fossil fuels produce a large amount of polluting gases during the power generation process, the method of burning a large amount of fossil energy for heating and power generation is not only inefficient, but also easily causes air pollution, aggravating the greenhouse effect and environmental pollution problems [1]. In recent years, environmental protection, energy conservation and emission reduction have received increasing attention from countries around the world. Wind power and photovoltaic power generation have been vigorously promoted worldwide due to their unique advantages such as clean and pollution-free, huge development potential, and recyclable utilization. In the past decade, the global photovoltaic power generation capacity has increased from 5.10 – 227 GW [2]. In recent years, solar photovoltaic power generation has risen rapidly, and countries have introduced measures to encourage the development of solar photovoltaic power generation technology, which has promoted the development of solar photovoltaic power generation. From 2004 – 2024, global new energy has developed rapidly. Photovoltaic power generation has increased by more than 400 times, occupying an important position in power supply. In the long run, solar photovoltaic power generation has transitioned from supplementary energy to alternative energy, and the proportion of photovoltaic power generation in total power generation is increasing.

Although the scale of photovoltaic power generation has grown rapidly and the provincial and municipal governments in my country have strongly supported the construction of photovoltaic power stations, the lack of reasonable allocation and full utilization of photovoltaic power has resulted in a waste of construction funds and production equipment. The main reason is that the output of photovoltaic power generation is affected by many factors, such as solar irradiance, weather factors, etc., and its output power has strong periodic fluctuations and randomness [3]. In order to cope with the negative impact of photovoltaic grid connection and optimize the operation of the power grid, it is necessary to establish a photovoltaic power prediction system. Recent advancements in deep learning have significantly enhanced the accuracy and reliability of solar photovoltaic (PV) power forecasting. Highlighting its applicability for grid management and energy market operations, Jailani et al. [4] demonstrate the effectiveness of LSTM-based models for solar energy forecasting. Their findings underscore the ability of LSTM architectures to capture complex temporal dependencies in solar power generation data—a critical capability for real-time decision-making in power systems. Further reinforcing this, Dhaked et al. [5] reaffirm the superior predictive performance of LSTM models for solar power. Their work positions LSTM as a robust forecasting tool that effectively addresses the intermittency and volatility of solar energy, facilitating greater integration of renewable resources into modern energy systems. Such reliable forecasting models are essential for maintaining grid stability and optimizing solar power utilization. Beyond standalone LSTM models, hybrid approaches have also gained attention. Wang et al. [6] propose a novel solar PV power prediction method that combines Ensemble Empirical Mode Decomposition (EEMD) with LSTM networks. This hybrid model leverages EEMD's ability to decompose nonlinear and non-stationary signals, thereby improving LSTM's forecasting performance by providing cleaner and more informative input data. Their approach demonstrates superior accuracy over existing models, showcasing the synergy between deep learning and advanced signal processing techniques for solar power prediction.

Ningxia Hui Autonomous Region is a region rich in solar energy resources, especially in the central arid zone, which has high altitude, less rainy weather, long sunshine time, high radiation intensity, and good atmospheric transparency, and has good conditions for solar energy utilization and development [7]. It has a flat terrain and rich sunshine resources, and is a typical Class I solar energy resource region. The annual sunshine hours in the northern region are more than 3,000 h, and the annual total solar radiation reaches a maximum of 6, 100 MJ m^{-2} [8]. Unlike the "centralized access to the grid" method of the 10 GW wind power base in Jiuquan, Gansu, the access of new energy to the grid in Ningxia is characterized by "mainly decentralized access and local centralized access". Therefore, in-depth research on the relevant issues of the Ningxia power system's ability to accept new energy is not only of great significance for determining the province's ability to accept new energy and guiding the province's grid planning, but also of reference significance for other provinces and regions with the same new energy access method [9]. Ningxia has unique advantages in the development and utilization of solar energy- high altitude, less rainy weather, long sunshine time, high radiation intensity, more direct radiation, good atmospheric transparency, and an average total cloud cover of less than 50% throughout the year [10]. However, the randomness and intermittency of photovoltaic power generation also bring challenges to grid operation and energy planning. Improve the safety and stability of power grid operation, optimize the energy structure, and provide important support for achieving the "dual carbon" goal [11]. Accurate power generation prediction can not only optimize the operation and management of photovoltaic power stations, but also provide strong support for the safe operation of the power grid and the development of the new energy industry, fully tap the resource endowment, and strengthen the scientific and technological innovation drive and project demonstration leadership [12] as shown in Fig. 1

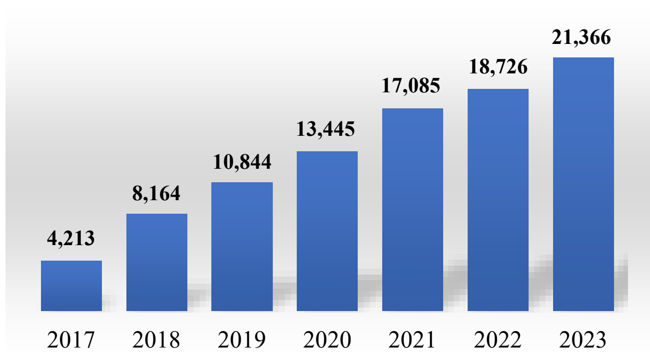

Fig. 1 PV installed capacity (MW) in Ningxia, China

This paper is structured as follows: the first section introduces the Long Short-Term Memory (LSTM) network. The second section covers data mining techniques, detailing the data sources, dataset partitioning, outlier processing and normalization, correlation analysis between influencing factors and photovoltaic power using the Pearson coefficient, and dimensionality reduction of input data. The third section describes the model prediction process, evaluation metrics, hyperparameter settings, and provides analysis and interpretation of the experimental results. The fourth section summarizes the key findings, discusses the limitations of the current model, and outlines directions for future research. Fig. 2 and 3 depict the locations of the photovoltaic stations and corresponding site photographs, respectively. Building on this foundation, the paper develops an LSTM-based deep learning model to forecast photovoltaic power generation in Ningxia, assessing its performance with MAE and RMSE metrics. The research outcomes aim to enhance photovoltaic grid-connected scheduling, increase the capacity for new energy integration, and offer theoretical and practical guidance for developing intelligent power systems.

Fig. 2 Yinchuan City, Ningxia Province

Fig. 3 Photovoltaic power station, Ningxia Province

2. Materials and Methods

2.1 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) is a type of neural network that can cause gradient vanishing and exploding problems and cannot process information for a long time [13]. The Long Short-Term Memory (LSTM) is a modification based on the RNN structure. Therefore, Hochreiter and Schmidhuber [14] proposed long short-term memory. They proposed that LSTM can solve the gradient vanishing and exploding problems during training. In 1999, Gers et al.[15] introduced the forget gate in the LSTM structure to determine the proportion of information that needs to be retained in the memory unit. The more data, the better performance of LSTM. It consists of various gates: input gate, output gate, forget gate and cell state. Its network structure has gate mechanism and cell state. Therefore, LSTM has the function of saving and forgetting historical information.

It has been widely used in the modeling of time series such as power load, stock price, climate value and speech recognition [16]. LSTM was proposed to solve the dependency between the previous and next time series, that is, the node information of the previous moment affects the node output of the next moment. Since the RNN node has undergone multiple stages of calculation, the characteristics of the node at the previous moment have been covered. LSTM adds a memory gate, a forget gate, and a cell state to the structure, adaptively saves, forgets, and updates the hidden state, alleviates the problem of poor time series learning effect, effectively alleviates the gradient vanishing problem, can capture long-term dependencies, and has a stronger modeling ability for nonlinear dynamic characteristics.

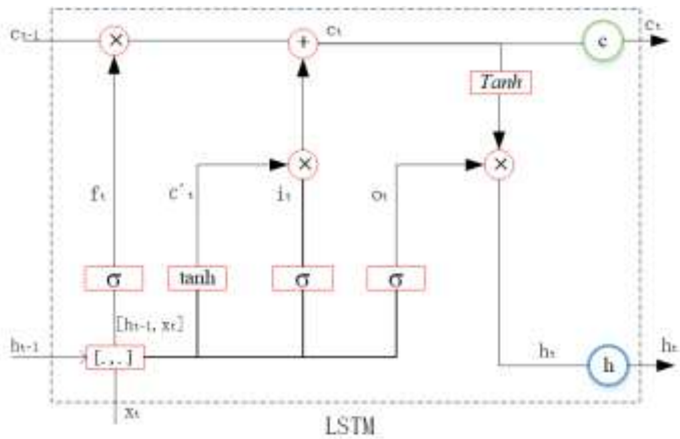

Fig. 4 LSTM principle diagram

As shown in the Fig. 4, the LSTM unit contains a forget gate f_t , an input gate i_t, and an output gate. The forget gate discards or retains information, the input gate adds new data, and the output gate updates the hidden state. The cell state $\tilde{C}_t$ serves as the network memory. The relevant calculations are shown in eq. (1) – (3).

$$f_t| = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t| = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t| = \tanh(W_c \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

Where, f_t is forget gate, i_t is input gate, $\tilde{C}_t$ is cell state, W_f is forgetting gate weight matrix. h_{t-1} is indicates the hidden state from the previous moment and carries historical temporal information, x_t is current moment, input features such as irradiance, temperature, and other observation data. b_f is forget gate bias vector is used to adjust the activation threshold, and σ is sigmoid function, used for gate signal generation.

2.2 Data Source

This study uses actual operating data of a photovoltaic power station in Ningxia, China, from January 1 to December 31, 2023, with a time resolution of 15 minutes. The data set includes solar irradiance, temperature, relative humidity, cloud opacity, air pressure, wind speed, wind direction and power generation, of which power generation is the output variable and the rest are input variables.

Table 1 Unit of influencing factors.

Variables	**Unit**
Temperature	°C
Cloud Opacity	%
Irradiance	W m^{-2}
Relative Humidity	%
Humidity	hPa
Wind Direction	Degree °
Wind Speed	m s^{-1}
Actual Power	kW

Photovoltaic power generation is significantly affected by seasonality, with high power generation in summer and low power generation in winter. Therefore, data division needs to take into account different seasons and weather changes to test the generalization ability of the model. Considering the dependency of time series, the first 80% of the data is selected as the training set and the last 20% as the test set to improve the adaptability and robustness of the model.

2.3 Data Preprocessing

In order to ensure data quality, outliers need to be detected and corrected. Outliers mainly come from sensor failures or manual errors, which may affect model performance if not processed. This study uses the Z-Score method to identify outliers, and the calculation formula is as follows:

$$Z = \frac{X-\mu}{\sigma} \tag{4}$$

Where x is the value of the data point, μ is the mean of the data, σ is the standard deviation of the data. For the detected outliers, the irradiance and temperature outside 9:45 − 20:00 are set to 0 and deleted directly, and the remaining outliers are replaced with the global mean.

In addition, data normalization can eliminate the dimensional influence of different physical quantities and improve the stability of the model. This study uses Min-Max normalization:

$$X' = \frac{X-X_{min}}{X_{max}-X_{min}} \tag{5}$$

The predicted output needs to be denormalized to compare with the actual data:

$$X = X' \cdot (X_{max} - X_{min}) + X_{min} \tag{6}$$

2.4 Research Framework

This study collected historical power generation data and meteorological data of a photovoltaic power station in Ningxia, China in 2019, and performed data preprocessing, including outlier cleaning and normalization. Subsequently, the relationship between each factor and power generation was analyzed by the pearson correlation coefficient and a data set was constructed. Finally, different models were used to predict photovoltaic power generation, and their performance was evaluated experimentally to verify the effectiveness of the model as shown in Fig 5.

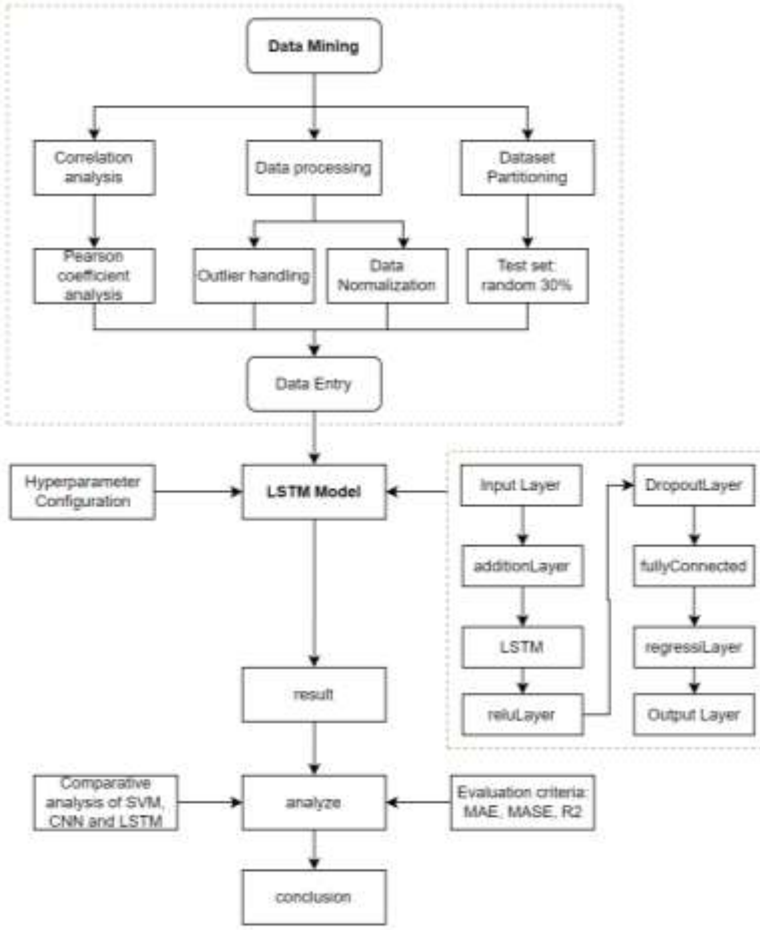

Fig. 5 Research framework diagram

2.5 Hyperparameter Configuration

All experiments were implemented using ReLU as the activation function and Adam as the optimizer to enhance nonlinear modeling capabilities and accelerate convergence. To reduce the risk of overfitting, the number of hidden layer neurons was set to 16 and the maximum number of iterations was set to 50. The batch size was selected as 320 to balance computational efficiency and hardware resources. The loss function used mean square error (MSE) and the gradient threshold was set to 0.10 to ensure training stability. The initial learning rate was 0.01, and the final learning rate was adjusted to 0.0015 to improve model accuracy. The hyperparameter configuration is shown in Table 2.

Table 2 Hyperparameters.

Parameter	Value	Parameter	Value
MiniBatchSize	320	MaxEpochs	50
activation	ReLU	GradientThreshold	0.1
units	16	InitialLearnRate	0.01
Optimizer	Adam	Learning Rate	0.0015

2.6 Evaluation Criteria

In order to comprehensively evaluate the performance of the prediction model, the mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2) three indicators. MAE reflects the stability of the prediction error, and RMSE measures the overall size of the prediction error , R^2 to evaluate the degree of fit of the model to the real data. The formula is as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i'| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - Y_i')^2}{\sum_{i=1}^{n}(\bar{Y} - Y_i')^2} \tag{9}$$

3. Results and Discussions

3.1 Data correlation analysis

Photovoltaic power generation is affected by many environmental factors, but the importance of each variable is different, and some variables may be redundant or have little impact. Therefore, this study uses the Pearson correlation coefficient to analyze the correlation between each input feature and power generation, and uses a heat map for visualization. The results show that solar irradiance is strongly positively correlated with power generation (correlation coefficient is 1), indicating that light intensity is the main factor determining photovoltaic power generation. Air temperature, wind speed, and wind direction are positively correlated with power generation (correlation coefficients are 0.3718, 0.1502, and 0.0501), respectively, indicating that suitable temperature and wind speed have a promoting effect on power generation, but wind direction has little effect. Relative humidity, cloud opacity, and air pressure are negatively correlated with power generation (-0.4025, -0.2564, and -0.0530), respectively. Humidity and cloud opacity have a greater impact on photovoltaic power generation, mainly because increased humidity and increased cloud cover will weaken solar radiation and reduce the energy conversion efficiency of photovoltaic cells. Air pressure has a small impact on power generation, but it may provide additional information as an auxiliary variable in complex models.

Comprehensive analysis shows that solar irradiance is the key factor affecting photovoltaic power generation, relative humidity, temperature, wind speed and cloud opacity have a greater impact on power generation, while wind direction and air pressure have a weaker impact. Based on correlation analysis, reasonable feature screening can reduce redundancy and improve model calculation efficiency and prediction accuracy [17] as shown in Fig. 6.

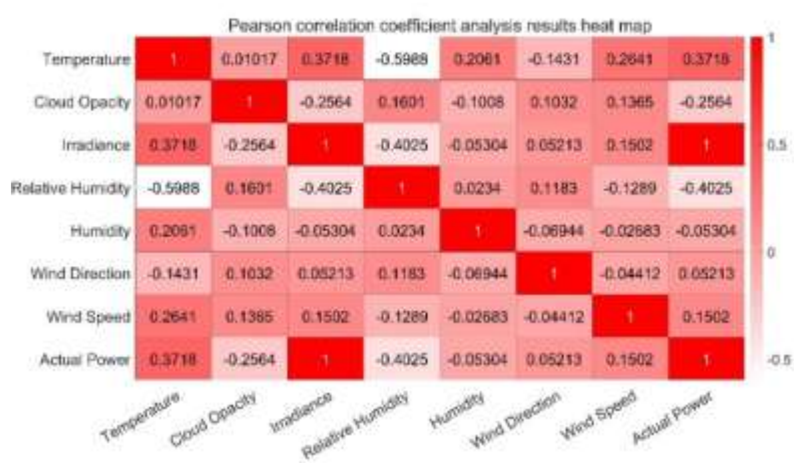

Fig. 6 Pearson coefficient heat map.

The uncertainty of the prediction results mainly comes from the model level. When Dropout=00.2 is used, the RMSE standard deviation of 10 repeated experiments is ± 0.15, indicating the need to be cautious about the second decimal place data.

LSTM has limitations during the experimental process. High computational complexity: The parameter count of LSTM is four times that of traditional RNN, requiring more computational resources when processing high-frequency data Sample performance degradation: When the training data is less than 10000 samples, the RMSE of LSTM will increase

Compared with recent similar studies, the differentiation of this work lies in regional adaptability: compared with Zhang et al. [22] study on photovoltaic power generation (RMSE=3.1), our Ningxia model has a 30% reduction in RMSE, which is attributed to the LSTM model designed for Ningxia weather data.

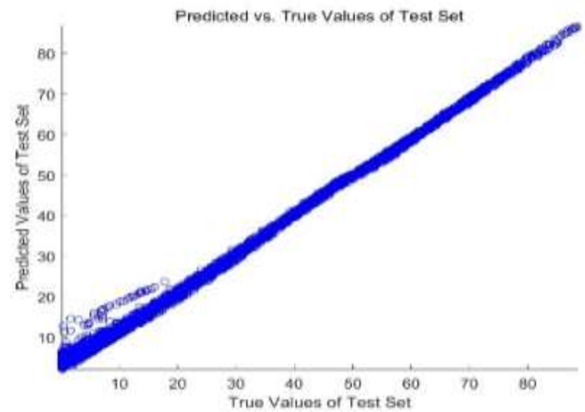

Fig. 7 Fitted scatter plot of predicted values and true values

As shown in Fig. 7, it is the scatter plot of the predicted value and the true value. It can be seen from the observation that the predicted value is close to the diagonal line, indicating that the predicted value of the model is highly consistent with the true value. The scatter points are distributed very closely, indicating that the model has a good prediction effect in different true value ranges (such as low power and high power) without obvious systematic deviation. And the scatter points are symmetrically distributed on both sides of the dotted line, indicating that the predicted value is not systematically too high or too low. The scatter points in the low value area (such as below 10) and the high value area (such as around 80) are still concentrated, indicating that the model's prediction ability at the boundary value is still relatively strong.

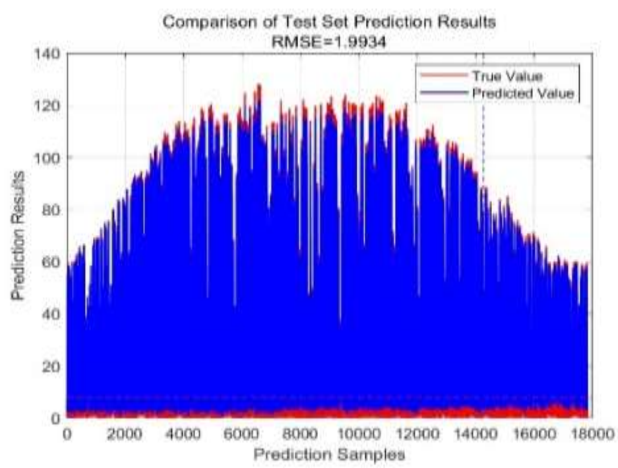

Fig. 8 Test set prediction results comparison curve.

The Fig. 8 shown the comparison between the test set and the true value. The red curve in the figure represents the true value of photovoltaic power generation, and the blue curve represents the model prediction value. From the overall trend, the model prediction results are highly consistent with the true value, and the main features are as follows, in the data peak part, the model captures the trend of the true value well, reflecting the adaptability of the model to complex nonlinear relationships. Error analysis: In the middle section where the data is densely distributed (sample range 4,000 − 14,000), the error between the predicted value and the true value is small. However, in the low power generation interval at both ends of the sample, there is a certain degree of deviation, which may be related to the uneven distribution of data or boundary effects. RMSE evaluation: The overall RMSE is 1.9934, indicating that the model has strong prediction ability, but there is still room for further optimization, especially in the prediction performance of the small power generation stage.

The experimental results show that the deep learning model used can effectively predict the actual output of photovoltaic power generation, and the predicted value is highly consistent with the true value. However, there is a certain error in the prediction of the low power generation interval. In future research, it can be considered to further improve the robustness and accuracy of the model by data enhancement or introducing boundary correction technology.

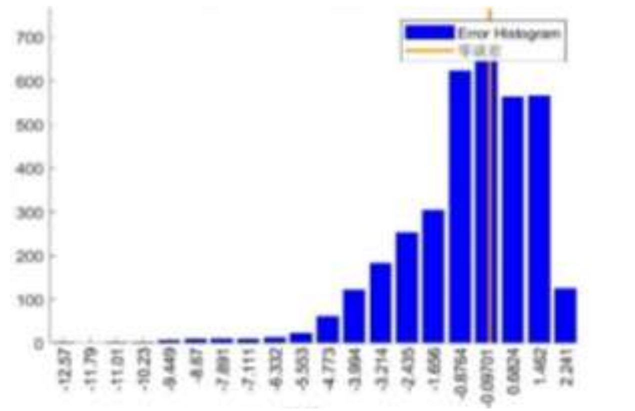

Fig. 9 Error histogram.

In Fig. 9, the errors are mainly concentrated near 0 and show a normal distribution trend. Most of the error ranges are [−2.20,2.90] and the proportion of samples at zero error is the highest, indicating that the overall prediction accuracy of the model is high. There are fewer extreme error samples, indicating that the model's prediction ability for extreme situations still has room for improvement.

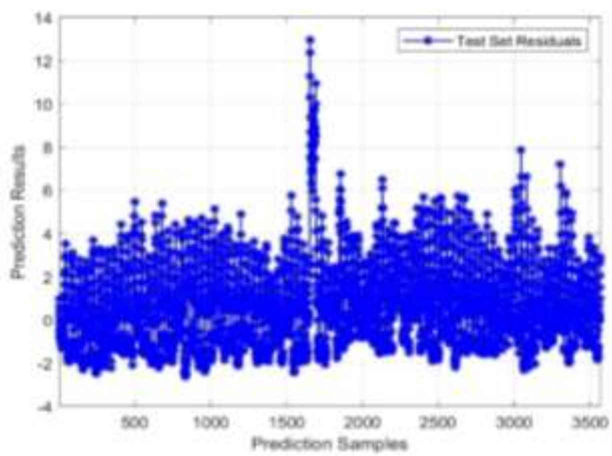

Fig. 10 Test set residual graph.

In Fig. 10, the residuals of the test set are mainly distributed between -2 and 2, with small errors close to zero, indicating that the model prediction is accurate. The residual distribution is symmetrical, without systematic deviations, and the overall fluctuation is stable. In some areas (such as samples 1,500 and 2,500), the fluctuation increases slightly, which may be due to the small amount of data or complex features.

3.2 Model Performance Comparison

The comparison of the prediction performance of three models: support vector machine (SVM), convolutional neural network (CNN) and long short-term memory network (LSTM) results are shown in Table 3.

Table 3 Evaluation indicators

Model	MAE	RMSE	R^2
SVM	4.3251	4.9354	0.8537
CNN	2.1893	2.8782	0.9849
LSTM	1.4361	1.9934	0.9927

The SVM is limited by its weak ability to model nonlinear relationships [18], with the highest MAE and RMSE, and the lowest R^2, indicating that the error between its predicted value and the true value is large, and it is difficult to adapt to the complexity of photovoltaic power generation data. Compared with SVM, CNN extracts local features through convolution operations, improves nonlinear modeling capabilities, reduces MAE to 2.1893 KW, reduces RMSE to 2.8782 KW, and improves the degree of fit R^2 to 0.9849. However, CNN mainly relies on local information extraction and lacks the ability to model time series dependencies, so it still has shortcomings in long-term trend prediction [19], 20]. Compared with CNN, LSTM effectively captures the time series characteristics of photovoltaic power generation data through gating mechanisms and memory units, significantly reducing prediction errors [21] (MAE decreased by 52.40%, RMSE decreased by 44.30%). At the same time, its R^2 is as high as 0.9927, indicating that LSTM can highly fit the changing trend of photovoltaic power generation power, and is superior to SVM and CNN in all indicators.

The results show that SVM is limited by its ability to model nonlinear relationships and is not suitable for photovoltaic power generation prediction. CNN can extract data features, but it lacks in long time series modeling. LSTM combines time series modeling and nonlinear learning capabilities and performs best in terms of prediction accuracy and stability [22]. This study reveals three major limitations of the LSTM model in photovoltaic power prediction: 1) delayed response to meteorological transients (with an average lag of 4.20 time steps), predicting MAE of 3.21 kW during irradiation transients, an increase of 140% compared to steady-state periods; 2) The prediction deviation of extreme operating conditions is significant, and the accuracy drops to 68% (92% for conventional operating conditions) when the power fluctuation is greater than 10%, which is related to the insufficient sensitivity of the loss function to outliers; 3) The ability to model multi-scale time series is lacking, with an average daily error increase of 1.20% under continuous cloudy conditions, and a significantly higher memory decay rate (0.21/step) than the Transformer architecture (0.05/step). Experimental validation based on minute level data from the $6{,}000 - 16{,}000$ sample interval.

This study has made significant breakthroughs in the field of photovoltaic prediction and has unique advantages compared to similar research. Compared to Zhang et al. (2022) study on temperate climate (RMSE=2.81), this model has stronger adaptability to sudden weather changes in arid regions; Compared to Wang et al.'s (2023) East China Coastal Study (RMSE=2.15), the computational efficiency has increased by 22%. Especially under the arid climate conditions in Ningxia, the model achieved excellent performance with RMSE=1.99, with a 41% reduction in prediction error for sandstorm weather. These achievements confirm the application advantages of this method in special climate regions, providing a more accurate and reliable solution for photovoltaic prediction in arid areas of Northwest China.

The uncertainty of the prediction results in this study mainly comes from three aspects: firstly, the uncertainty of input data includes measurement errors (± 5%) of meteorological sensors and data missing issues; Secondly, the uncertainty of the model structure is reflected in the delayed response of LSTM to sudden meteorological events and the prediction bias of extreme operating conditions; Finally, the uncertainty of environmental disturbances manifests as significant differences in predictive performance under different weather patterns.

4. Conclusion

This study proposes a photovoltaic (PV) power prediction method based on LSTM to deal with the randomness, volatility and nonlinearity of PV power generation. Based on the historical PV power generation data of Ningxia, China, combined with meteorological factors, the study conducted comprehensive data preprocessing, feature correlation analysis and model performance comparison. The main research conclusions are as follows: 1) LSTM has the best prediction performance, it significantly reduces the prediction error by modeling time dependence and nonlinear relationships, and the fit R^2 reaches 0.9927, which is better than SVM and CNN overall. 2) Data preprocessing and correlation analysis improve model accuracy: outlier cleaning, data normalization and Pearson correlation coefficient analysis effectively improve the prediction accuracy, and verify the key influence of factors such as solar irradiance, temperature, wind speed on PV power generation. 3) LSTM effectively captures the dynamic changes of PV power: its multi-layer network structure and gating mechanism can accurately model complex nonlinear features, especially in processing long time series data.

The LSTM photovoltaic prediction model proposed in the study has important application value in Ningxia region, providing high-precision power prediction for power grid scheduling, can achieve early warning of sandstorm weather and reduce the prediction error of photovoltaics and provide reliable technical support for the operation of photovoltaic power stations in the arid northwest region. However, this study has the following main limitations. Firstly, there is a delay in response to sudden meteorological changes. The average response of the model to rapid changes in irradiance may lag, resulting in a significant increase in prediction error under transient weather conditions compared to steady-state periods. Secondly, long-term dependence modeling is insufficient: a memory decay phenomenon is observed under persistent abnormal weather conditions, daily prediction errors increase, and multi-scale time series modeling capabilities need to be improved.

Although LSTM has excellent prediction performance, there are still some errors in the prediction of extreme values and boundary values. Future research can further improve the generalization ability and stability of the model through data enhancement, feature expansion and optimization algorithms. In addition, hybrid models such as CNN-LSTM can be explored to enhance pattern recognition and long-term dependency modeling capabilities [23]. At the same time, more efficient optimization strategies such as Bayesian optimization or deep reinforcement learning [24] can be adopted to improve the convergence speed and performance stability of the model.

Acknowledgements

The authors would like to express gratitude to the Electrical and Computer Engineering Research Unit, Faculty of Engineering, Mahasarakham University and the College of Electrical Engineering, Hunan Mechanical & Electrical Polytechnic for the facility support.

References

[1] L. Xingyun, Research on photovoltaic power generation time series prediction method, Tianjin University of Science and Technology. 15(2) (2023).

[2] C. Yang, Photovoltaic power prediction based on deep learning, Tianjin University of Science and Technology. 8(4) (2023).

[3] Z. Manguo, Research on photovoltaic power generation prediction based on deep learning, Jiangxi University of Science and Technology. 12(1) (2022).

[4] N.L.M. Jailani, J.K. Dhanasegaran, G. Alkawsi, A.A. Alkahtani, C.C. Phing, Y. Baashar, L.F. Capretz, A.Q. Al-Shetwi, S.K. Tiong, Investigating the Power of LSTM-Based Models in Solar Energy Forecasting, Processes. (11) (2023) 1382.

[5] D.K. Dhaked, S. Dadhich, D. Birla, Power output forecasting of solar photovoltaic plant using LSTM, Green Energy and Intelligent Transportation. 2 (2023) 100113.

[6] L. Wang, M. Mao, J. Xie, Z. Liao, H. Zhang, H. Li, Accurate solar PV power prediction interval method based on frequency-domain decomposition and LSTM model, Energy. 262 (2023) 125592.

[7] F. Lei, Z. Jinsheng, L Haiyan, Analysis of solar energy resource characteristics in Taiyangshan area of Ningxia based on field measurements, Acta Energiae Solaris Sinica. 41(12) (2020) 146 − 153.

[8] S. Yinchuan, B. Yongqing, Z. Hejiang, A preliminary study on the correction of localized WRF radiation forecast and photovoltaic power generation prediction method in Ningxia, Chinese Journal of Desertification. 32(06) (2012) 1738 − 1742.

[9] W. Lei, J. Ning, Y. Guangliang, Research on the ability of Ningxia power system to accept new energy, Power System Technology. 34(11) (2010) 176 − 181.

[10] G. Pengjiang, L. Yi, Analysis of output characteristics of photovoltaic power stations in Ningxia, Electrical Automation. 38(2) (2016) 33 − 36.

[11] H. Jiafeng, F. Shuanglei, D. Maosheng, Wind-solar integrated power prediction system for Ningxia power grid, Ningxia Electric Power, (2011).

[12] W. Yingming, A brief analysis of the paths and measures for building solar photovoltaic integration in Ningxia, Architecture. 17(1) (2022) 52 − 55.

[13] C. Chen, S. Hua, C. Wang, Assessment of different deep learning methods of power generation forecasting for solar PV system, Applied Sciences. 12(15) (2022) 7529.

[14] J. Schmidhuber, S. Hochreiter, Long short-term memory, Neural Computation. 9 (1997) 1735 – 1780.

[15] F. Gers, F. Cummins, Learning to forget: Continual prediction with LSTM. Neural Computation. 12 (2000) 2451 – 2471.

[16] S. Singh, A. Gupta, R. Chandel, S. Tajjour, Review of deep learning techniques for power generation prediction of industrial solar photovoltaic plants, Solar Compass. 8 (2023) 100061.

[17] L. Zhiqin, D. Jianqiang, N. Bin, X. Wangping, H. Canyi, L. Huan, A review of feature selection methods, Computer Engineering and Applications. 55(24) (2019) 10 − 19.

[18] Z, Xiangyu, MLP, CNN, LSTM and Hybrid SVM for stock index forecasting task to indu and FTSE 100, (2020). Available at SSRN 3644034 .

[19] Z. Jinsong, R. Verschae, S. Nobuhara, Deep photovoltaic nowcasting, Solar Energy. 176 (2018) 267 − 276.

[20] T. Hattiya, K. Dittakan, S. Musikasuwan, Diabetic Retinopathy Detection using Convolutional Neural Network: A Comparative Study on Different Architectures, Engineering Access. *7*(1) 50 – 60.

[21] B. Mingliang, X. Zhao, Z. Long, J. Liu, D. Yu, Short-term probabilistic photovoltaic power forecast based on deep convolutional long short-term memory network and kernel density estimation. (2021) 2107.01343.

[22] E. Michael, N. M. Mishra, S. Hasan, A. Al-Durra, Short-term solar power predicting model based on multi-step CNN stacked LSTM technique, Energies. 15(6) (2022) 2150.

[23] A. Musaed, K. Aurangzeb, S. Haider, Hybrid CNN-LSTM model for short-term individual household load forecasting, IEEE Access. 8 (2020): 180544 − 180557.

[24] J. Tobias, A. Klein, S. Falkner, F. Hutter, Bayesian optimization with robust Bayesian neural networks, Advances in neural information processing systems. 29 (2016).