Journal of Climate Change
ISSN 2395-7611 Volume 12 (2026) No. 1. pp. 32–45

https://doi.org/10.70917/jcc-2026-002
Article

Ultra-short-term power prediction of photovoltaic power systems based on multi-climatic environments

Zhengqi Fu[1,*]

[1] School of Electrical Engineering and New Energy, China Three Gorges University, Yichang, Hubei, 443000, China
* Correspondence author: 2822336882@qq.com

Abstract: Ultra-short-term power prediction of photovoltaic (PV) power generation system is an important basis for grid scheduling and energy management. In order to adapt to the complex and changing climatic conditions, this paper proposes an MHSA-LSTM ultra-short-term power prediction model for PV power generation system that integrally considers all kinds of climatic factors. In this study, the raw data of a PV power plant is processed by multiple interpolation, and four key climate environment variables, namely, temperature, irradiance, relative humidity and atmospheric pressure, which affect the power generation, are extracted from them. The time series data containing the influencing variables are used as samples, and the MHSA module is utilized to filter the importance of the data at different historical moments, and the strong and weak weights of each environmental variable are adaptively assigned to the LSTM network to obtain the prediction results. The results show that climatic factors such as solar irradiance, temperature, relative humidity and atmospheric pressure can affect the PV power, with irradiance having the most significant effect.The MHSA-LSTM model is oriented to the climatic conditions of different seasons and weather, and the predicted ultrashort-term power is closer to the actual power value. Taking spring conditions as an example, the MAPE\RMSE of the MHSA-LSTM coupled model is reduced by 10.33%\2.771kW, 23.12%\6.296kW, 5.34%\1.234kW, compared to BP-LSTM, LSTM and LSTNet, respectively.The fluctuation of the ultra-short-term power prediction model based on deep learning is basically the same as that of the actual fluctuation. It is more adapted to the subsequent grid power scheduling and operation requirements.

Keywords: photovoltaic power generation; ultra-short-term power prediction; deep learning; MHSA-LSTM; multi-climatic environment

1. Introduction

With the promotion of the “dual carbon” strategy and the development of photovoltaic (PV) technology, PV installations continue to grow. It is predicted that the global PV installation will reach 1700 GW in 2030, and it is expected to be responsible for about a quarter of the global electricity supply in 2050 (Li et al., 2022; Jo et al., 2024). This puts higher requirements on the stability, scheduling and power prediction of PV power generation. PV power is stochastic, fluctuating and intermittent, affected by system factors (e.g., conversion efficiency, aging) and environmental factors, with the former varying slowly and the latter transiently and variably. After grid connection, these changes are easy to disturb the grid, damage equipment, and cause economic losses. Power prediction can anticipate the changes in advance, optimize dispatch, reduce standby and operation costs, and improve grid quality and reliability (Khan et al., 2023). Therefore, it has become a research hotspot to construct a power prediction model with local sensitivity modeling, physical consistency guarantee and multi-climate adaptation capability.

PV power ultra-short-term prediction methods are categorized into three types: physics, statistics and artificial intelligence (Tsai et al., 2023). Among them, physics methods combine key

meteorological parameters and module characteristics to construct physical mechanism constraint models (Zhang et al., 2025). For example, Yang & Huang (2018) in their study proposed an ultra-short-term prediction method for PV power, which considers the periodicity of solar irradiance, extracts its periodic component, utilizes a locally sensitive hash algorithm for fast classification based on weather type, and introduces Euclidean distance for prediction. Such methods do not rely on historical generation data, but their prediction accuracy is heavily dependent on the quality of weather data and model parameter settings. In contrast, statistical methods are based on the regularity of historical data and construct mapping relationships between data and power output with the help of, e.g., regression analysis and its extended models (Jagadeesh et al., 2020). For example, the ultra-short-term prediction method of PV power proposed by Souza, G constructs a nonlinear mixed-effects model based on the log-logistic Steele function, which reflects the nonlinear relationship between power generation and environmental conditions and realizes efficient prediction of power generation (Souza et al., 2022). Abdullah et al. (2024) compared the performance of various regression algorithms in PV power prediction and found that the multilayer perceptron regressor performs best and is able to identify the key factors affecting PV power generation. Such methods have certain advantages in trend extraction and cycle modeling, and are relatively simple to implement. However, the prediction ability is limited when facing nonlinear and nonsmooth time series data.

In order to overcome the limitations of the above methods, artificial intelligence techniques have been widely introduced into the PV power prediction task, mainly including 2 types of machine learning and deep learning. Machine learning algorithms commonly used for PV power prediction are, including support vector machine, random forest, decision tree, gradient boosting decision tree, etc. For example, Rinesh et al. (2024) proposed a solar PV power prediction model based on extreme gradient boosting regression, which utilizes key features such as humidity, temperature, clear-sky index, and moments, and is able to effectively deal with complex nonlinear relationships. Saxena et al. (2024) fused K-nearest neighbor and SVM algorithms to construct a hybrid machine learning model to improve the prediction accuracy of solar photovoltaic power. Zazoum (2022) explored the relationship between multiple input parameters and output power when predicting solar photovoltaic power using a machine learning model, while comparing two machine learning methods, SVM and Gaussian process regression. Niu et al. (2020) used RF to screen the key factors affecting power generation, passed the importance weights computed by RF into an improved gray ideal value approximation model, and screened the similar days of different weather types to improve the quality of the training set; subsequently, the original power sequence was decomposed into components of different frequencies by complementary ensemble empirical mode decomposition to reduce the sequence volatility. Banik & Biswas (2023) proposed a hybrid model combining RF and CatBoost algorithms for long term prediction of solar irradiance and PV power and the combined model performed excellently with a 6% improvement in accuracy compared to the existing models. Mahesh et al. (2022) proposed a machine learning algorithm based on decision tree regression for maximum power point tracking in stand-alone photovoltaic power systems. Machine learning has some advantages over traditional statistical models in dealing with nonlinear input-output relationships, but its practical prediction is limited due to its lack of ability to model time-series dependent structures.

In contrast, deep learning methods are able to utilize their powerful feature extraction and representation capabilities to more effectively mine the temporal dependence and high-dimensional nonlinear patterns between PV power and meteorological variables (Huang & Yang 2023). For example, Su et al. (2025) proposed a deep learning model based on a gated recurrent unit network with a dual attention mechanism for ultra-short-term prediction of PV power, which is able to better extract the key information by adopting an encoder-decoder framework and incorporating a feature and temporal dual attention mechanism. Wang et al. (2024) proposed a deep learning model named ABCGRU for ultra-short-term prediction of regional distributed PV plant power, which effectively exploits the spatio-temporal characteristics of PV power generation by integrating a self-attention mechanism, a bidirectional convolutional gated loop unit, and an encoder-decoder structure. Yong et al. (2025) proposed a multimodal model called ConvODE-Mixer for improving the accuracy of ultra-short-term prediction of photovoltaic power, which incorporates a convolutional neural network and god-often differential equations. Zhou et al. (2025) proposed a hybrid model for ultra-short-term prediction of photovoltaic power, which utilizes CEEMDAN to decompose the original power series to reduce data volatility, and constructs a combined AM-TCN-BiLSTM network to extract features and learn them, optimizes the model hyper-parameters by using the RIME algorithm, and finally reconstructs the prediction results of each sub-sequence. Liu et al. (2025) proposed an ultra-short-term PV power forecasting model based on wavelet decomposition, dual-attention mechanism and bi-directional long- and short-term memory network, which utilizes wavelet decomposition to deal with the volatility and nonlinear characteristics of power data, and enhances the integrated learning capability of time series

through BiLSTM and dual-attention mechanism.

In recent years, Transformer and its derived models have demonstrated a strong capability in handling long-time dependencies and complex nonlinear features. Tang et al. (2025) proposed a hybrid model for short-term PV power prediction based on feature construction and an improved Transformer architecture, which demonstrated high stability and generalization performance in handling both short- and long-series tasks. Xiong et al. (2025) proposed a probabilistic PV power ultrashort-term prediction framework combining sky images and historical power data, which is based on Transformer's multimodal model, which combines a sequence of predicted images with historical power data to derive a probabilistic distribution of future power. Ye et al. (2026) proposed a multiscale CNN-Transformer network called MSCT-RCM for ultra-short-term prediction of PV power, which incorporates a parallel multiscale CNN for extracting local temporal features, a Transformer encoder for capturing long-range dependencies, and a residual correction module for dynamically correcting the prediction based on historical error patterns. Zhang et al. (2026) proposed an ultrashort-term PV power prediction method based on a temporal attention-variable parallel fusion coding network in response to the significant degradation of the ultrashort-term prediction accuracy of the existing methods at 3-4 h. The model corrects the temporal prediction mainly through the introduction of numerical weather forecasting data. Zhai et al. (2025) stacked a long- and short-term neural network with the Transformer structure while integrating variational modal decomposition and sparrow search algorithms to achieve effective capture of complex relationships between meteorological variables and power time series. Transformer and its variants in the ultra-short-term prediction of photovoltaic still exists in the local mutation response insensitive, non-smooth modeling is limited and the output rationality is insufficient, etc., and the stability and generalization ability of the model under the complex multi-climate is challenged.

In this paper, the data of a photovoltaic power station from March 1, 2023 to February 28, 2025 are collected, and the anomalous missing values in them are complemented based on the multiple interpolation method. Combined with the photovoltaic output power characterization, the key environmental variables are screened out based on the Pearson correlation coefficient method, and multiple types of weather and seasons are divided to construct an ultra-short-term power prediction model of the PV power generation system for the multi-climate environment. The model consists of two parts: multi-head attention mechanism (MHSA) and long and short-term memory neural network (LSTM), MHSA mines out the importance of historical data moments, establishes the corresponding weights, and comprehensively takes into account the contribution of strong and weak characteristic variables to the power prediction, while the LSTM takes the weighted power generation time-series data as the input, and predicts the power output through multi-step learning, and finally the prediction results of multi-climatic environment are compared with those of BP-LSTM. The prediction results are finally compared with models such as BP-LSTM.

2. Photovoltaic power generation system raw data acquisition and pre-processing

In this paper, the data of a PV power station from March 1, 2023 to February 28, 2025 are selected as samples, and the meteorological conditions such as irradiance, temperature, humidity and barometric pressure are collected by weather stations, and the output power is measured by PV inverters/meters, and the raw data are automatically obtained by a set sampling frequency. For the abnormalities and missing measurements in the raw data, pre-processing is required to eliminate deviations. The missing measurement items adopt multiple interpolation method to make up the average value of the credible observations before and after the missing point, and the mathematical form is shown in equation (1).

$$\hat{\chi}_s = \frac{\chi_{i-1} + \chi_{i+1}}{2} \tag{1}$$

where $\hat{\chi}_s$ is the missing value estimate and χ_{i-1} and χ_{i+1} are the data values at the moments before and after the missing position, respectively.

3. Characterization of photovoltaic output power

Pearson's correlation coefficient is a common measure of the linear association between two variables, with a value ranging from -1 to 1. Take X and Y as an example: a correlation coefficient of 0 means no linear association; a positive correlation in [0,1], where an increase in X leads to an increase in Y; and a negative correlation iS where an increase in X leads to a decrease in Y. The closer the absolute value is to 1, the stronger the correlation is. The closer the absolute value is to 1, the stronger the correlation is. In this paper, the Pearson correlation coefficient is used to assess the impact of

meteorological factors on PV power, extract meteorological and power information based on historical data, calculate the full-cycle mean and daily mean to construct the data set, and obtain the correlation coefficient between PV power and each meteorological factor according to equation (2).

$$r = \frac{\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{P}_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(\overline{x}_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(\overline{P}_i - \overline{P})^2}} \tag{2}$$

where x denotes the meteorological element, P is the power, $\overline{x}_i$ is the meteorological data of the i th day, $\overline{p}_i$ is the average power of the i th day, $\overline{x}$ is the average power for the whole year, and $\overline{p}$ denotes the number of dates. n is the number of days. This study was calculated with the help of SPSS and the results are summarized in Table 1 below. The analysis shows that solar irradiance has the most significant effect on PV power, showing the strongest positive correlation; relative humidity and humidity are next in line, showing a negative correlation; and atmospheric pressure is positively correlated with power.

Table 1. Pearson coefficient of photovoltaic output power and meteorological factors

Data type	Pearson coefficient
Solar radiation intensity(W/m^2)	0.905
Temperature(℃)	-0.634
Wind speed(m/s)	0.189
Humidity(%)	-0.341
Atmospheric pressure	0.143

3.1. Effect of season on photovoltaic output power

Figure 1 illustrates the intraday time-series variation of PV power during the four seasons under clear conditions. The power is higher in spring and fall; in summer the power is lower than in spring and fall but the output duration is longer; and in winter the power is the lowest and shortest in duration. Differences in sunrise and sunset and solar radiation angle in different seasons make the peak power vary seasonally. Therefore, categorizing the prediction by season can effectively reduce the difficulty of prediction.

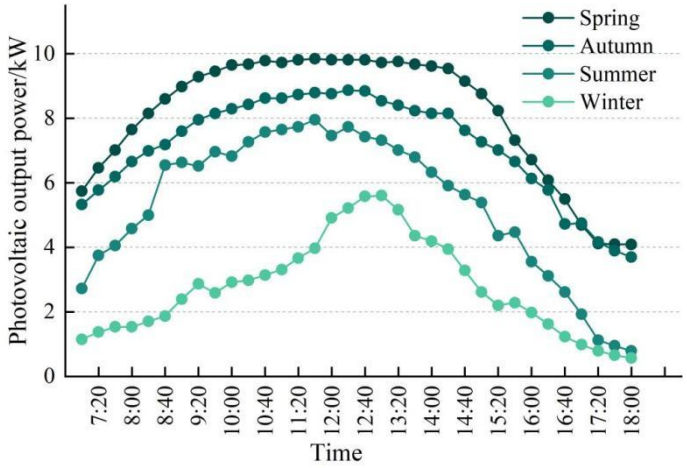

Figure 1. Comparison of photovoltaic power generation across different seasons

3.2. Effect of weather type on photovoltaic output power

The weather category constrains the day's temperature and irradiation intensity to specific fluctuation intervals, which in turn affects the variation characteristics and power output range of the PV system. Accordingly, this paper reclassifies the weather into three categories: sunny, rainy and cloudy.

The real-time power output of a PV plant and its variation process are affected by meteorological conditions. Figure 2 demonstrates the time series of power generation under three types of weather,

namely, sunny, rainy and cloudy (cloudy) days. Among them, sunny days have the highest peak output and the longest duration; there are some differences compared to cloudy days, while the most significant difference is with rainy days, which have the lowest peak power and the shortest duration. In addition, the power curve on a sunny day is relatively smooth, with no obvious ups and downs; the curve fluctuation is intensified under cloudy conditions, which is due to the frequent changes in the thickness and coverage of the cloud layer, resulting in large fluctuations in the solar irradiance that reaches the surface of the module; and under rainy conditions, the power series has the most violent oscillations and the largest number of change points. It can be seen that the pattern and amplitude of power changes under the same category of weather are basically the same, while the differences between different weather are obvious, so the power prediction should be categorized according to the weather category first.

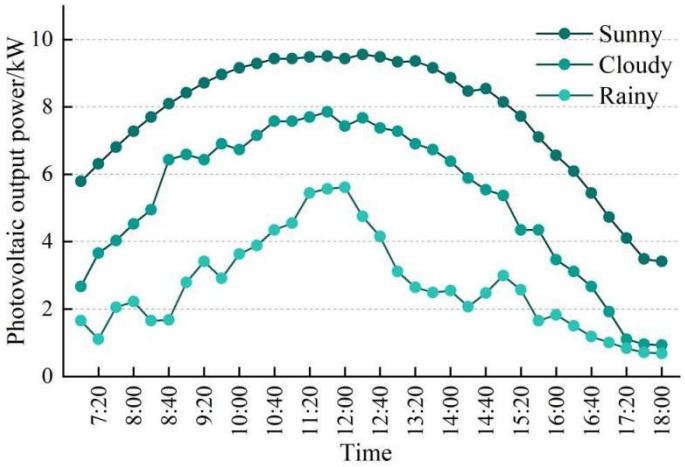

Figure 2. Comparison chart of photovoltaic output power under different climates

4. PV power prediction model based on MHSA-LSTM

4.1. Submodels

4.1.1. Multi-pronged self-attention mechanisms

Figure 3 illustrates the structure of multi-head attention. This multi-head self-attention is mapped to obtain a matrix of three types of queries Q, keys K and values V: Q interacts with K to generate the attention weights, K extracts the key information in the sequence by linear transformation, and V carries the content representations of each position. Different aspects of the sequence are attended to through h parallel self-attention heads, and the outputs of each head are subsequently spliced in the feature dimension, and then the splicing results are mapped to the target dimension in the following form:

$$H_t = \{h_{t-k}, h_{t-k}, \cdots, h_t\} \tag{3}$$

$$Q = HW^Q \tag{4}$$

$$K = HW^K \tag{5}$$

$$V = HW^V \tag{6}$$

$$Attention(Q,K,V) = \sigma\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

$$W_{MultiHead} = Concat(head_1, \cdots, head_h) \tag{8}$$

where H_t is the LSTM model output vector. k is the size of the model time step. W^Q, W^K, and W^V are three coefficient matrices. d_k is the dimension size of matrices Q and K. $head_i$ is the attention output of the i th head. $Concat$ is the splicing function.

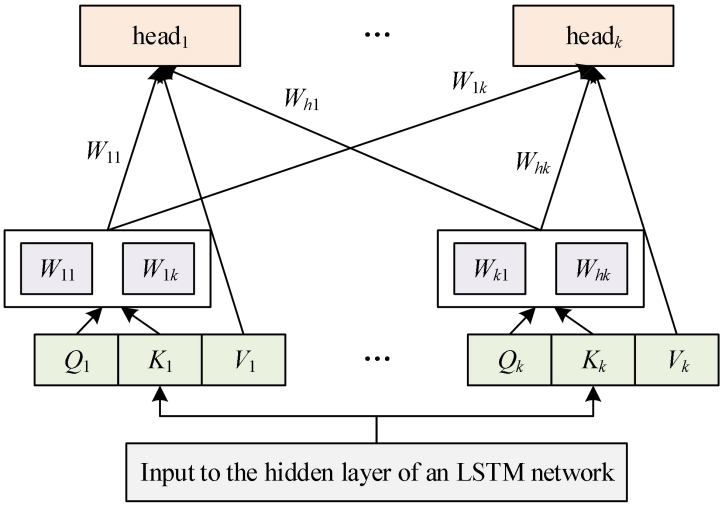

Figure 3. Head Attention Mechanism

4.1.2. Long and short-term memory networks

Long Short-Term Memory (LSTM) is a specialized form of recurrent neural network, which overcomes the problem of gradient vanishing and exploding by introducing gating units, and has the ability to model long term dependence, and its structure is shown in Fig. 4.

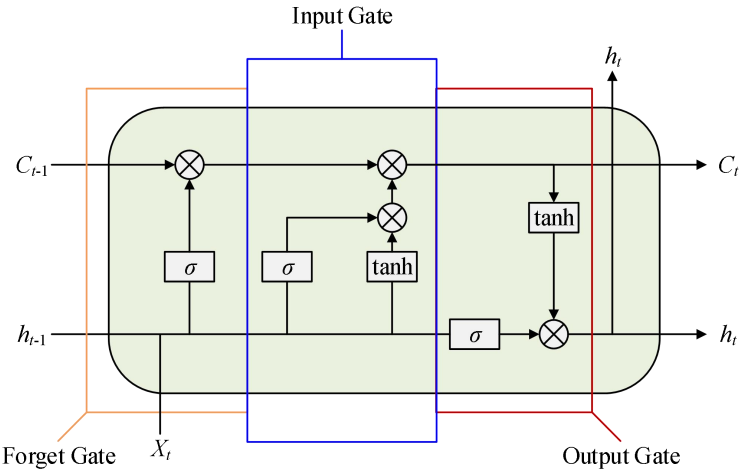

Figure 4. The network structure of LSTM

Fig. 4 illustrates the processing flow of the LSTM network: after inputting the PV power data, the sequence is traversed by time steps to extract the hidden states at each step. The model update formula is shown in Eq. (9)-Eq. (14).

$$f_t = \sigma(W_F \cdot [h_{t-1}, x_t] + b_F) \tag{9}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{10}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{11}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{C}_t \tag{12}$$

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \tag{13}$$

$$h_t = o_t \odot \tanh(c_t) \tag{14}$$

In Eq. f_t denotes the output of the forgetting gate, and σ table is the sigmoid activation function. W_F and b_F correspond to the weight matrix and bias vector of the forgetting gate,

respectively. h_{t-1} denotes the hidden state of the previous time step and x_t denotes the input of the current time step. i_t denotes the output of the input gate and $\tilde{C}_t$ is the candidate memory value to be updated. c_t denotes the state of the memory cell at the current moment. W_i, W_c and b_c are the corresponding weight matrices with bias terms, respectively. o_t is the output of the output gate and h_t is the hidden layer state at the current moment. W_o, b_o are the corresponding weight matrices with bias terms, and c_t is the cell state at the current time step.

4.2. Constructing the MHSA-LSTM model

In this chapter, a hybrid model combining MHSA and LSTM is proposed for PV power prediction, see Fig. 5, in which MHSA enhances the key information and suppresses the invalid features by assigning weights to historical weather and power data. Meanwhile, LSTM adopts a cell-by-cell information transfer approach, which leads to a long information transfer path between distant locations in the sequence and makes it difficult to fully retain long-range dependencies; while MHSA can directly portray the correlation between any two locations in the sequence, which significantly shortens the information interaction path between the two points, and thus better captures the long-range dependencies.

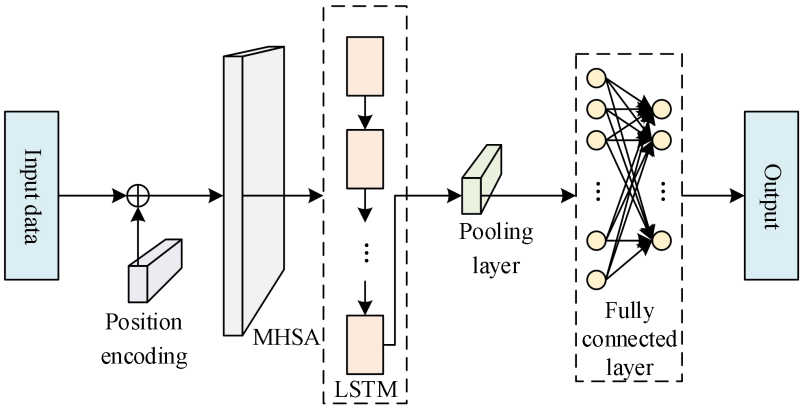

Figure 5. Structural diagram of MHSA-LSTM

4.3. Ultra-short-term prediction of PV power based on MHSA-LSTM modeling

Photovoltaic power generation power presents various characteristics in different seasons with the weather conditions of the day. Given that the weather factor can be regarded as a component of seasonal factors, in this paper, after completing the data preprocessing, the collected data from a PV power plant is divided into four types of sub-datasets by season and inputted into the MHSA-LSTM coupled model for training and power prediction respectively.

As shown in Fig. 6, the experimental flow of PV power prediction using the MHSA-LSTM model has been given, and the detailed steps of its prediction are as follows:

Step 1: The historical weather and power data are processed with vacancies and anomalies and normalized to a uniform scale.

Step 2: Evaluate the preprocessed data using Pearson's correlation coefficient and screen the features that are strongly correlated with PV power as model inputs.

Step 3: Split the input data into four parts by four seasons, which are used as the training set for each sub-model of MHSA-LSTM.

Step 4: For the test set, predictions are made using the sub-models of the corresponding seasons, and the outputs are inverse normalized to obtain the final predicted values for each season.

Step 5: On the basis of preprocessing, screen the data from 7:00 to 18:00 every day, construct the training set and test set according to different prediction steps, and train the MHSA-LSTM sub-models under each step one by one.

Step 6: Call the trained sub-models to generate predictions according to different prediction steps, and reduce the output to the original scale to obtain the final prediction results.

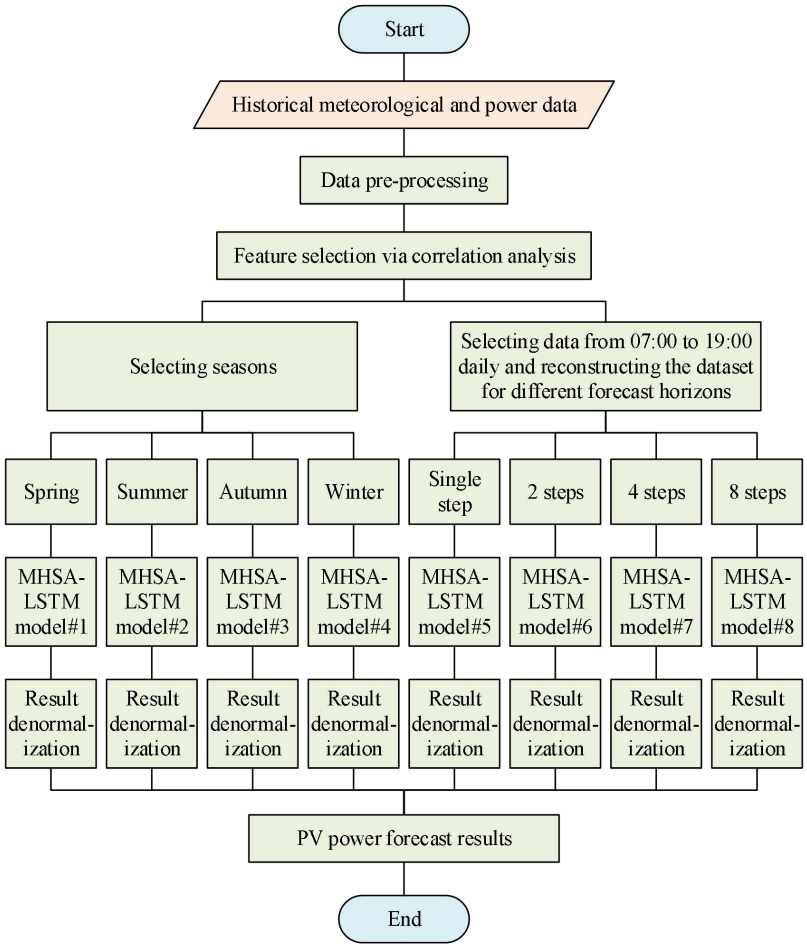

Figure 6. Schematic diagram based on MHSA-LSTM PV power prediction method

5. Ultra-short-term prediction of photovoltaic power generation in multi-climatic environments

Based on the four seasonal subsets that have been delineated, the data from March 1, 2023 to February 29, 2024 were selected as the training set, and the data from March 1, 2024 to February 28, 2025 were selected as the test set, respectively. The data sampling interval of this PV plant is 20 minutes, and a total of 72 time points are collected every day.The hyperparameters of the MHSA-LSTM model are set as follows: the input sequence length is 64 time steps, the batch size is 24, the number of training rounds is 16, and the learning rate is 0.0001; the LSTM module contains 1 hidden layer and the number of hidden units is 256; and the MHSA module is 1 layer with a total of 8 attention heads. The loss function adopts mean square error, and the optimizer chooses Adam.

5.1. Predictive Model Evaluation Indicators

Since the output power of photovoltaic power plants is susceptible to obvious fluctuations due to weather, temperature, solar radiation, atmospheric pressure and other factors, prediction errors are inevitable, so it is necessary to assess the prediction accuracy. The following three indicators are selected for this study:

(1) Mean Absolute Percentage Error (*MAPE*)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y'_i - y_i}{y'_i}\right| \times 100\% \tag{15}$$

In Eq. (15), n is the number of prediction points, and y_i and y'_i denote the predicted value and the actual value, respectively.MAPE is able to intuitively reflect the degree of relative deviation between predicted and actual values.

(2) Root Mean Square Error (*RMSE*)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{16}$$

In Eq. (16), n is the number of predicted points, and y_i and y'_i denote the predicted and actual values, respectively. The root mean square error is a measure used to assess the level of error in a regression model that reflects the degree of dispersion between the predicted and true values.

5.2. Power prediction and result analysis under different seasons

In order to test the robustness of the MHSA-LSTM model in different seasons, the data of four representative months are firstly selected for training and analysis, and then April 12, July 9, October 28, and December 27, 2024 are used as the predicted sample days (all sunny days) in spring, summer, autumn, and winter, respectively, to carry out the power generation prediction based on four methods, namely, BP-LSTM, LSTM, LSTNet, and MHSA-LSTM. LSTM, LSTM, LSTNet and MHSA-LSTM methods. The corresponding prediction curves are shown in Figs. 7-10, and the error statistics are shown in Table 2. The results show that MHSA-LSTM has the highest agreement with the measured power curves under the four seasons, and the prediction of the power peaks remains stable and consistent. By enhancing the weights of key features, MHSA effectively compensates for the shortcomings of LSTM in information extraction, which makes MHSA-LSTM show good robustness under different seasons.

As can be seen from Table 2, the MAPE of the MHSA-LSTM model was reduced by 10.33%, 23.12%, and 5.34%, and the RMSE values were reduced by 2.771kW, 6.296kW, and 1.234kW, respectively, compared to BP-LSTM, LSTM, and LSTNet in springtime. The present model also shows different degrees of performance advantages in summer, fall and winter, although the lead is most obvious in spring. The prediction error of the combined MHSA-LSTM model is always the smallest under all seasons, indicating that the model can stably maintain a high prediction accuracy in PV power prediction in all seasons.

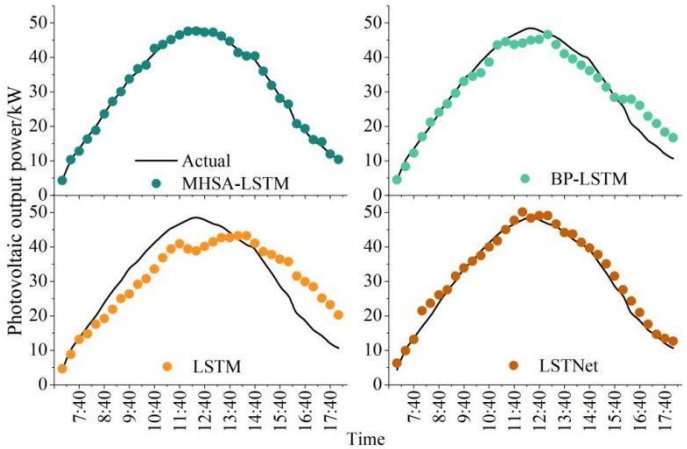

Figure 7. Prediction of PV in spring (2024-04-12)

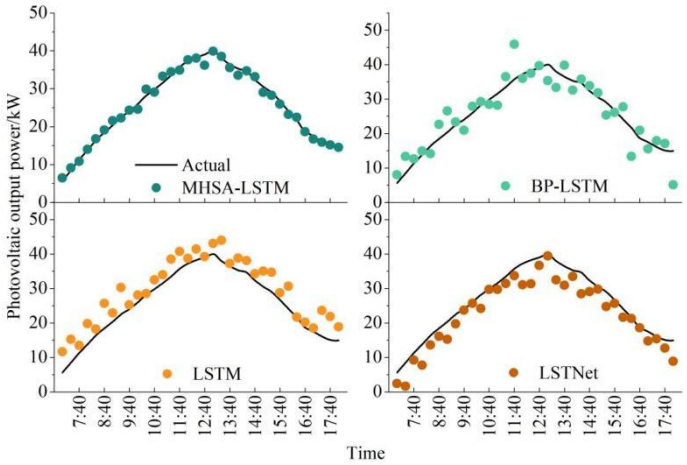

Figure 8. Prediction of PV in summer (2024-07-09)

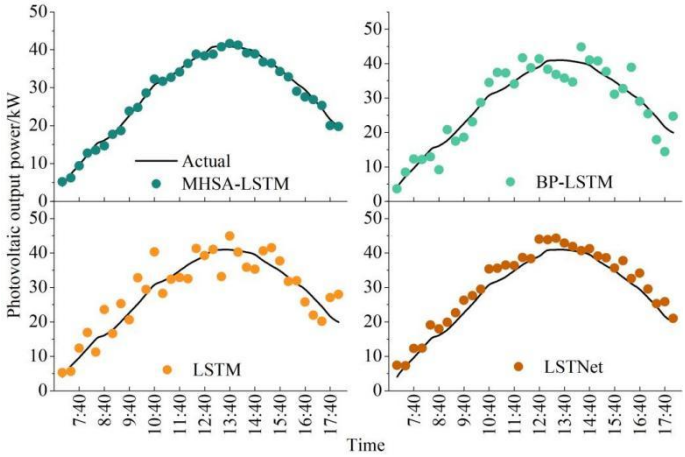

Figure 9. Prediction of PV in autumn (2024-10-28)

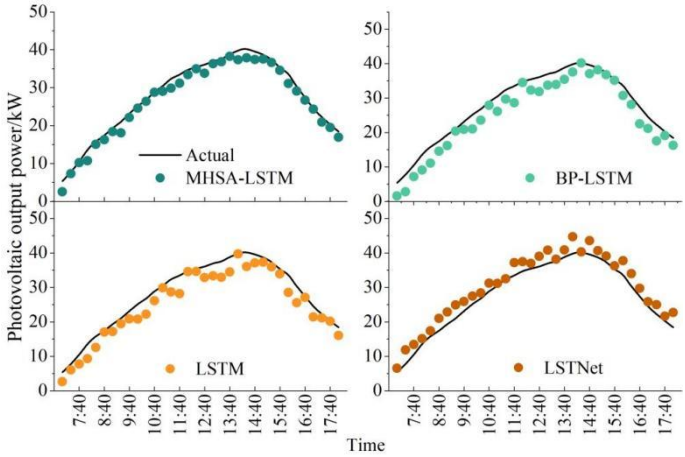

Figure 10. Prediction of PV in winter (2024-12-07)

Table 2. Comparison of Daily Average Error of Four Seasonal Forecast

Season (Date)	Model	RMSE/kW	MAPE/%
Spring (2024-04-12)	MHSA-LSTM	0.620	1.90
	BP-LSTM	3.391	12.23
	LSTM	6.916	25.02
	LSTNet	1.854	7.24
Summer (2024-07-09)	MHSA-LSTM	0.967	3.36
	BP-LSTM	3.901	15.10
	LSTM	4.302	20.01
	LSTNet	3.628	15.06
Autumn (2024-10-28)	MHSA-LSTM	0.996	4.09
	BP-LSTM	3.867	12.87
	LSTM	4.352	15.89
	LSTNet	2.771	11.29
Winter (2024-12-07)	MHSA-LSTM	1.545	6.21
	BP-LSTM	3.178	14.46
	LSTM	2.981	11.38
	LSTNet	2.778	11.43

The MHSA-LSTM model fully integrates the advantages of the two sub-models, and has high accuracy in both global and local feature prediction, and outperforms many traditional single models in both RMSE and MAPE error metrics.

5.3. Power prediction in different weather and analysis of results

To evaluate the prediction feasibility of the MHSA-LSTM model, sunny, rainy and cloudy weather data in mid-August 2023 are selected for training, and the power generation prediction is performed from 07:00 to 18:00 on August 8 (sunny), August 18 (rainy), and August 13 (cloudy), 2024, respectively. The daily average meteorological data for the three weather conditions are shown in Table 3.

Table 3. Predicted Daily Weather Parameters

Date	Solar radiation intensity (W/m²)	Temperature (℃)	Humidity (%)	Atmospheric pressure
2024-08-08	205	29.6	65	1.016
2024-08-18	134	27.8	82	0.992
2024-08-13	180	28.2	72	1.001

The PV power prediction was carried out along the multiple prediction models described in the previous section, and the results are shown in Figures 11-13. It can be seen that the best prediction results are obtained under sunny conditions. Under cloudy weather, clouds intermittently shade the sunlight, resulting in large fluctuations in irradiance during the day, and the error curve shows obvious ups and downs accordingly. During rainy days, the atmospheric conditions are unstable and change frequently, and the power fluctuation is significant in a short period of time, so that the PV power shows fluctuating changes. Compared with other models, MHSA-LSTM is more effective in capturing the trend of power curve changes, and the overall prediction effect is better. The error results of each prediction model under different weather types are shown in Table 4.

Under cloudy or rainy weather, the power output is affected by the large fluctuation of transmittance and solar irradiance, which makes the RMSE and MAPE of the model higher than that of sunny air conditions. Compared with other single models, the MHSA-LSTM with the introduction of the multi-attention mechanism shows better prediction performance under sunny, cloudy and rainy conditions. In particular, the RMSE and MAPE are 1.054 kW and 3.80% for clear sky, for example.

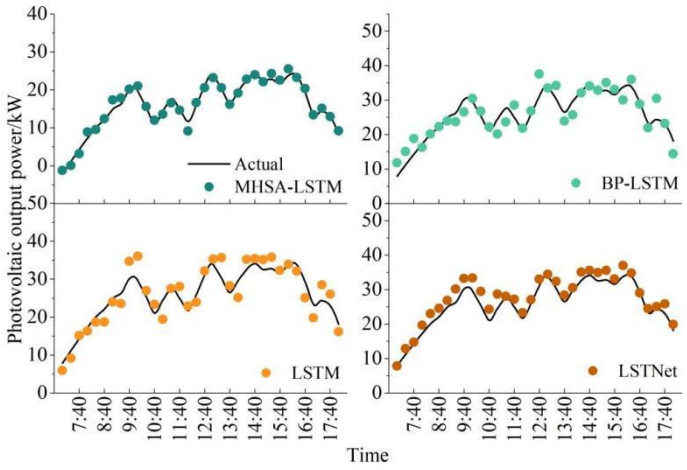

Figure 11. Prediction of PV on August 8 (Sunny)

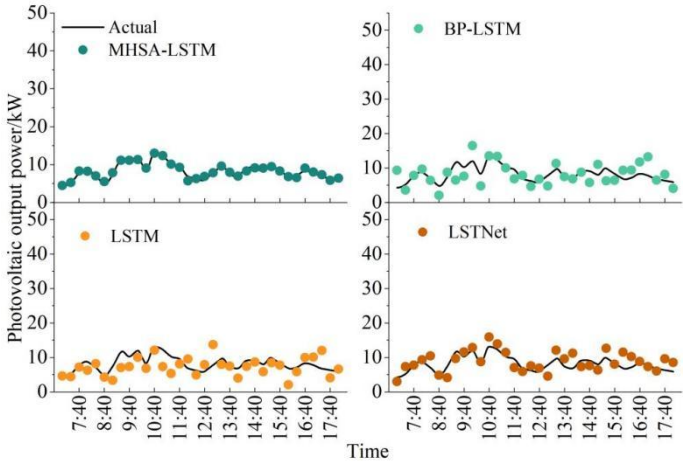

Figure 12. Prediction of PV on August 18 (Rainy)

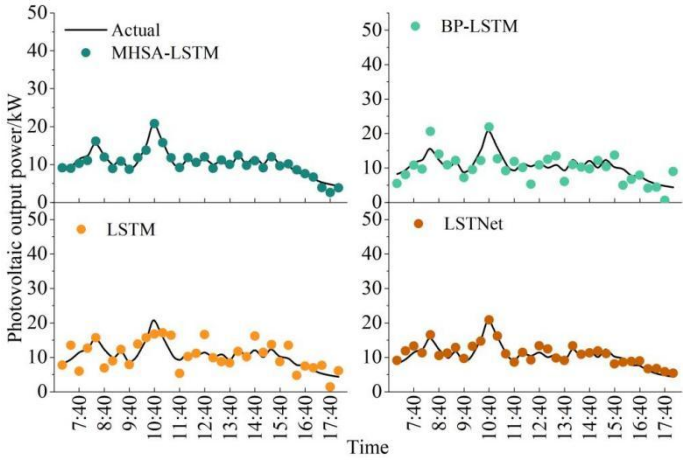

Figure 13. Prediction of PV on August 13 (Cloudy)

Table 4. Comparison of Daily Average Error of Three Different Kinds of Weather

Season (Date)	Model	RMSE/kW	MAPE/%
Sunny (2024-08-08)	MHSA-LSTM	1.054	3.80
	BP-LSTM	2.883	10.51
	LSTM	2.891	10.12
	LSTNet	2.303	8.13
Rainy (2024-08-18)	MHSA-LSTM	1.534	5.63
	BP-LSTM	2.971	27.27
	LSTM	2.984	26.86
	LSTNet	2.698	24.15
Cloudy (2024-08-13)	MHSA-LSTM	1.218	7.58
	BP-LSTM	2.895	24.06
	LSTM	2.902	21.66
	LSTNet	2.435	12.22

6. Conclusion

On the basis of systematically combing the influencing factors of PV power generation system, this paper constructs a MHSA-LSTM model for ultra-short-term PV power prediction. Using the preprocessed experimental data, the model is trained and validated in two dimensions, namely, seasonal category and weather category, and its prediction performance is evaluated in comparison with that of BP-LSTM, LSTM and LSTNet.

The MHSA module can effectively filter the key information in historical weather and power data, and enhance the ability of LSTM to capture remote dependencies in sequence features. Compared with the standard LSTM, the RMSE and MAPE of MHSA-LSTM are reduced by 6.296 kW/23.12%, 3.335 kW/16.65%, 3.356 kW/11.8%, and 1.436 kW/5.17% for the four seasons, respectively, indicating that

the introduction of MHSA in the LSTM framework can improve the prediction performance significantly. Meanwhile, compared with the BP-LSTM model with nonlinear feature mapping, the MHSA-LSTM model also shows better prediction results, and its prediction results are closer to the real value curves under different climatic conditions. Although LSTNet can capture the fluctuation characteristics of periodic time series data, it is prone to error accumulation in multi-step prediction.

The prediction model proposed in this paper has strong versatility and good adaptability to seasonal changes and weather types, and is able to realize high-precision prediction of ultra-short-term PV power under complex and variable climatic conditions, which has the potential to be widely promoted in PV power plants.

References

Abdullah, B. U. D., Khanday, S. A., Islam, N. U., Lata, S., Fatima, H., & Nengroo, S. H. (2024). Comparative analysis using multiple regression models for forecasting photovoltaic power generation. Energies, 17(7), 1564.

Banik, R., & Biswas, A. (2023). Improving solar PV prediction performance with RF-CatBoost ensemble: a robust and complementary approach. Renewable Energy Focus, 46, 207-221.

Huang, C., & Yang, M. (2023). Memory long and short term time series network for ultra-short-term photovoltaic power forecasting. Energy, 279, 127961.

Jagadeesh, V., Venkata Subbaiah, K., & Varanasi, J. (2020). Forecasting the probability of solar power output using logistic regression algorithm. Journal of Statistics and Management Systems, 23(1), 1-16.

Jo, H. H., Kim, J., & Kim, S. (2024). Enhancing the power generation performance of photovoltaic system: Impact of environmental and system factors. Applied Thermal Engineering, 240, 122221.

Khan, Z. A., Hussain, T., & Baik, S. W. (2023). Dual stream network with attention mechanism for photovoltaic power forecasting. Applied Energy, 338, 120916.

Li, L., Lin, J., Wu, N., Xie, S., Meng, C., Zheng, Y., ... & Zhao, Y. (2022). Review and outlook on the international renewable energy development. Energy and Built Environment, 3(2), 139-157.

Liu, M., Wang, X., & Zhong, Z. (2025). Ultra-short-term photovoltaic power prediction based on BiLSTM with wavelet decomposition and dual attention mechanism. Electronics, 14(2), 306.

Mahesh, P. V., Meyyappan, S., & Alla, R. R. (2022). Maximum power point tracking using decision-tree machine-learning algorithm for photovoltaic systems. Clean Energy, 6(5), 762-775.

Niu, D., Wang, K., Sun, L., Wu, J., & Xu, X. (2020). Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study. Applied soft computing, 93, 106389.

Rinesh, S., Deepa, S., Nandan, R. T., Sachin, R. S., Thamil, S. V., Akash, R., ... & Kumar, A. S. (2024). Prediction and classification of solar photovoltaic power generation using extreme gradient boosting regression model. International Journal of Low-Carbon Technologies, 19, 2420-2430.

Saxena, N., Kumar, R., Rao, Y. K., Mondloe, D. S., Dhapekar, N. K., Sharma, A., & Yadav, A. S. (2024). Hybrid KNN-SVM machine learning approach for solar power forecasting. Environmental Challenges, 14, 100838.

Souza, G., Santos, R., & Saraiva, E. (2022). A Log-logistic predictor for power generation in photovoltaic systems. Energies, 15(16), 5973.

Su, Z., Gu, S., Wang, J., & Lund, P. D. (2025). Improving ultra-short-term photovoltaic power forecasting using advanced deep-learning approach. Measurement, 239, 115405.

Tang, H., Kang, F., Li, X., & Sun, Y. (2025). Short-term photovoltaic power prediction model based on feature construction and improved transformer. Energy, 320, 135213.

Tsai, W. C., Tu, C. S., Hong, C. M., & Lin, W. M. (2023). A review of state-of-the-art and short-term forecasting models for solar PV power generation. Energies, 16(14), 5436.

Wang, G., Sun, S., Fan, S., Liu, Y., Cao, S., & Guan, R. (2024). A spatial–temporal data-driven deep learning framework for enhancing ultra-short-term prediction of distributed photovoltaic power generation. International Journal of Electrical Power & Energy Systems, 160, 110125.

Xiong, B., Chen, Y., Zhao, X., Su, Z., Fu, J., Chen, D., & Zhang, D. (2025). Multimodal ultra-short-term probabilistic solar power forecasting with generative AI and transformer. Advances in Applied Energy, 100250.

Yang, M., & Huang, X. (2018). Ultra-short-term prediction of photovoltaic power based on periodic extraction of PV energy and LSH algorithm. IEEE Access, 6, 51200-51205.

Ye, X., Yin, J., Zhang, J., Li, A., Liu, Z., Chen, B., ... & Li, H. (2026). A Multi-Scale CNN-Transformer Network with Residual Correction for Ultra-Short-Term Photovoltaic Power Forecasting. Processes, 14(5), 759.

Yong, B., Zhang, Y., Shen, J., Ren, A., Zhou, X., & Zhou, Q. (2025). ConvODE-Mixer: A multimodal deep learning model for ultra-short-term PV power forecasting. Solar Energy, 300, 113777.

Zazoum, B. (2022). Solar photovoltaic power prediction using different machine learning methods. Energy Reports, 8, 19-25.

Zhai, C., He, X., Cao, Z., Abdou-Tankari, M., Wang, Y., & Zhang, M. (2025). Photovoltaic power forecasting based on VMD-SSA-Transformer: Multidimensional analysis of dataset length, weather mutation and forecast accuracy. Energy, 324, 135971.

Zhang, J., Zhao, Z., Guo, R., Hu, X., Qu, T., Ge, C., & Yan, J. (2026). A Novel Ultra-Short-Term PV Power Forecasting Method Based on a Temporal Attention-Variable Parallel Fusion Encoder Network. Energies, 19(1), 274.

Zhang, R., Wu, Y., Zhang, L., Xu, C., Wang, Z., Zhang, Y., ... & Chen, Q. (2025). A multiscale network with mixed features and extended regional weather forecasts for predicting short-term photovoltaic power. Energy, 318, 134792.

Zhou, D., Liu, Y., Wang, X., Wang, F., & Jia, Y. (2025). Combined ultra-short-term photovoltaic power prediction based on CEEMDAN decomposition and RIME optimized AM-TCN-BiLSTM. Energy, 318, 134847.