

# Photovoltaic power generation prediction based on k-means clustering analysis and GRO-CNN-LSTM Attention

Fangcheng Jin

Northeast Petroleum University, Heilongjiang 163000, China

---

**Abstract:** In order to improve the accuracy of photovoltaic power prediction, this paper proposes a GRO-CNN-LSTM Attention prediction model, which improves the prediction accuracy by introducing the Gold Rushing Optimization Algorithm (GRO) and SE attention mechanism. Secondly, the photovoltaic data is divided into three categories based on weather types using the k-means clustering analysis algorithm: sunny, cloudy, and rainy. Finally, photovoltaic power generation predictions were conducted under three different weather types, and the proposed prediction model was compared and analyzed with other prediction models. The results showed that the CNN-LSTM prediction model, which introduced GRO and SE attention mechanism optimization, had a decrease in MAE, RMSE, and MAPE compared to the original model under three different weather types.

**Keywords:** K-means clustering analysis; Gold rush optimization algorithm; Attention mechanism; CNN-LSTM; Photovoltaic power generation prediction.

---

## 1. Introduction

With the proposal of China's "dual carbon" goals, the development of photovoltaic power generation technology has significant strategic significance [1]. The instability and randomness of photovoltaic power generation pose great challenges to the smooth operation of power systems [2]. The improvement of accuracy in predicting photovoltaic power generation can effectively solve this problem.

There are two main methods for predicting photovoltaic power generation both domestically and internationally: physical modeling methods [3] and deep learning methods [4]. The deep learning method uses neural network models to train, learn, and model a large amount of input data for power prediction. This method can capture complex nonlinear relationships between data and can be adjusted and improved according to actual situations, so it has a very wide range of applications. Reference [5] proposed a prediction model based on MGA-VMD-SNS-Attention GRU, Reference [6] proposed an HBA-Bi-LSTM-KELM prediction model, and optimized it using an improved similar day. Reference [7] used multi-channel input and PCNN-Bi-LSTM model for ultra short term prediction of photovoltaic power generation. Reference [8] used graph convolutional network GCN and gated linear unit GLU to extract spatiotemporal characteristics of power generation and trained the prediction model with historical data of photovoltaic power plants for ultra short term power prediction. Although deep learning models and their combination models used for photovoltaic power generation

prediction have shown good prediction accuracy in recent years, the combination model is generally superior to a single prediction model [9-11]. Therefore, how to improve the optimization algorithm to further enhance the prediction accuracy of the combination model becomes the key to solving the problem of photovoltaic power generation prediction.

This article proposes a GRO-CNN-LSTM Attention prediction model for photovoltaic power generation. After introducing the gold mining optimization algorithm and channel attention mechanism, CNN-LSTM has faster convergence speed and better optimization performance. In order to verify the optimization ability of the improved algorithm, experimental comparative analysis was conducted with other models under three different weather types. The effectiveness and superiority of the GRO-CNN-LSTM Attention model proposed in this paper were demonstrated by comparing the predicted data.

## 2. Construction of predictive models

### 2.1. Gold Rushing Optimization Algorithm

Gold Rush Optimizer, GRO is the latest intelligent optimization algorithm proposed in 23 years, The inspiration for GRO comes from the gold mining behavior of gold miners, simulating three key behaviors during their gold mining period: migration, gold mining, and collaboration.

#### 1) Set initial population

In the initial stage of gold exploration, a random batch of gold prospectors is first generated, and the location data of

these prospectors is stored in the matrix  $M_G$ . The expression of the matrix  $M_{GP}$  is shown in equation (1):

$$M_{GP} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} \end{bmatrix} \quad (1)$$

In the formula:  $x_{ij}$  represents the position of gold miner  $i$  in the  $j$  dimension;  $d$  is the size of the dimension;  $n$  is the number of gold miners.

In the process of updating the location, it is necessary to optimize and evaluate the position of the gold miners. Therefore, an objective function needs to be used to evaluate the gold miners, and the evaluation results should be stored numerically in the matrix  $M_F$ . The expression of the matrix  $M_F$  is shown in equation (2):

$$M_F = \begin{bmatrix} f(x_{11} & x_{12} & \dots & x_{1d}) \\ f(x_{21} & x_{22} & \dots & x_{2d}) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_{n1} & x_{n2} & \dots & x_{nd}) \end{bmatrix} \quad (2)$$

In the formula:  $f()$  is the evaluation function.

## 2) Migration of gold miners

After discovering the gold mine, the gold miners began to migrate towards the location of the gold mine. The location of the richest gold mine is the optimal solution in the algorithm search space. Due to its exact location being unknown, the location of the richest gold deposit is estimated based on the optimal location of the gold digger. The formula for updating the position of gold miners is shown in equations (3) and (4):

$$\vec{D}_1 = \vec{C}_1 \cdot \vec{X}^*(t) - \vec{X}_i(t) \quad (3)$$

$$\overline{X}_{new_i}(t+1) = \vec{X}_i(t) + \vec{A}_1 \cdot \vec{D}_1 \quad (4)$$

In the formula:  $\vec{X}^*$  represents the location of the richest gold deposit;  $\vec{X}_i$  For the location of gold digger  $i$ ;  $t$  is the current number of iterations;  $\overline{X}_{new_i}$  is the latest location for gold digger  $i$ ;  $\vec{D}_1$  Is the transfer vector coefficient;  $\vec{A}_1$ 、 $\vec{C}_1$  Is the vector coefficient.  $\vec{A}_1$  and  $\vec{C}_1$  expression of is shown in equations (5) and (6):

$$\vec{A}_1 = 1 + l_1 \left( \vec{r}_1 - \frac{1}{2} \right) \quad (5)$$

$$\vec{C}_1 = 2\vec{r}_2 \quad (6)$$

In the formula:  $\vec{r}_1$ 、 $\vec{r}_2$  is a random vector with a value range of  $[0,1]$ ;  $l_1$  is the convergence component, and its expression is shown in equation (7):

$$l_e = \left( \frac{\max_{iter} - iter}{\max_{iter} - 1} \right)^e \left( 2 - \frac{1}{\max_{iter}} \right) + \frac{1}{\max_{iter}} \quad (7)$$

In the formula:  $\max_{iter}$  is the maximum number of iterations. When the value of  $e$  is equal to 1,  $l_e$  decreases linearly from 2 to  $\frac{1}{\max_{iter}}$ ; When the value of  $e$  is greater than 1, Nonlinear reduction of  $l_e$ .

## 3) Gold exploration

After finding the gold mine, the gold prospector will mine it, and at this time, the location of the gold mine will be considered as the location of the gold mine. The expression for updating the position of gold miners during gold mining is shown in equations (8) and (9):

$$\vec{D}_2 = \vec{X}_i(t) - \vec{X}_r(t) \quad (8)$$

$$\overline{X}_{new_i}(t+1) = \vec{X}_i(t) + \vec{A}_2 \cdot \vec{D}_2 \quad (9)$$

In the formula:  $\vec{X}_i$  is the position of the random gold miner  $r$ ;  $\vec{D}_2$  is the gold mining vector coefficient;  $\vec{A}_2$  is the vector coefficient. The expression of  $\vec{A}_2$  is shown in equation (10):

$$\vec{A}_2 = 2l_2\vec{r}_1 - l_2 \quad (10)$$

The use of  $l_2$  instead of  $l_1$  here is to improve the optimization ability of the algorithm in the gold mining stage.

## 4) Gold Rusher Collaboration

Sometimes, gold miners collaborate with each other. When three random prospectors  $i$ 、 $g_1$  and  $g_2$  cooperate, the expression is shown as (11) and (12):

$$\vec{D}_3 = \vec{X}_{g_2}(t) - \vec{X}_{g_1}(t) \quad (11)$$

$$\overline{X}_{new_i}(t+1) = \vec{X}_i(t) + \vec{r}_1 \cdot \vec{D}_3 \quad (12)$$

In the formula:  $\vec{X}_{g_1}$  and  $\vec{X}_{g_2}$  is the location of gold diggers  $g_1$  and  $g_2$ ;  $\vec{D}_3$  is the cooperation vector coefficient.

## 5) Gold digger location update

In order to get the most gold, gold miners compare the value of different mines and choose whether to go to other mines. This process evaluates the value of the gold deposit through the objective function to determine whether the prospector will update the location. The expression of the position update mechanism for gold diggers is shown in equation (13):

$$\begin{aligned} \vec{X}_i(t+1) &= \overline{X}_{new_i}(t+1) \\ \text{if } f(\overline{X}_{new_i}(t+1)) &< f(\vec{X}_i(t)) \end{aligned} \quad (13)$$

## 2.2. SE attention mechanism

SE attention mechanism is a feature channel-based attention

mechanism, which is used to enhance the network's attention to different feature channels, thereby improving the performance of the network. Specifically, the SE attention mechanism mainly consists of two key steps: Squeeze and Excitation.

#### 1) Transformation

After transformation operation, the input feature graph  $X$  is changed into the output feature graph  $U$ .  $F_{tr}$  can be regarded as a standard convolution operator. The formula is defined as follows:

$$U_c = V_c * X = \sum_{s=1}^C V_c^s * X^s \quad (14)$$

In the formula:  $X$  is the output feature map;  $U$  is the output feature map;  $V$  is a set of filter nuclei;  $V_c$  is the parameter of the  $c$  filter;  $V_c^s$  is a 2D space core;  $*$  is a convolution operation.

#### 2) Squeeze

The feature graph of  $W \times H \times C$  containing global information is compressed into a  $1 \times 1 \times C$  feature vector  $z$  by global averaging pooling for each channel. The feature of each channel is compressed into a scalar value, which represents the importance of the feature on the channel. The definition formula is as follows:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (15)$$

In the formula:  $z_c$  is the  $c$ -th element of the eigenvector  $Z$ ;  $F_{sq}()$  is the Squeeze operation;  $u_c$  is the output feature diagram of channel  $c$ .

#### 3) Excitation

The feature vector  $Z$  obtained in the previous step is processed by two fully connected layers  $W_1$  and  $W_2$ , and the channel weight value  $s$  is obtained. In order to make the connection layer use two nonlinear functions ReLU and Sigmoid at the same time, the parameter  $r$  is introduced to reduce the number of channels. The specific calculation method is as follows:

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (16)$$

In the formula:  $s$  is the channel attention weight vector;  $F_{ex}()$  Excitation is the Excitation operation;  $z$  is the eigenvector of  $1 \times 1 \times C$ ;  $W_1$  and  $W_2$  are two fully connected layers;  $\delta()$  is the activation function ReLU;  $\sigma()$  is the activation function Sigmoid.

#### 4) Scale

After the Squeeze and Excitation operations, we also need to merge the results with the output feature graph  $U$ , which weights the attention weight obtained earlier to the features of each channel. The method of incorporation is as follows:

$$\hat{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (17)$$

In the formula:  $\hat{x}_c$  is the feature map weighted by attention weight;  $F_{scale}()$  is the Scale operation;  $u_c$  is the  $c$  element of the  $c$  vector  $s$  in the attention weight vector  $s$ ;  $s_c$  is the  $C$ -th element in the output feature map

To sum up, the principle flow chart of SE attention mechanism is shown in Figure 1:

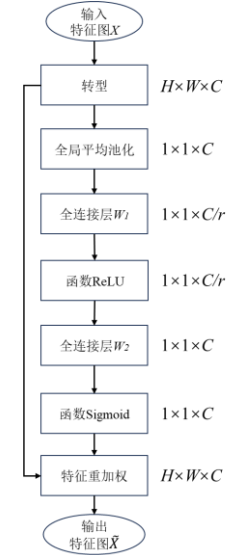


Figure 1. SE Attention Mechanism Flowchart

### 2.3. Construction of GRO-CNN-LSTM-Attention prediction model

#### 1) CNN

With the advantage of convolutional operation, convolutional neural network can express the original data at a higher level and in a more abstract way. In this paper, the convolutional neural network is mainly used to extract features from photovoltaic power data. By using two convolutional layers, each of which has a different number of convolutional cores and filter sizes, these convolution cores can learn and capture local features in the input sequence, and extract these features through convolution operations. Then, the extracted features are mapped and abstracted nonlinearly by the activation functions ReLU and Sigmoid of the convolutional layer, so that the model can learn more advanced abstract feature representations. The output of convolutional layer feature extraction is:

$$Y = \sigma(W \cdot X + b) \quad (18)$$

In the formula:  $Y$  is the extracted feature;  $\sigma()$  is the activation function Sigmoid;  $W$  is the weight matrix;  $X$  is the time series;  $b$  is the bias vector.

#### 2) LSTM

The long short-term memory network has memory and forgetting mechanisms, so it is better able to process time series

data to extract useful long - and short-term features from the input sequence. The LSTM layer is responsible for timing modeling the features from the CNN portion. It can capture the time dependencies in the utilization input sequence and make predictions based on them. The LSTM layer is mainly divided into the forgetting gate, the input gate and the output gate, and its expression is as follows:

$$z^f = \sigma(W^f \otimes [x^t, h^{t-1}]) \quad (19)$$

$$z^i = \sigma(W^i \otimes [x^t, h^{t-1}]) \quad (20)$$

$$z^o = \sigma(W^o \otimes [x^t, h^{t-1}]) \quad (21)$$

In the formula: is the multiplication of elements in the operation matrix;  $W^f$ 、 $W^i$ 、 $W^o$  are the weight coefficients of the forgetting door, the input door and the output door respectively;  $x^t$  is the input of the current neuron;  $h^{t-1}$  is the hidden state of the previous neuron;  $z$  is the value of the result converted by activation function.

### 3)GRO-CNN-LSTM-Attention

The GRO-CNN-LSTM-attention prediction model consists of four parts: data preprocessing, parameter setting of GRO optimization algorithm, CNN unit based on SE Attention mechanism, and LSTM unit. The details are described as follows:

Step.1 Import data set, set the ratio of training set to test set according to 4:1 ratio, perform data normalization, data tiling and data format conversion.

Step.2 Call GRO optimization algorithm, set parameters of optimization algorithm, set the number of optimization algorithms, the maximum number of iterations, the number of optimization parameters and the upper and lower bounds of parameter values.

Step.3 The input layer is established, the input data structure is [f,1,1], and the sequence input and sequence folding are carried out to transform the sequence data into the form suitable for CNN processing. Where, f is the number of input features in the dataset, and f=5 when predicting the power interval of photovoltaic power generation.

Step.4 The convolution layer and activation layer are established, and the model extracts features from the input sequence. In this paper, two convolution cores are used, each of which learns different feature representations, thus facilitating the extraction of multi-scale feature information.

Step.5 The global average pooling layer can obtain the importance information of each channel by reducing the dimension of the output of the convolution layer. The SE attention mechanism is then used to dynamically adjust the

weight of each channel.

Step.6 Sigmoid activation function and multiplication layer were used to achieve the feature weighting of different channels, and the weighted results were reduced to a form suitable for LSTM processing by sequence expansion and flattening, and then input into LSTM layer.

Step.7 In LSTM layer, autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to analyze the correlation between the historical data and the current value, so as to effectively process the time series data and capture the time correlation in the series.

Step.8 Finally, the fully connected layer and the regression layer make the final prediction according to the output of the LSTM layer, and analyze the error of the prediction results.

To sum up, the GRO-CNN-LSTM-Attention PV power prediction model architecture is shown in Figure 2:

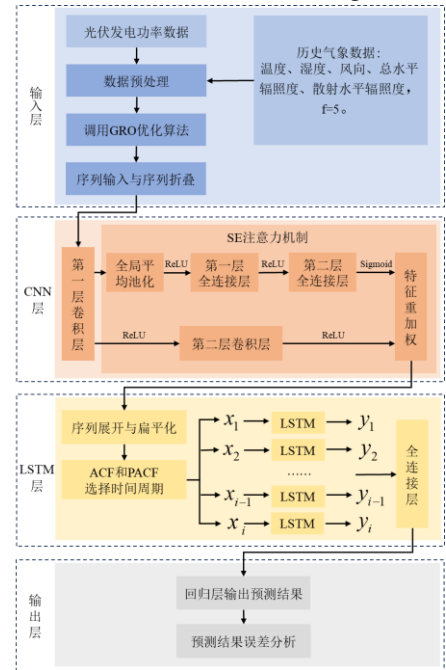


Figure 2. Architecture of photovoltaic power generation power prediction model

## 3. Selection of cluster analysis method

### 3.1. Commonly used cluster analysis method

At present, the commonly used cluster analysis methods at home and abroad include K-means algorithm, hierarchical clustering algorithm, DBSCAN algorithm, spectral clustering algorithm, Gaussian mixture model, etc. These algorithms have their own characteristics and are suitable for different data types and clustering requirements. In practical application, it is very important to select the appropriate clustering algorithm according to the data characteristics and task objectives.

K-means algorithm is an intuitive and easy to understand clustering algorithm. The basic idea is to divide the data points into K clusters so that the similarity of the data points within

the clusters is the highest and the similarity between the clusters is the lowest. At the same time, the calculation efficiency of K-means algorithm is higher, especially for large-scale data sets. Because it is updated iteratively, the calculation speed is faster. This algorithm mainly relies on calculating the distance between data points. At the same time, by using approximation algorithm or sampling method, K-means algorithm can handle data with high dimensional characteristics. Therefore, it performs well when dealing with large data such as photovoltaic power generation.

### 3.2. k-means cluster analysis

The calculation steps of k-means clustering analysis algorithm are as follows:

Step.1 First you need to determine the number of clusters K, that is, how many clusters you want to divide the data into.

Step.2 K data points are randomly selected as the initial clustering center. These initial clustering centers can be selected randomly from the data set or through other heuristic methods.

Step.3 For each data point, the distance between it and all cluster centers is calculated and assigned to the nearest cluster center. Commonly used distance measurements include Euclidean distance, Manhattan distance, etc.

Step.4 For each cluster, the mean of all data points in that cluster is calculated and that mean is used as the new cluster center.

Step.5 Repeat steps 3 and 4 until certain stopping criteria are met, such as the cluster center is stable without significant change, or the maximum number of iterations is reached.

Step.6 You end up with K clusters, and each data point is assigned to one of them. The clustering results can be further analyzed and interpreted as needed.

The schematic diagram of its principle is as follows:

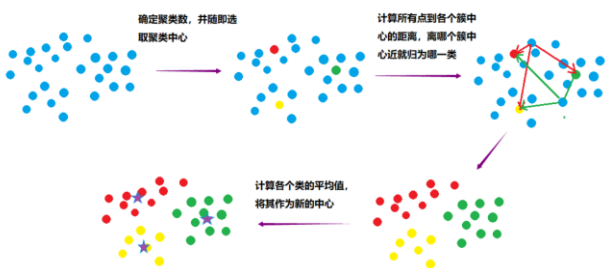


Figure 3. k-means schematic

A total of 365 sets of photovoltaic power generation data of a photovoltaic power station in Australia were selected for one year and divided into three categories according to weather type: sunny day, cloudy day and rainy day by K-means clustering analysis algorithm. After removing bad data and missing data, the number of clusters K=3 was set to obtain cluster analysis

results. The clustering analysis results are shown in the figure below:

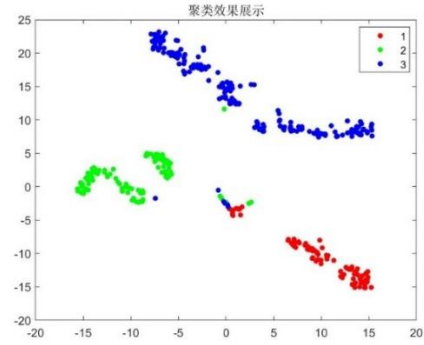


Figure 4. Cluster analysis results

## 4. Analysis of numerical examples

### 4.1. Data Processing

In Australia, the solar energy center DKSCAD photovoltaic power station measured data as an example. Based on the Pearson coefficient between meteorological factors and photovoltaic power generation power, five sets of meteorological data with high correlation coefficients of temperature, humidity, wind direction, total horizontal irradiance and scattering level irradiance were selected for power prediction. Due to the high quality of the selected data, after eliminating part of the incomplete data, the data of the photovoltaic power generation system output period 6:00 to 18:00 is selected, the sampling interval is 5 minutes, a total of 145 sampling points per day, a total of 300 days of data are selected. Finally, the data are normalized according to the ratio of training set to test set 4:1, and these data are used to predict photovoltaic power generation.

### 4.2. Evaluation index

In the comparative analysis of prediction errors, this paper uses mean absolute error (MAE), mean square root error (RMSE) and mean absolute error rate (MAPE) to compare the advantages and disadvantages of the algorithm models. The calculation formula is as follows:

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i| \quad (22)$$

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (23)$$

$$e_{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\tilde{y}_i - y_i}{y_i} \right| \quad (24)$$

In the formula:  $n$  is the number of predicted results;  $y_i$  is the actual value;  $\tilde{y}_i$  is the predicted value.

### 4.3. Comparative analysis of prediction results

Through the k-means clustering analysis of the photovoltaic power data of photovoltaic power stations and the corresponding meteorological data, the data can be roughly divided into three categories according to the weather types: sunny, cloudy and rainy days. This data processing method helps to better show the influence of different weather types on photovoltaic power generation, and provides a basis for the establishment of models and comparative experiments. By examining the prediction results of various models under

different weather types, the effectiveness and performance advantages of the proposed improved algorithm under different weather conditions can be more comprehensively evaluated. Further, the prediction model proposed in this paper was compared and analyzed with the LSTM, CNN-LSTM and CNN-LSTM-ATTENTION prediction models. In the prediction results, three weather types of sunny day, cloudy day and rainy day were selected respectively for visualization to obtain the comparison of photovoltaic power point prediction results, as shown in Figure 5.

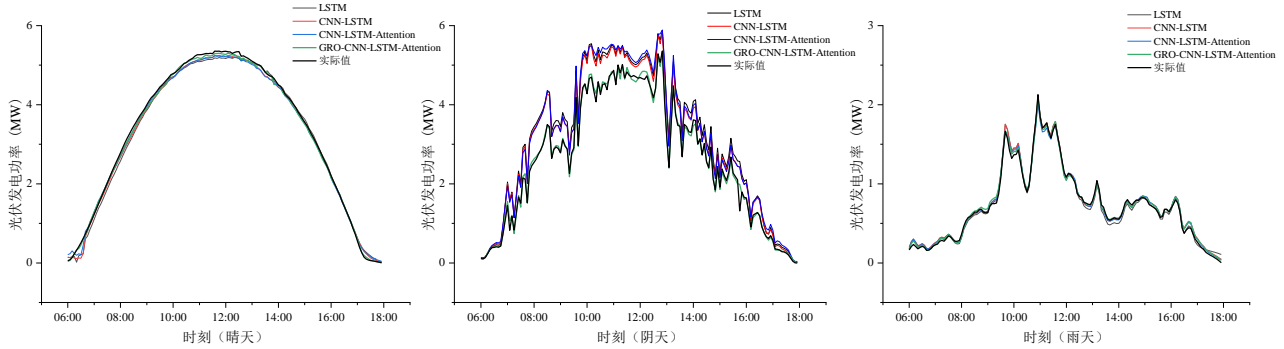


Figure 5. Comparison chart of point prediction results

According to Figure 5, we find that the selected PV power data set has high data quality in sunny and rainy days, and good prediction results have been achieved. However, because the sampling interval of this dataset is 5 minutes, the PV power data fluctuates greatly when the weather type is cloudy, which makes it difficult for the point prediction model to extract data features, thus causing fluctuations in the prediction results and affecting the prediction accuracy.

Specific analysis shows that when the weather type is sunny, although the results of each forecasting model are generally better, there are large errors in the two periods of 6:00-7:00 and 17:00-18:00, except for the point forecasting model proposed in this paper. In other words, at the beginning and end of the prediction, the results of the other models are not accurate enough. Secondly, when the weather type is rainy, on the whole, the prediction models do not show obvious differences, except in the middle part of the forecast, the prediction results of other

prediction models are slightly up and down relative to the true value. Finally, when the weather type is cloudy, other models fail to stand the test due to the sharp fluctuation of the collected data, and the predicted results are significantly deviated from the real values. However, the point prediction model proposed in this paper is almost unaffected, showing significant performance advantages.

In summary, by bolding the prediction results of the GRO-CNN-LSTM-Attention point prediction model proposed in this paper and the visual line of the actual value in the figure, it can be clearly seen that the two lines almost coincide, which further verifies the validity of the prediction model proposed in this paper.

In order to more intuitively demonstrate the advantages of the forecast model proposed in this paper, the prediction error results of each forecast model under three different weather types were visualized, as shown in Figure 6.

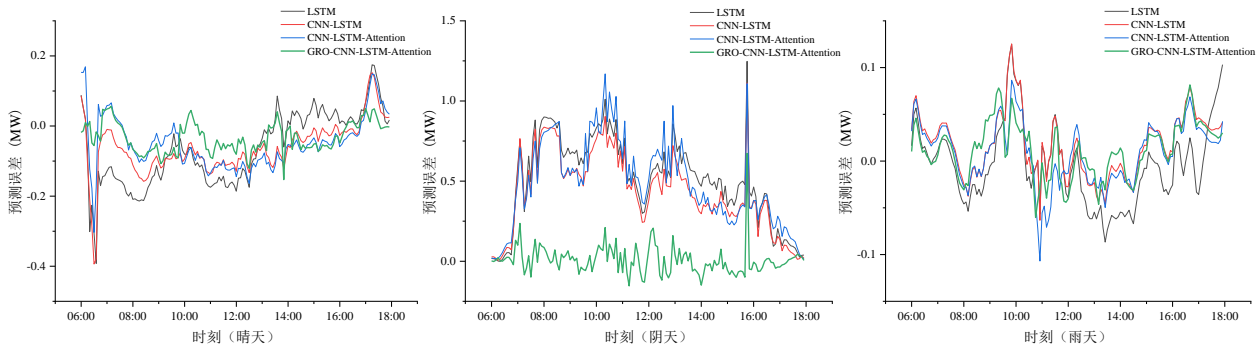


Figure 6. Comparison chart of point prediction error results

As can be seen from Figure 6, no matter what kind of weather type, each forecasting model has a large error fluctuation, and

even a sudden change occurs at very few forecasting points. Compared with other models, the GRO-CNN-LSTM-Attention model proposed in this paper exhibits smaller errors and error fluctuations. Although this advantage may not be obvious under sunny and rainy weather types, the error results of the other three prediction models fluctuate around 0.5MW on cloudy days, while the error of the prediction model proposed in this

paper is controlled around 0MW, achieving good prediction results. Meanwhile, the mean absolute error (MAE), mean square root error (RMSE) and mean absolute error rate (MAPE) of the forecast results of each forecasting model under three different weather types were calculated in this paper. The specific data are listed in Table 1.

**Table 1.** Comparison of prediction errors of different models under different weather types

Weather	Model	MAE/MW	RMSE/MW	MAPE/%
Sunny day	LSTM	0.11	0.13	2.22
	CNN-LSTM	0.08	0.10	1.28
	CNN-LSTM-Attention	0.05	0.07	1.88
	GRO-CNN-LSTM-Attention	<b>0.03</b>	<b>0.06</b>	<b>1.05</b>
Overcast day	LSTM	0.53	0.58	9.47
	CNN-LSTM	0.49	0.53	5.84
	CNN-LSTM-Attention	0.45	0.54	4.28
	GRO-CNN-LSTM-Attention	<b>0.25</b>	<b>0.19</b>	<b>2.89</b>
Rainy day	LSTM	0.07	0.06	7.64
	CNN-LSTM	0.05	0.05	3.71
	CNN-LSTM-Attention	0.03	0.03	3.44
	GRO-CNN-LSTM-Attention	<b>0.02</b>	<b>0.03</b>	<b>1.85</b>

Specific data are shown in Table 1. Compared with the original model, the CNN-LSTM prediction model optimized with GRO and SE attention mechanisms showed a decrease in MAE, RMSE and MAPE of the forecast results under three different weather types. Among them, MAE, RMSE and MAPE decreased by 0.05MW, 0.04MW, 0.23%, 0.24MW, 0.34MW, 2.95% and 0.03MW, 0.02MW, 1.86% year-on-year. In summary, the above data strongly verify the effectiveness, reliability and stability of the point prediction model proposed in this paper, and it has better performance advantages than other prediction models.

## 5. Conclusion

With the proposed "double carbon" goal, the demand for the accuracy of photovoltaic power prediction is getting higher and higher. To solve this problem, a prediction method based on improved kernel density estimation and GRO-CNN-LSTM-Attention is proposed. The CNN-LSTM model with GRO and SE attention mechanism optimization showed a decrease in MAE, RMSE and MAPE in three different weather types compared with the original model. The validity, reliability and stability of the prediction model proposed in this paper are effectively verified, and it has better performance advantages than other prediction models.

## References

- [1] LI M D, LI J W. The Development Status and Prospects of Distributed Photovoltaic Power Generation in China under the "Dual Carbon" Goal[J]. Solar energy, 2023(05):5-10.
- [2] WANG Y. Analysis of the impact of distributed photovoltaic grid connection on the distribution network[J]. Lamps and lighting, 2023(08):135-137.
- [3] DAI J, XIAO W P, HU F Y et al. Research on Physical Prediction Models for Photovoltaic Power Generation Performance[J]. Chinese journal of power sources, 2018, 42(02):262-266.
- [4] Mellit A, Pavan A M, Lughi V. Deep learning neural networks for short-term photovoltaic power forecasting[J]. Renewable Energy, 2021, 172: 276-288.
- [5] Li H, Gao B P. SHORT-TERM PV POWER FORECASTING BASED ON IMPROVED VMD AND SNS-ATTENTION-GRU[J]. Acta energiae solaris sinica, 2023, 44(8): 292-300.
- [6] LI C R, PAN C P, YANG W R et al. Photovoltaic power generation power prediction based on improved similar day optimization of HBA-BiLSTM-KELM[J/OL]. Acta energiae solaris sinica, 1-9[2023-10-23].
- [7] BI G H, ZHAO X, CHEN C P. Ultrashort-term prediction of photovoltaic power generation based on multi-channel input and PCNN BiLSTM[J]. Power system technology, 2022, 46(09): 3463-3476.
- [8] LI L G, SUN L Q, DAI Y S et al. Ultrashort-term prediction method for photovoltaic power generation based on spatiotemporal graph convolutional neural network[J]. Journal of engineering for thermal energy and power, 2023, 38(09):152-157+173.
- [9] JIA L Y, YUN S N, ZHAO Z N et al. Research progress in the application of neural networks for short-term photovoltaic power

- generation prediction[J]. *Acta energiae solaris sinica*, 2022,43(12):88-97.
- [10] Xue J, Shen B. Dung beetle optimizer: A new meta-heuristic algorithm for global optimization[J]. *The Journal of Supercomputing*, 2023, 79(7): 7305-7336.
- [11] Dong S, Wang Y, \*\* for high-precision ADC calibration[J]. *Microelectronics Journal*, 2021, 116: 105259.
- [12] CHAI Y, ZHU Y, RENG S. An Improved Whale Optimization Algorithm Based on Multi strategy Collaboration[J]. *Computer science engineering*,2023,45(07):1308-1319.
- [13] WANG Y W, WANG W L,YANG Y G et al. An Improved Marine Predator Algorithm Based on Multi strategy Fusion and Its Engineering Application[J/OL]. *Computer integrated manufacturing system*:1-21[2023-10-23].
- [14] GAO J L, LI H, DENG M.Short term power prediction of wind farm based on GAVMD-SGRU model[J]. *Journal of Jilin University(information science edition)*, 2021, 39(6): 647-655.
- [15] XIANG L, LI J X, WANG P H, et al.Wind speed multistep interval forecasting based on VMD-FIG and parameter-optimized GRU[J]. *Acta energiae solaris sinica*, 2021, 42(10): 237-242.
- [16] LIU B L, ZHANG H L, WANG C, et al.Ultra-short-term wind speed prediction based on sequence-to-sequence and attention mechanism[J]. *Acta energiae solaris sinica*, 2021, 42(9): 286-294.
- [17] Mellit A. An Overview on the Application of Machine Learning and Deep Learning for Photovoltaic Output Power Forecasting[C]//*Proceedings of the 2nd International Conference on Electronic Engineering and Renewable Energy Systems: ICEERE 2020*, 13-15 April 2020, Saidia, Morocco. Springer Singapore, 2021: 55-68.
- [18] Wang K, Qi X, Liu H. Photovoltaic power forecasting based LSTM-Convolutional Network[J]. *Energy*, 2019, 189: 116225.
- [19] Sharadga H, Hajimirza S, Balog R S. Time series forecasting of solar power generation for large-scale photovoltaic plants[J]. *Renewable Energy*, 2020, 150: 797-807.
- [20] Zhou H, Zhang Y, Yang L, et al. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism[J]. *Ieee Access*, 2019, 7: 78063-78074.