

# A TPA-TCN Prediction Model Applied In Photovoltaic Power Generation Field

Jianxia Wang<sup>1</sup>, Shangyue Wang\*, Xi Meng<sup>†</sup>, Jiaqing Cao<sup>2</sup>, Junjie Wang<sup>1</sup>, and Zhiguo Liu<sup>2\*</sup>

<sup>1</sup>Hebei University of Science and Technology, China

<sup>2</sup>Shijiazhuang University, China

**Abstract:** To solve the problem of large fluctuation and instability of photovoltaic power generation, a deep learning prediction model (TPA-TCN) based on temporal pattern attention mechanism (TPA) and temporal convolutional network (TCN) is proposed, and then applied to photovoltaic power generation. First of all, the k-means clustering algorithm is used to cluster historical data to obtain three typical weather types, and the model is trained by dividing test sets according to the clustering results. After TPA is introduced into the TCN model, which can capture the influence of each variable on the predicted series of the model, help the model pay better attention to the key features in the time series, improve the model's ability to understand the data, and thus efficiently and accurately predict the short-term photovoltaic power. Combined with the measured data, the experiment results show that the TPA-TCN model has good generalization ability and high precision in different weather types.

## 1 Introduction

With the increasing shortage of traditional fossil energy and the increasingly serious problem of environmental pollution, clean renewable energy, especially solar energy [1], has been developed and utilized on a large scale [2]. However, the volatility of photovoltaic power generation has brought many problems to the security of the grid, such as: too little photovoltaic power generation will lead to too little power transmission in the grid transmission process; The phenomenon of light abandonment occurs in photovoltaic power stations, resulting in a large amount of waste of solar energy resources. Accurate photovoltaic power prediction is conducive to the reasonable dispatch of photovoltaic power stations and power grids, increase the power grid's absorption capacity, and improve the stability of power grids [3].

Accurate prediction of photovoltaic power output is a current research hotspot. The development of deep learning is very rapid, and many scholars have made great contributions in the field of deep learning. Many scholars make use of the classic Recurrent Neural Networks (RNN) [4,5], Long Short-Term Memory (LSTM) [6-8] and Gated Recurrent Unit (GRU) [9,10] have achieved a lot of results, which can effectively capture the dependence of time series and achieve good results in short-term PV output prediction. Babbar et al. combined the Adaboost algorithm with the LSTM model to classify and predict PV power under different weather types with greater accuracy than a single model.[11] Hu et al. proposed a comprehensive similarity index that combines Euclidean distance and

grey correlation to select similar days[12]. Limouni et al. combined the LSTM model with the TCN model to form the LSTM-TCN model, and LSTM extracted time features from the input data to further improve the accuracy of the TCN model[13]. Li et al. used MOSMA for feature selection to improve the accuracy of the TCN model[14].

Although the above literature improves the accuracy of the model, some of them do not take into account the generalization ability of the model. In order to better model long-distance dependencies and improve the generalization ability of TCN model, this paper applies the TPA [15] attention mechanism to TCN models to make the networks better capture long-distance dependencies in time series and dynamically adjust the weights according to the importance of different time steps. The main contributions of this study are as follows:

- The k-means clustering algorithm is used to cluster the historical data of photovoltaic power generation, and the historical data of three typical weather types, sunny, cloudy and rainy, are obtained respectively. Three different types of weather are used to detect the performance of the model.
- A deep learning predictive model TPA-TCN based on temporal pattern attention mechanism and temporal convolutional network is proposed. By combining TPA with TCN model, the model can better focus on important time steps to better capture long-term time dependence.
- Compared with RNN, LSTM and TCN models, the MSE value and MAE value of TPA-TCN model have different degrees of reduction under sunny, cloudy and rainy conditions, respectively, indicating that the TPA-

\* Corresponding author: [sjzlg@163.com](mailto:sjzlg@163.com)

TCN model has higher accuracy and generalization ability.

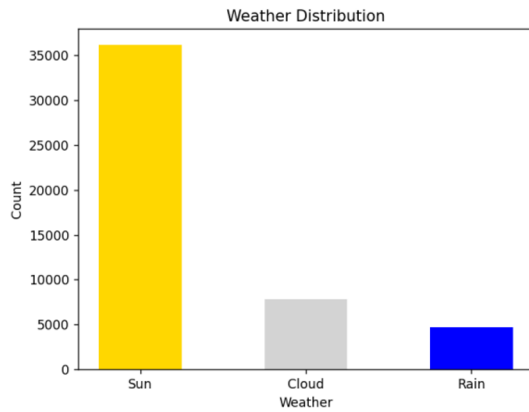


Fig. 1. Classification of weather types.

## 2 Pv power prediction model based on TPA-TCN

The data set used in this paper is the public data of DKASC's Alice Hot spring photovoltaic system in Australia in 2017-2018. Because photovoltaics do not generate electricity at night, the data used is between 7 a.m. and 6 p.m.. The sampling period is 5 min, and a total of 48,678 data sets are used. Through the classification of k-means algorithm, the classification of weather types as shown in Figure 1 is obtained, in which sunny weather accounts for the most, with 36,176 data. 7851 data on cloudy days; There are 4651 data points for rainy weather.

Time series prediction involves a variety of neural network models, which differ in internal structure, training methods and hyperparameter setting. These structural differences directly affect the fit degree, generalization ability, computational efficiency and error level of the model. Even when predictions are made on the same data set, the results of different models can vary significantly. Therefore, when selecting the appropriate prediction model, it is necessary to consider the requirements of different scenarios, data characteristics and specific prediction requirements.

### 2.1 Data preprocessing

Historical PV data and meteorological data contain a wealth of feature information, directly input the original data into the model may not be able to fully excavate the correlation between features, resulting in the model difficult to learn and affect the prediction effect. Therefore, before training and testing the model, it is necessary to pre-process historical PV and meteorological data to enhance the applicability and trainability of the data.

First of all, the data set of DKASC Australian Alice Hot spring photovoltaic system is processed for missing outliers. Due to the particularity of time series data, the data containing missing values will be deleted by direct deletion method. For data containing outliers (such as

negative photovoltaic power, sudden power failure of photovoltaic power stations at some time, etc.), the window average method is used to correct.

Then Pearson correlation analysis is used for influencing factors in the data set (as shown in Formula 3). Screen out temperature, relative humidity, global horizontal radiation, diffuse horizontal radiation, rainfall and active power are the main features of this paper.

$$P_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

(3)

In the formula,  $P_{xy}$  represents the Pearson correlation coefficient between variable  $x$  and variable  $y$ ,  $x_i$  and  $y_i$  represents the values of variable  $x$  and variable  $y$  at the  $i$  th data point in the sample data set, respectively,  $\bar{x}$  and  $\bar{y}$  represents the mean of variable  $x$  and variable  $y$ , respectively,  $n$  represents the number of data points in the sample data set.

The problem of weight imbalance caused by different data dimensions is avoided. In this paper, other features except active power are normalized, and the normalization formula is shown as follows:

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

(4)

Where,  $x_{norm}$  indicates the normalized data,  $x$  indicates the original data,  $\min(x)$  indicates the minimum value of data set  $x$ ,  $\max(x)$  indicates the maximum value of data set  $x$ .

In order to reduce the noise and fluctuation in the data and better show the trend and periodicity of the data, triple exponential smoothing is performed on the active power, which is very helpful for predicting the future generation power or analyzing the long-term trend. The triple exponential smoothing formula is shown as follows:

$$S_t = \alpha \cdot Y_t + (1 - \alpha) \cdot (S_{t-1} + b_{t-1}) \quad (5)$$

$$b_t = \beta \cdot (S_t - S_{t-1}) + (1 - \beta) \cdot b_{t-1} \quad (6)$$

$$F_{t+m} = S_t + m \cdot b_t \quad (7)$$

In the triple exponential smoothing formula,  $S_t$  represents the smooth value at time  $t$ ,  $Y_t$  represents the original observation at time  $t$ ,  $b_t$  represents the trend value at time  $t$ ,  $F_{t+m}$  Represents the predicted value at time  $t+m$ ,  $\alpha$  and  $\beta$  are the smoothing coefficient and trend coefficient, respectively, controlling the degree of smoothness and the rate of change of the trend.

### 2.2 TCN model

Sequential convolutional network is a kind of structure for sequence modeling, which combines the characteristics of one-dimensional full convolutional network, causal convolutional network and extended convolutional network. Compared with recurrent neural network, it solves the problem of gradient disappearing or gradient explosion effectively, has the advantages of parallel computation, low memory consumption, and can control the sequence memory length by changing the receptive field. Fig.2 shows the schematic diagram of causal expansion convolution in TCN.

Given an input sequence  $X$ , and a filter  $F$ , the expansion convolution of  $x$  at time  $t$  is calculated as follows:

$$F(x_t) = \sum_{k=1}^K f_k \cdot x_{t-(K-k)d} \quad (8)$$

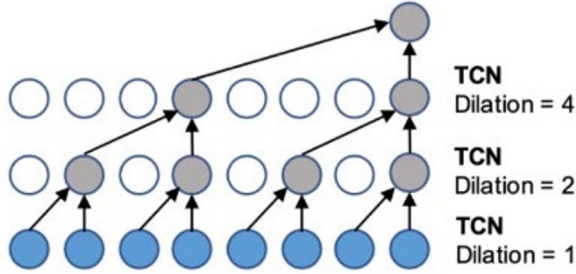


Fig.2. Schematic diagram of the TCN convolution structure.

Where  $f_k$  represents the  $K$ th weight parameter of the convolution kernel, which is used for the convolution operation on the input sequence,  $x_{t-(K-k)d}$  represents an element in the input sequence  $x_t$ , where  $k$  is the length of the convolution kernel,  $k$  is the index of the convolution kernel,  $d$  is the dilatation factor, which is used to control the sensitivity field of the convolution kernel.

Because of the depth of the neural network easy to appear gradient disappear or gradient explosion problem, so the residual connection is proved to be an effective approach of deep web training. Therefore, when designing the TCN model, the researchers chose to use residual blocks instead of a single convolutional layer. The residual block is implemented by adding the output from a series of transformations  $F$  to the input of the block, as follows:

$$y = x + F(x) \quad (9)$$

Where  $x$  represents the input feature,  $F(x)$  represents the transformation function of the residual block, and  $y$  represents the output of the residual block.

### 2.3 TPA-TCN model

Due to the large fluctuation and instability of PV power generation data and the fixed receptive field of convolution operation, the TCN model is limited by the long-term time dependence of capture. Based on this premise, it is proposed to combine the TPA with the TCN model, so that the model can better focus on the important time steps, improve the generalization ability of TCN model, the model structure of TPA-TCN is shown in Fig. 3.

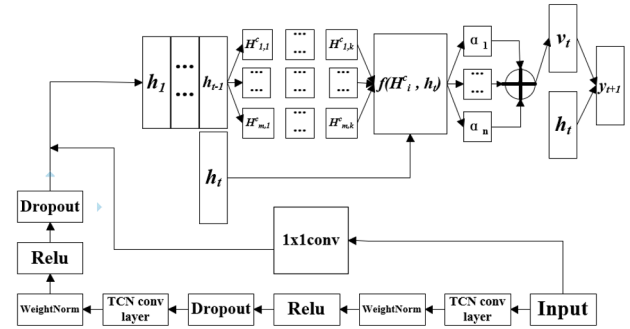


Fig.3. The model structure of TPA-TCN.

The TPA attention mechanism is an attention mechanism for processing time series data. It works by introducing the concept of time patterns on the basis of traditional attention mechanisms to better capture important patterns and features in time series. In the TPA attention mechanism, given a time series data, it is represented as  $H = \{h_1, h_2, \dots, h_n\}$ , where  $h_i$  represents the observed value at time  $i$ , using the TCN convolution to capture the variable signal pattern, the formula is as follows:

$$H_{i,j}^C = \sum_{l=1}^w H_{i,l-1} \cdot C_{j,l} \quad (10)$$

In the formula,  $H_{i,j}^C$  represents the weighted representation of attention at time step  $i$ ,  $H_{i,l-1}$  represents the feature vector at time step  $i$ ,  $C_{j,l}$  represents the context feature representation at time step  $j$ , and  $w$  represents the window size in the calculation of attention weight.

The extracted feature sequence is passed to the time mode encoder, and the attention mechanism is applied on the output of the time mode encoder, the functional formula for calculating attention weight is as follows:

$$f(H_i^C, h_t) = (H_i^C)^T W_\alpha h_t \quad (11)$$

Where,  $h_t$  represents the hidden state at the current time step  $t$ , and  $W_\alpha$  represents the parameter matrix of attention weight, which is used to linearly transform the hidden state shown in the context table  $H_i^C$  at the time step  $h_t$  and the current time step  $t$ .

Then the sigmoid function is normalized to calculate the attention weight to select among the multiple variables:

$$\alpha_i = \text{sigmoid}(f(H_i^C, h_t)) \quad (12)$$

Using the attention weight, the weighted summation of each line of  $H_i^C$  gives the context vector  $v_t$ :

$$v_t = \sum_{i=1}^m \alpha_i H_i^C \quad (13)$$

The context vector is spliced with the hidden state, and the final prediction result is generated by matrix multiplication:

$$y_{t+1} = W_h h_t + W_v v_t \quad (14)$$

In the TPA-TCN model, TCN will carry out the first feature extraction after receiving photovoltaic data, and capture the dependency between photovoltaic power and multiple features at the current moment through the receptive field. The residual block in TCN can effectively solve the problems of gradient disappearance and gradient explosion, and then through the TPA attention mechanism, enables the model to better focus on critical time steps, thereby capturing long-term time dependencies more effectively.

### 3 Example analysis

In this paper, the data set of DKASC's Alice hot spring photovoltaic system in Australia is selected as the research object. After data pre-processing, the data set is classified into sunny, cloudy and rainy days by k-means algorithm. The TPA-TCN model is used to predict the photovoltaic power generation in each weather. It is compared with TCN model, LSTM model and RNN model. All models are evaluated using Pytorch 1.13.1, cuda 12.3, and cudnn 10.0 on GeForce RTX 3080 GPU10G and Intel Xeon(R) Bronze 3204 CPU @ 1.90GHz×6. In the experiment, the number of residual modules used by the TCN model and the TPA-TCN is 3, the iteration times of the TPA-TCN model is 40, and the unified iteration times of other models is 80. dropout rate for all models is set to 0.5 and batch size to 16.

#### 3.1 Model prediction evaluation metrics

In order to verify the validity of the model, Mean Squared Error (MSE), Mean Absolute Error (MAE) and Coefficient of Determination ( $R^2$ ) were used to evaluate the prediction effect of the model, and the formula is as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (17)$$

Where  $n$  is the number of samples,  $y_i$  is the true value,  $\hat{y}_i$  is the predicted value, and  $\bar{y}$  is the mean of the true value.

#### 3.2 Forecast in sunny weather

After classification by k-means algorithm, 272 sunny days are obtained, and the sampling time is from 7 a.m. to 6 p.m., with a sampling interval of 5 minutes, a total of 36,176 data pieces are obtained. From Table I, we can see that the proposed TPA-TCN model has the best prediction accuracy under sunny weather. Compared with RNN model, MSE value and MAE value decreased by 24.49% and 36.75% respectively. Compared with LSTM model, MSE value and MAE value decreased by 22.04% and 35.02% respectively. Compared with the basic model TCN, MSE value and MAE value decreased by 13.03% and 29.75% respectively. Fig. 4-7 shows the test results.

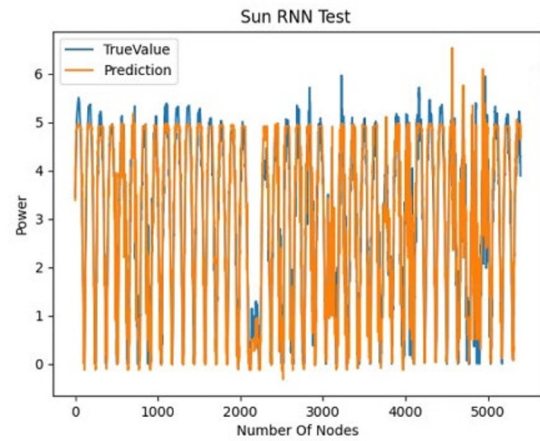


Fig.4. Performance of RNN model on test set in sunny weather.

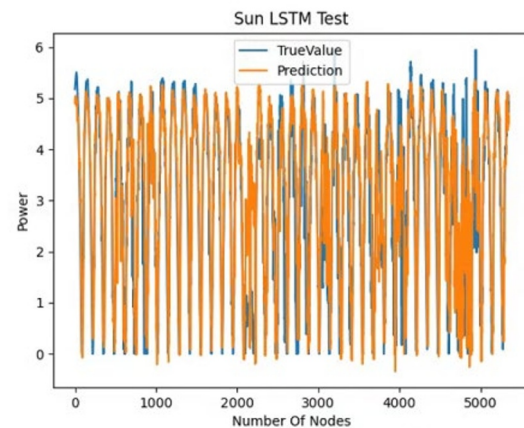


Fig.5. Performance of LSTM model on test set in sunny weather.

Table 1. Comparison of error in sunny weather prediction.

| Method  | MSE    | MAE    | $R^2$  |
|---------|--------|--------|--------|
| RNN     | 0.0698 | 0.1986 | 0.9269 |
| LSTM    | 0.0676 | 0.1933 | 0.9362 |
| TCN     | 0.0606 | 0.1788 | 0.9758 |
| TPA-TCN | 0.0527 | 0.1256 | 0.9881 |

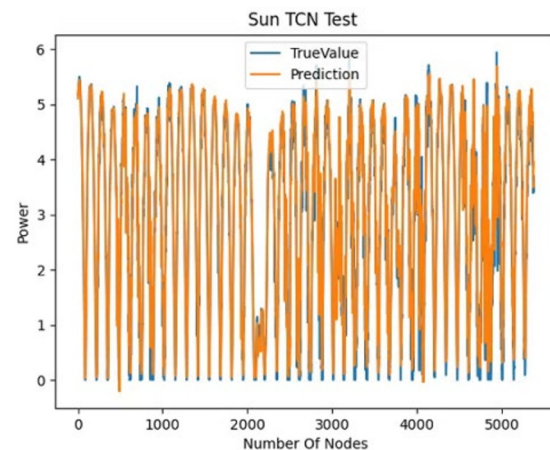
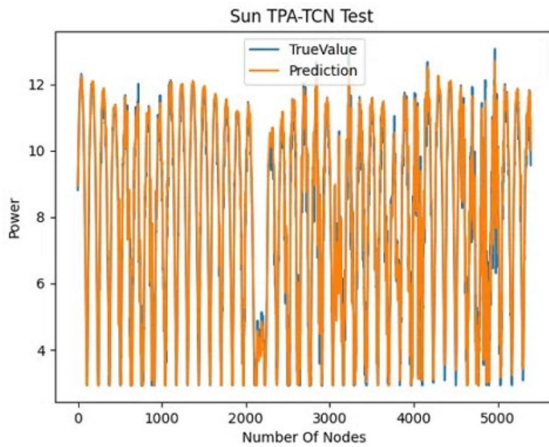


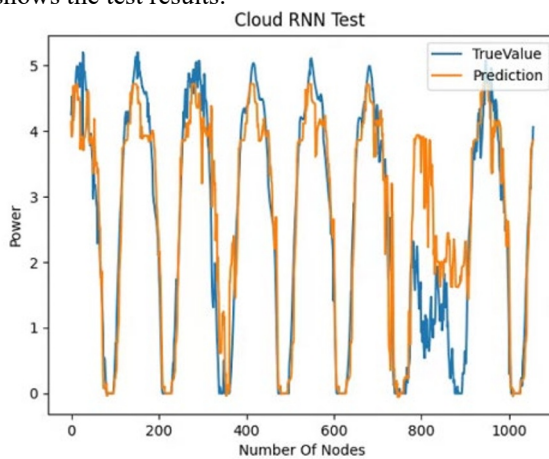
Fig.6. Performance of TCN model on test set in sunny weather.



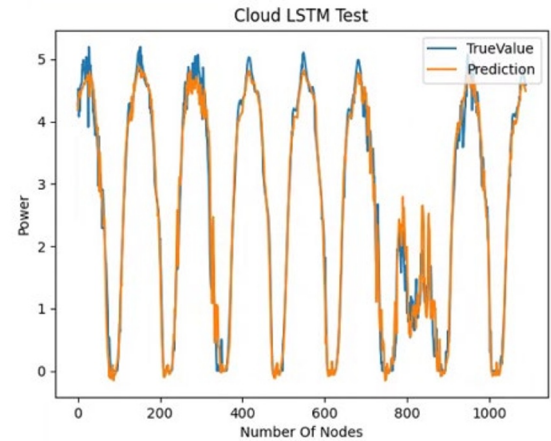
**Fig.7.** Performance of TPA-TCN model on test set in sunny weather.

### 3.3 Forecast in cloudy weather

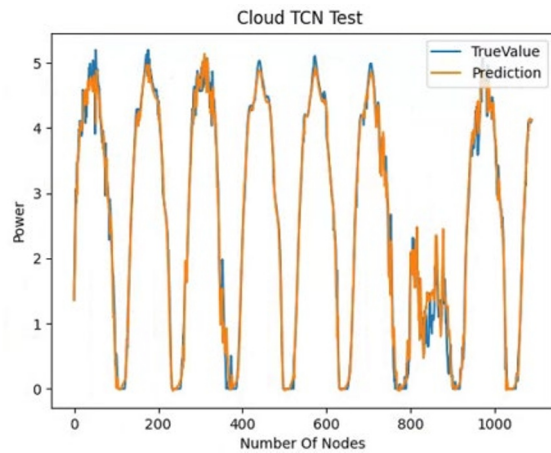
After k-means classification, 57 cloudy days are obtained, and the sampling time is from 7 a.m. to 6 p.m., with a sampling interval of 5 minutes, there are 7851 pieces of data. It can be seen from Table 2 that the proposed TPA-TCN model has the best prediction accuracy under cloudy weather. Compared with RNN model, MSE value and MAE value decreased by 86.63 % and 61.19 % respectively. Compared with LSTM model, MSE value and MAE value decreased by 78.18% and 30.28% respectively. Compared with the basic model TCN, MSE value and MAE value decreased by 13.35% and 9.62% respectively. Fig. 8-11 shows the test results.



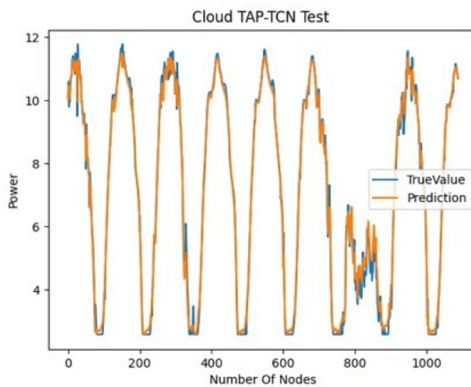
**Fig.8.** Performance of RNN model on test set in cloudy weather



**Fig.9.** Performance of LSTM model on test set in cloudy weather



**Fig.10.** Performance of TCN model on test set in cloudy weather



**Fig.11.** Performance of TPA-TCN model on test set in cloudy weather

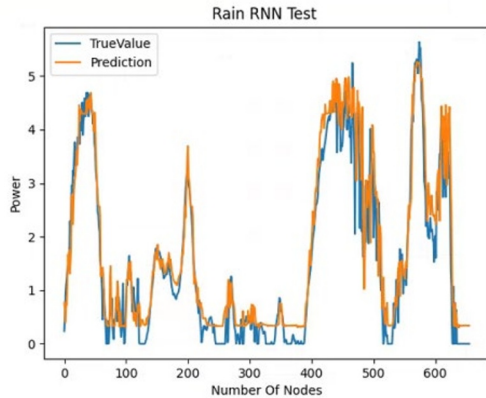
**Table 2.** Comparison of error in cloudy weather prediction.

| Method  | MSE    | MAE    | $R^2$  |
|---------|--------|--------|--------|
| RNN     | 0.1893 | 0.2806 | 0.8081 |
| LSTM    | 0.1160 | 0.1562 | 0.9763 |
| TCN     | 0.0293 | 0.1205 | 0.9854 |
| TPA-TCN | 0.0253 | 0.1089 | 0.9923 |

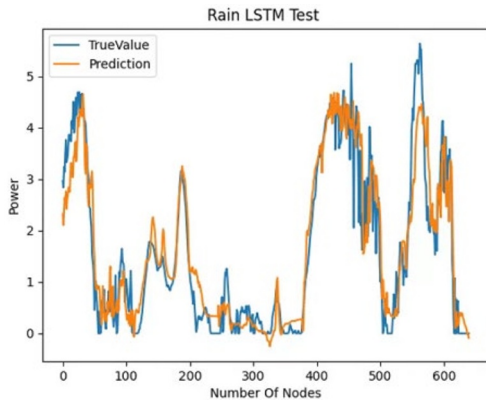
### 3.4 Forecast in rainy weather

After classification by k-means algorithm, 34 days of rainy days are obtained, and the sampling time is from 7 a.m. to 6 p.m., with a sampling interval of 5 minutes,

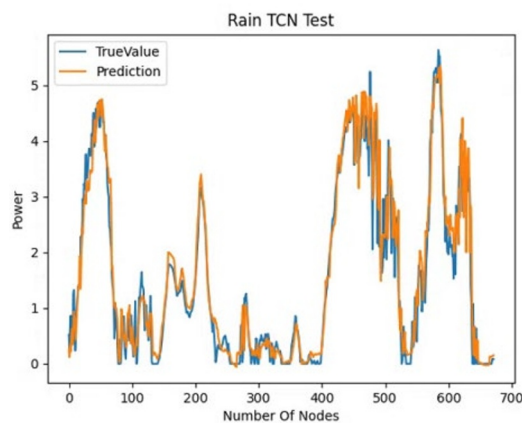
there are 4651 pieces of data. It can be seen from Table 3 that in rainy weather, the proposed TPA-TCN model has the best prediction accuracy. Compared with RNN model, MSE value and MAE value decreased by 66.25 % and 43.48 %, respectively. Compared with LSTM model, MSE value and MAE value decreased by 81.95% and 63.75% respectively. Compared with the basic model TCN, MSE value and MAE value decreased by 53.95% and 28.49%, respectively. Fig. 12-15 shows the test results.



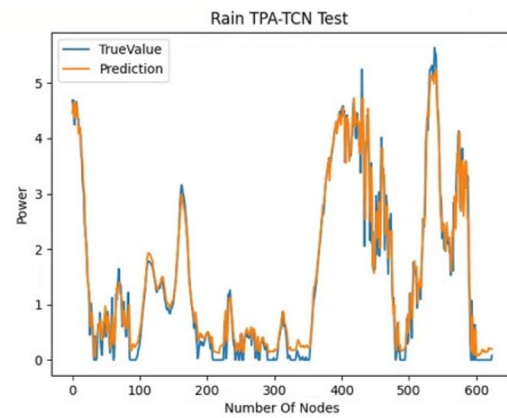
**Fig.12.** Performance of RNN model on test set in rainy weather



**Fig.13.** Performance of LSTM model in rainy weather



**Fig.14.** Performance of TCN model on test set in rainy weather



**Fig.14.** Performance of TPA-TCN model on test set in rainy weather

**Table 3.** Comparison of error in rainy weather prediction.

| Method  | MSE    | MAE    | $R^2$  |
|---------|--------|--------|--------|
| RNN     | 0.0655 | 0.1918 | 0.9174 |
| LSTM    | 0.1225 | 0.2991 | 0.8823 |
| TCN     | 0.0480 | 0.1516 | 0.9407 |
| TPA-TCN | 0.0221 | 0.1084 | 0.9720 |

## 4 Conclusion

In this study, a deep learning model TPA-TCN is established to solve the problem of large power fluctuation and instability of photovoltaic power generation. Due to the particularity of the time series, missing values are deleted on a daily basis. After that, Pearson correlation analysis is used to screen important features. After that, the active power is triple-smooth-processed to better show the trend and periodicity of the data. In order to test the accuracy of the model under different weather conditions, the k-means algorithm is used to divide the PV power generation data set into sunny, cloudy and rainy days, and the three data sets are tested separately. The TPA attention mechanism is combined with the TCN model to capture long-term dependencies in time series more effectively. The weights are dynamically adjusted according to the importance of each time step. Finally, through the test of three kinds of weather, it is proved that the prediction accuracy of the model in this dataset is better than that of RNN, LSTM and TCN.

However, the model was not tested separately on the data sets of the four seasons, and further testing of the generalization ability of the model and using other advanced optimization algorithms to optimize the hyperparameters of the model to improve the accuracy is one of the future research work.

This work was supported in part by the National Nature Science Foundation of China under Grant 62201195; and in part by the Doctoral Startup Foundation of SJZ university under Grant (23BS022, 21BS017). China University Industry Research Innovation Fund (2021ITA05002, 2021ITA05004).

## References

1. G. Chen, Y. Xu, X. Yuan, et al. Research on electrical parameter modeling of flexible photovoltaic cell arrays. *J.* **46**, 134-142 (2020).
2. W. Zheng, W. Zheng, Z. Li, et al. Current status and progress trends of online monitoring technology for photovoltaic power plants. *J.* **58**, 1-7 (2021).
3. Y. Zhang . Power dispatching can be predicted China Strategic Emerging Industries, *J.* 14-16 (2021).
4. Lateko, Andi AH, et al. Stacking ensemble method with the RNN meta-learner for short-term PV power forecasting, *Energies* 14. **16** , 4733 (2021).
5. M. N. Akhter , S. Mekhilef , H. Mokhlis , et al. An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants, *Energies, J.* **15** , 2243(2022).
6. Zhen, Hao, et al. Photovoltaic power forecasting based on GA improved Bi-LSTM in microgrid without meteorological information, *Energy* 231, 120908 (2021).
7. Qing, \*\*angyun, N. Yugang . Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM, *Energy* 148. 461-468 (2018).
8. Chen, Hailang, \*\*anfa Chang. Photovoltaic power prediction of LSTM model based on Pearson feature selection, *Energy Reports* 7. 1047-1054 (2021).
9. Dai, Yeming, et al. LOWESS smoothing and Random Forest based GRU model: A short-term photovoltaic power generation forecasting method, *Energy* 256. 124661(2022).
10. Lago, Jesus, F. D. Ridder, B. D. Schutter. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms, *Applied Energy* 221. 386-405 (2018).
11. S. M. Babbar, C. Y. Lau, K. F. Thang . Long term solar power generation prediction using adaboost as a hybrid of linear and non-linear machine learning model, *International Journal of Advanced Computer Science and Applications.* J.12(11) (2021).
12. W. Hu, X. Zhang, L. Zhu, et al. Short-term photovoltaic power prediction based on similar days and improved SOA-DBN model, *IEEE Access. J.* **9**, 1958-1971 (2020).
13. T. Limouni , R. Yaagoubi , K. Bouziane , et al. Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model, *Renewable Energy. J.* **205**, 1010-1024 (2023).
14. H. Li , G. Ma , Y. Meng . Short-term photovoltaic output prediction model based on MOSMA feature selection and TCN, *IEEE. C.* 1800-1806 (2023).
15. Shih, Y. Shun , K. S. Fan , L. Hung-yi . Temporal pattern attention for multivariate time series forecasting, *Machine Learning* 108. 1421-1441 (2019).