

Photovoltaic Short-term Power Prediction based on PSO-GRU Model

Jicheng Liu¹, Wenqing Zheng¹, Xiangnan Liu¹

School of Economics and Management, North China Electric Power University, Changping, Beijing, 102206, China.

Abstract

Faced with energy depletion and increasingly serious environmental pollution, photovoltaic power generation has gradually become an effective way of power supply and energy development, and short-term photovoltaic power prediction has gradually attracted attention. This paper proposes a short-term photovoltaic power prediction method based on the PSO-GRU model. After the PSO joins the adaptive mutation operation, the key hyperparameters in the GRU are optimized to reduce the training cost. Subsequently, GRU is used for power prediction, and compared with other algorithms to evaluate the performance of the PSO-GRU model. The results show that the PSO-GRU model is more suitable for short-term power prediction under photovoltaic environment, and the prediction effect is better.

Keywords

Photovoltaics; Short-term Power Prediction; PSO; GRU; Hyperparameter Optimization.

1. Introduction

With the extensive use of traditional energy, energy exhaustion and environmental pollution are becoming more and more serious, which strongly promotes the vigorous development of clean energy, and photovoltaic power generation has become one of the most potential clean energy power generation technologies [1-3]. However, photovoltaic power generation itself is vulnerable to the interference of external environment, which has the characteristics of randomness and instability, which affects the control of power grid system to a certain extent. Therefore, it is necessary to predict the power in the photovoltaic environment more accurately, to better control the operation of the power system and to the healthy development of the photovoltaic industry.

With the development of a new generation of information technology, people pay more attention to accurate and efficient short-term power prediction [4, 5]. In recent years, a large number of scholars have paid more and more attention to the application of prediction methods such as machine learning and deep learning, and used intelligent algorithms to optimize parameters and improve accuracy in prediction. However, the limitations of a single algorithm are difficult to overcome, and the prediction results have large errors. Therefore, the combination of algorithms to optimize a single algorithm's prediction method has gradually become the method that more scholars choose when optimizing parameters and improving prediction accuracy. In Literature[6], in order to solve the problem of data loss caused by too long time sequence, a power prediction method based on wavelet decomposition of two-way long and short-term memory network is proposed, which uses wavelet to decompose time domain information and frequency domain information, and then adopts two-way long and short time The memory network makes predictions. Literature [7] uses the improved particle swarm algorithm to search for the optimal weights of the DBN neural network and establishes the initial DBN network. The meteorological data and historical power generation data of the

predicted data are used as the training set to train the initial DBN network, and the established prediction model has higher prediction accuracy. Literature [8] proposed an ultra-short-term photovoltaic power prediction method based on the VMD-DESN-MSGP model. Use VMD to decompose the time series modally, and establish DESN forecasting models for each modal. In order to further improve the prediction accuracy of the model, MSGP is used to compensate the prediction error. Finally, the predicted value of the error and the predicted value of the original power are superimposed as the final prediction result.

Although there are many researches and methods for power prediction, there are still few researches on power prediction in photovoltaic environment. One of the main reasons is that photovoltaic power generation is easily affected by the external environment to produce abnormal values that affect the prediction results. Therefore, on the basis of considering the advantages of algorithm combination to make up for the shortcomings of a single algorithm, this paper proposes a power prediction model combining particle swarm optimization (PSO) and gated recurrent unit (GRU) network, and uses PSO algorithm to optimize the GRU algorithm. The key hyperparameters are used to weaken the influence of abnormal power values caused by the instability of the photovoltaic environment and improve short-term photovoltaic power prediction.

This article has made the following innovations both academically and practically. First, according to the literature, the predictive indicators that may be related to photovoltaic power are identified, and the corresponding data is collected using crawler technology. Secondly, in order to solve the problems of falling into the local optimal solution prematurely and the slow optimization speed in the later stage, this paper proposes to optimize the key hyperparameters in the GRU network through the PSO algorithm to improve the data training effect and reduce the data training cost. Compare the optimization results of the PSO algorithm in the scene. Next, the GRU network uses the optimized hyperparameters to predict the short-term power in the photovoltaic environment, and compares the prediction results of the PSO-GRU prediction model with other five algorithms to verify the credibility of the PSO-GRU prediction model. Adaptability. Finally, according to the prediction results and comparative analysis, the short-term power prediction under the photovoltaic environment is summarized.

2. GRU network prediction model based on PSO algorithm to optimize hyperparameters

2.1. PSO algorithm

The particle swarm optimization algorithm (PSO) is a global random search algorithm based on swarm intelligence. For optimization problems, the position of each particle in the particle swarm is a solution to the optimization problem. The particle swarm algorithm calculates the velocity and the velocity of each particle by iteration. Position to obtain the optimal solution[9]. Among them, the particle velocity and position update formula are as follows:

$$v_i^k = \omega v_i^{k-1} + c_1 rand_1() (pbest_i^k - x_i^{k-1}) + c_2 rand_2() (gbest_i^k - x_i^{k-1}) \quad (1)$$

$$x_i^k = x_i^{k-1} + v_i^{k-1} \quad (2)$$

Where ω is the inertia weight, c_1 and c_2 are the acceleration factors, and represents two random parameters, with a value range of [0,1], $rand_1()$ and $rand_2()$ are random number in the range of [0,1], and, $pbest_i^k$ and $gbest_i^k$ are the particles and the front k of the particle group, respectively The optimal solution for the second iteration.

2.2. Hyperparameter optimization based on PSO

Hyperparameters are parameters whose values are set before starting the learning process, not parameter data obtained through training. In recent years, with the continuous increase of deep

reinforcement learning models, the training cost, that is, the search space of hyperparameters, has also continued to increase, resulting in excessively high time costs for finding a better hyperparameter configuration [10-13]. Therefore, in order to improve prediction efficiency and reduce training costs, this paper proposes a PSO-based hyperparameter optimization method. The specific optimization steps are as follows:

Step 1 Initialize the number of particles n , the maximum number of iterations K , the number of hyperparameters m , and the velocity and position of the particles.

Step 2 Determine the fitness function according to the research problem, calculate the fitness of each particle, and update the particle position iteratively. The model is trained using the parameter information contained in the current position of the particle, and the fitness value of each particle is calculated.

Step 3 Mark the current optimal position of the individual particle and the optimal position of the group.

Step 4 Determine whether the maximum number of iterations has been reached. If not, update the particle velocity and position, and go to step 2. If yes, get the optimal hyperparameter and perform the assignment.

2.3. GRU network prediction based on optimal hyperparameters

The Gated Recurrent Unit (GRU) network is a variant of the LSTM network in the deep learning RNN algorithm. The principle behind GRU is very similar to LSTM, that is, the gating mechanism is used to control input, memory and other information, and to make predictions at the current time step. But different from the LSTM network, GRU has fewer training parameters than LSTM, and reduces the internal hidden state and related gate structure than LSTM. The former has three doors, while there are only two door structures in the GRU, namely the update door and the reset door, as shown in Figure 1. The update gate controls the extent to which the state information at the previous moment is retained in the current state, and the reset gate determines whether the current state is combined with the previous information. The specific calculation process is as follows:

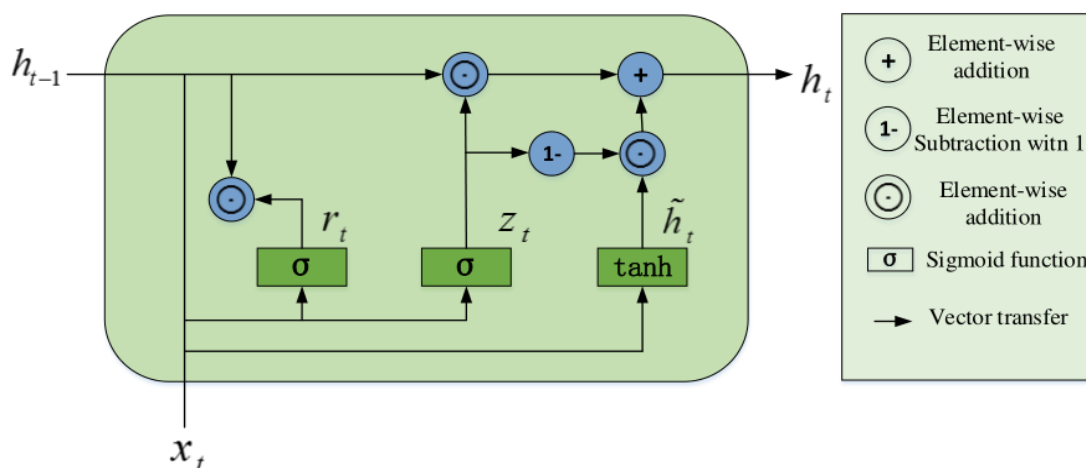


Figure 1. Basic structure of GRU

Firstly, calculate the update gate z_t at time t , as in formula (3). x_t is the input vector of the t -th time step, which will undergo a linear transformation (multiply by the weight matrix W^z), and h_{t-1} will save the information of the previous time step $t-1$. Secondly, the reset gate r_t is calculated at time t . The calculation process is similar to the update gate, as shown in formula (4). In the process of using the reset gate, the new content $\tilde{h}_t$ will use the data in the historical step stored in the reset gate, as shown in formula (5). The input x_t and the previous time step

information h_{t-1} first undergo a linear transformation, that is, the matrices W and U are multiplied to the right. Calculate the Hadamard product of the reset gate r_t and, that is, the product of the corresponding elements with. Finally, the network needs to calculate the final memory h_t of the current time step. This vector will retain the information of the current unit and transfer it to the next unit, as shown in formula (6).

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1}) \quad (3)$$

$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1}) \quad (4)$$

$$\tilde{h}_t = \tanh(r_t * U h_{t-1} + W x_t) \quad (5)$$

$$h_t = (1 - z_t) * \tilde{h}_t + z_t * h_{t-1} \quad (6)$$

In addition, this article adds a function to achieve exponential decay learning rate, and uses Adam gradient descent method to achieve the purpose of reducing the loss value. The calculation formula.

$$m_t = \eta[\beta_1 m_{t-1} + (1 - \beta_1)g_t] \quad (7)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) * \text{diag}(g_t^2) \quad (8)$$

Among them, m_t represents the first-order momentum, and v_t represents the second-order momentum. β_1, β_2 represents the random objective function. In the initial iteration stage, the two momentums have an offset to the initial value, which can be offset correction and gradient update.

3. Case analysis and result display

3.1. Preparation work

3.1.1. Determine the power prediction index

Photovoltaic power generation is relatively affected by external environmental factors. Therefore, when considering power prediction in a photovoltaic environment, it is unreasonable to only use historical power generation as a predictive indicator. Based on a large number of references, this article identifies as many factors as possible that affect photovoltaic power generation, that is, the predictive index of photovoltaic power prediction. In the end, this article has determined seven predictive indicators from two levels of photovoltaic resources and meteorological factors. The index of the photovoltaic resource layer mainly affects the power generation quality of photovoltaic resources and photovoltaic generators to change the power output, including solar radiation, sunshine hours and rainfall. The indicators of the meteorological factor layer mainly affect the power generation efficiency of photovoltaic generators to change the power output and affect the power prediction, including average temperature, average wind speed, average air pressure and relative humidity. This article will also use these 7 indicators to predict the power generation under the photovoltaic environment.

3.1.2. Data collection and processing

By browsing related websites, crawling the historical data of the corresponding standards using crawler technology. Finally, the historical data of a photovoltaic field in Jiangsu, China is determined as the forecast data of this article.

Before using the data to make predictions, the collected data must be preprocessed, that is, data cleaning and normalization. This paper calculates the mean value of missing features and replaces it with missing values as a method of data cleaning. This method can offset the missing data and produce better results. The normalization of data is to make the features between different dimensions have a certain numerical comparison. This article adopts the classic MinMaxScaler method, and its calculation method is

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (9)$$

3.2. PSO-GRU power prediction results

3.2.1. PSO algorithm optimizes hyperparameters

In order to reduce the training cost of the GRU algorithm and reduce the impact of abnormal values caused by the randomness and instability of the photovoltaic environment, according to the introduction in section 2.1, this paper uses the IPSO optimization algorithm to perform the key hyperparameters in the GRU algorithm. optimization. This paper chooses to optimize the learning rate and the number of nodes in the hidden layer in the GRU algorithm, and the learning rate search interval is set in the range of [0.001, 0.01]. Since this paper sets the hidden layer in the GRU algorithm to two layers, the PSO algorithm will optimize the number of nodes in the two hidden layers, and the search interval for the number of nodes is set within the range of [1-200]. According to the choice of hyperparameters, the fitness function in this paper is defined as the mean square error between the predicted power value and the actual value, which is defined as follows

$$f(x_i) = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x}_i(\theta))^2 \quad (10)$$

m is the number of samples in the test set, x_i is the power value of the i -th predicted sample, and $\bar{x}_i(\theta)$ is the predicted expected value of the i -th predicted sample. After confirming the fitness function, the improved PSO algorithm sets the number of initial particles $n=5$, the maximum number of iterations $k=10$, the maximum inertia weight $\omega_{max}=0.8$, the minimum inertia weight $\omega_{min}=0.2$, the acceleration factor c_1 and c_2 the selected value are 1.5, r_1 and r_2 represents the choice of 1.5 for both random parameters. After 10 iterations of optimization, the learning rate is 0.0054 (Figure 2(a)), and the optimal number of nodes in the first hidden layer is 33 (Figure 2(b)). The optimal number of nodes in the second hidden layer is 32 (Figure 2(c)).

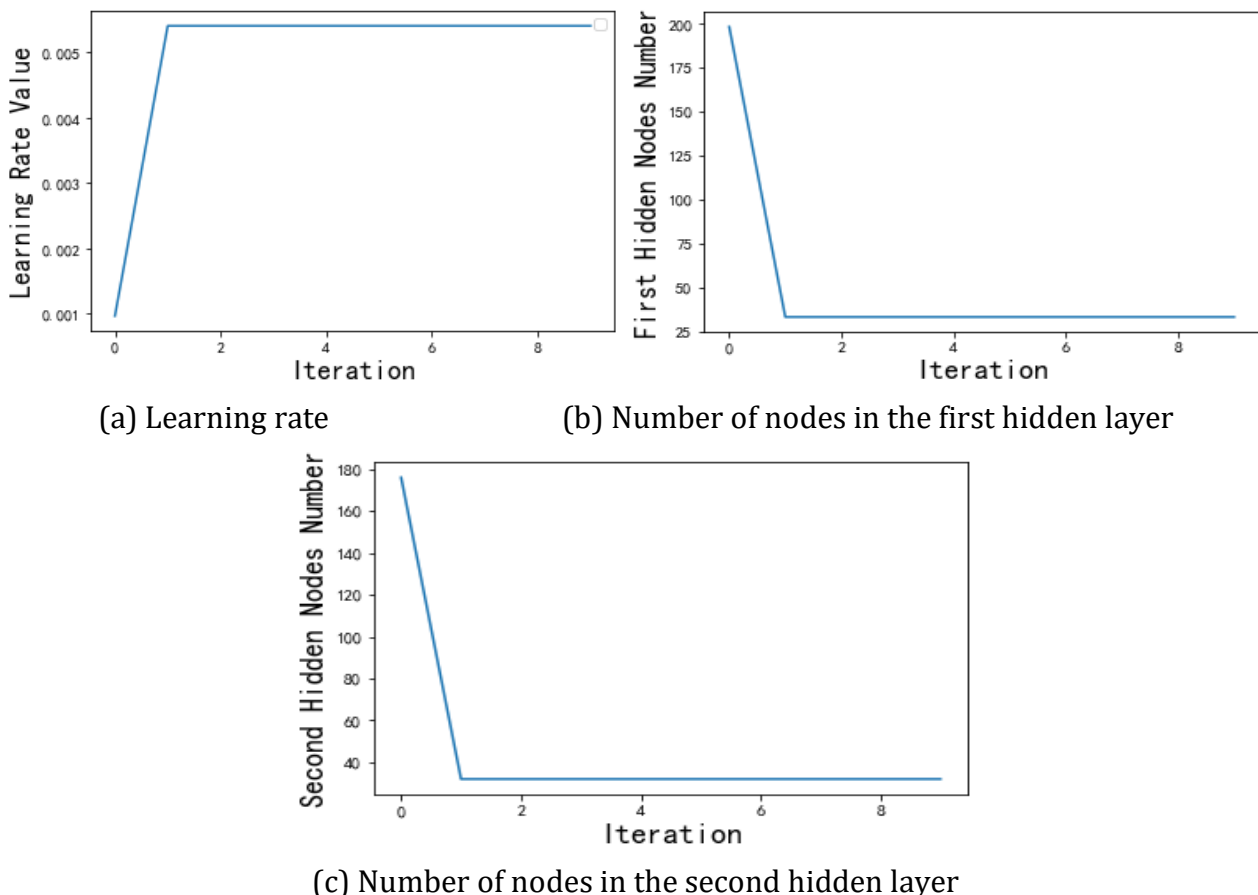


Figure 2. Results of hyperparameter optimization of PSO

In order to further illustrate the rationality of the design of the IPSO algorithm, this paper analyzes the fitness situation under different iteration times by comparing 4 sets of different acceleration factors. As shown in Figure 3, among the 4 sets of acceleration factor pairs, the fastest optimizing speed is $c_1=2.4/c_2=1.8$ acceleration factor pairs, but as the number of iterations increases, $c_1=1.5/c_2=1.5$ acceleration factor pairs in 150 iterations At this time, the optimization speed exceeds $c_1=2.4/c_2=1.8$ acceleration factor pair and other acceleration factor pairs, and it is also the fastest acceleration factor pair close to 0. This shows that the acceleration factor pair set in this article $c_1=1.5/c_2=1.5$ has a stronger ability to optimize the four sets of acceleration factor pairs, which helps to reduce the prediction of abnormal values caused by the instability of the photovoltaic environment to a certain extent error.

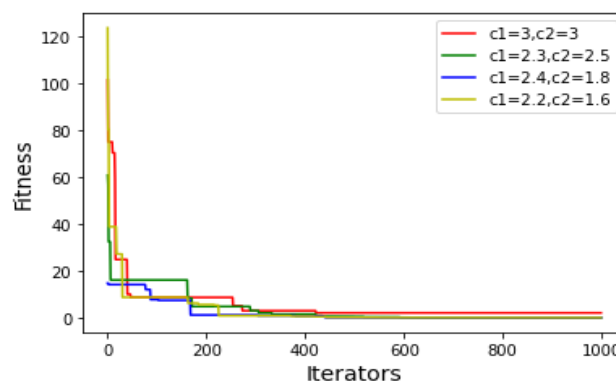


Figure 3. Acceleration factor fitness situation comparison

3.2.2. PSO-GRU algorithm power prediction

Through the hyperparameter optimization of the PSO algorithm, the number of nodes in the first hidden layer of GRU is set to 35, the number of nodes in the second hidden layer is set to 34, and the learning rate is 0.0054. In this paper, the iterative series epoch of GRU is set to 250, and the batch-size of a set of training data is 60.

Before making the actual prediction, it is necessary to complete the selection of the data set and the determination of the model parameters, which is of great significance to the calculation of the PSO-GRU model. In this article, the training set and the test set are divided according to the ratio of 8:2. After calculation, the goodness of fit of the test set can reach 0.98116 (Figure 4(d)), and the goodness of fit of the test set can reach 0.94419 (Figure 4(b)). The predicted value is in good agreement with the sample data. The proposed model has a good fitting effect. The red line in (a) (c) represents the sample data of the test set and training set, the blue line represents the predicted value of the test set and training set, and the red line in (b) (d) is the best fit obtained by the analysis line.

Figure 5 shows the relationship between the number of iterations and the prediction loss, which can reflect the convergence of the PSO-GRU prediction algorithm. In the course of 250 iterations, the loss of the training set dropped rapidly to below 0.05 after the 10th iteration. Although the loss of the test set fluctuates greatly during the iteration process because of the special value, the overall loss is also below 0.036, and after the 10th iteration, it also rapidly drops to 0.018, and finally drops to 0.0165. Comparing the loss of the training set and the test set, it can be found that although the initial loss of the test set is much smaller than the training set, the training set declines faster than the test set, and the final loss of the training set The degree is also smaller than the test set. Overall, the loss of PSO-GRU in the prediction process is very small. After the prediction of the model, the comparison result chart of the predicted result and the actual value is obtained (Figure 6), and the prediction accuracy can reach 97.915% in the end.

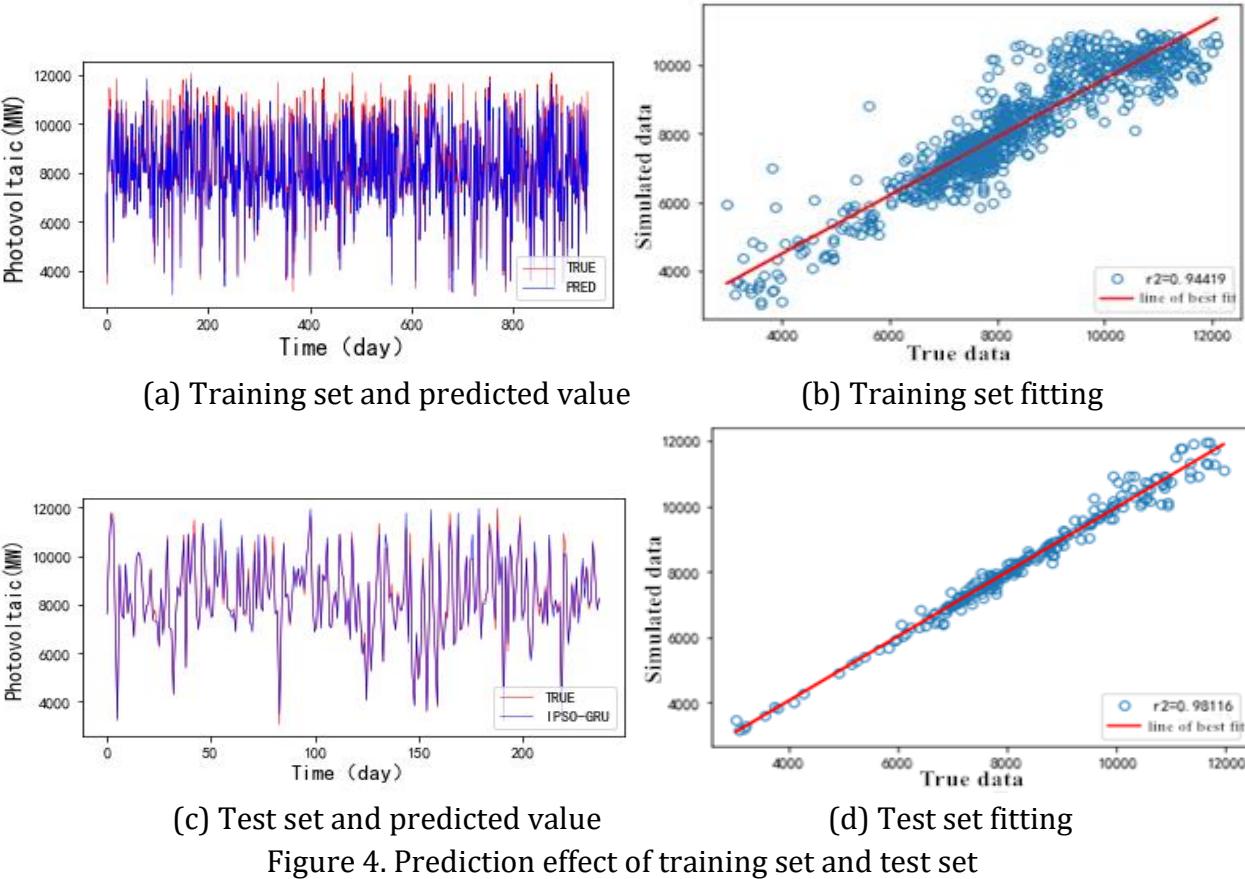


Figure 4. Prediction effect of training set and test set

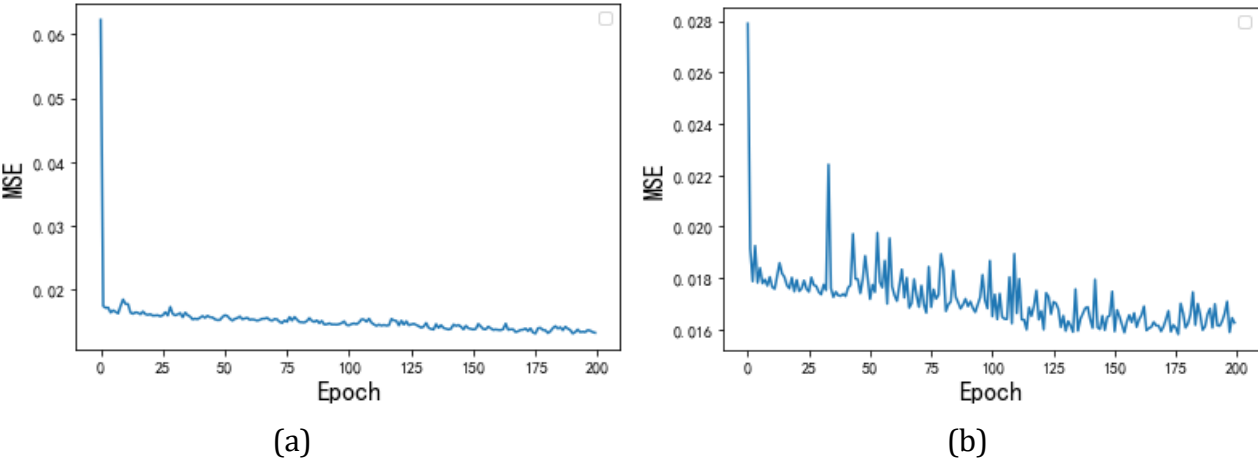


Figure 5. Training set and Test set loss curve

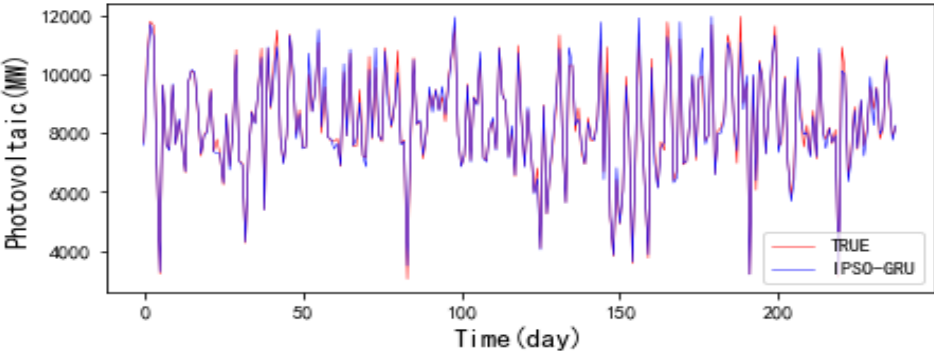


Figure 6. PSO-GRU forecast results

3.3. Algorithm comparison analysis

3.3.1. Algorithm evaluation index

The abnormal values caused by the randomness and instability of the photovoltaic environment will have error values when predicting the results, and these errors are unavoidable. In order to better analyze and evaluate the IPSO-GRU forecasting model, this paper selects 5 commonly used evaluation indicators, including average percentage error (MAPE), root mean square error (RMSE), average absolute error (MAE), mean square error (MSE) and goodness of fit (R^2).

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \bar{y}_l}{y_i} \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_l)^2} \quad (12)$$

$$\text{MAE} = \frac{\sum_{i=1}^n e_i}{n} = \frac{\sum_{i=1}^n |y_i - \hat{y}_l|}{n} \quad (13)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_l)^2 \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_l)^2}{\sum_{i=1}^n (y_i - y_{true})^2} \quad (15)$$

Among them, n is the number of predicted samples, y_i is the predicted value, $\hat{y}_l$ is the actual value, and $\bar{y}_{true}$ is the average of the actual values.

3.3.2. Comparison of prediction results between PSO-GRU model and other algorithms

In order to better illustrate the superiority of the IPSO-GRU model, this paper also selects the GRU algorithm, LSTM algorithm, RNN algorithm, random forest (RF) algorithm, support vector machine (SVM) five algorithms to compare with the IPSO-GRU model. And according to the evaluation index in 3.3.1, 6 algorithms are analyzed.

Table 1. Algorithm comparison table

	IPSO-GRU	GRU	LSTM	RNN	RF	SVM
MAPE	0.01868	0.01987	0.01891	0.02984	0.11924	0.07984
RMSE	245.13518	301.0918	387.6546	495.2398	571.5464	465.0231
MAE	161.39802	159.90007	234.65740	329.234	446.7682	389.6781
MSE	245.13518	249.79034	239.79034	457.34645	512.1209	415.5231
R^2	0.98116	0.97067	0.98067	0.96234	0.95781	0.96101

It can be seen from Table 1 that the MAPE value of the PSO-GRU model is 0.01868, which is the smallest among the six algorithms, indicating that the prediction error rate of the PSO-GRU model is the smallest. To a certain extent, it reflects that the PSO-GRU model can be more effectively deal with the abnormal values caused by the instability of the photovoltaic environment and improve the prediction accuracy. Among the six algorithms, the smallest RMSE value is also the PSO-GRU model. It can be seen that the RMSE value of the GRU is greater than the PSO-GRU model, which also reflects that the optimization of the PSO algorithm makes the prediction result of the GRU algorithm more accurate. The MSE value of the PSO-GRU model is 245.13518, which is only smaller than the 239.79034 of the LSTM algorithm. This is because the GRU algorithm is essentially a simplified version of the LSTM algorithm. It is a bit rougher than the LSTM algorithm when processing data, so the loss value is higher. However, the MSE value of the PSO-GRU model as a whole is lower than most algorithms, which also shows that the loss of the PSO-GRU model is a bit lighter. At the same time, the PSO-GRU model is the smallest among the MAE values of all algorithms, indicating that the PSO-GRU model can better reflect the actual situation of the predicted value error.

In addition, the PSO-GRU model with a goodness of fit of 0.97116 is the best fit, which is greater than the goodness of fit of other algorithms, indicating that the prediction value of the IPSO-GRU prediction model is compared with other algorithms. It is more accurate with the actual value. On the whole, the PSO-GRU model is more suitable for power prediction in a photovoltaic environment, and has advantages in exception handling and accuracy improvement. Although the GRU algorithm is highly efficient, its prediction accuracy is lower than that of the PSO-GRU algorithm and the LSTM algorithm, but higher than the RNN algorithm with more complex processing.

4. Conclusion

As one of the fastest-growing clean energy sources, photovoltaic resources can not only supply energy demand at a lower cost, but also improve environmental pollution. Accurate and efficient short-term power prediction is a prerequisite for meeting power demand, and it is also the basis for ensuring and supporting the stability and safety of the power system.

This paper proposes a power prediction algorithm based on the combination of PSO algorithm and GRU network, which reduces the cost of prediction data training and improves the accuracy and efficiency of the prediction results. The PSO-GRU algorithm selects the three key hyperparameters of the learning rate and the number of nodes in the two hidden layers in the GRU algorithm for optimization, and adds an adaptive variable operation to reduce the probability of mutation. On this basis, the selected predictive indicators and historical power generation data are input into the GRU prediction algorithm for prediction. Finally, compare the results of the PSO-GRU model with other five algorithms. From the perspective of the prediction results, the PSO-GRU model is more suitable for power prediction in the photovoltaic environment, which is beneficial to eliminate errors caused by abnormal values, and the accuracy of the prediction results is also more obvious. Through the comparison of the prediction results of different algorithms, it can be seen that how to more effectively eliminate the abnormality of photovoltaic resource data caused by geographic location and how to more effectively reduce the abnormal value in the power output caused by the instability of the photovoltaic environment is optimized. The true direction of the prediction algorithm is also an important concern for improving the accuracy of the prediction results.

Although innovations have been made and more precise results have been obtained, this article also has some limitations and deficiencies. In subsequent studies, in order to make predictions more accurate, we will think about the selection of predictive indicators, and reduce the subjectivity of indicator selection by interpreting structural models or structural equation models.

Acknowledgments

This work is supported by National Nature Science Foundation of China, grant number is 71771085.

References

- [1] BAE K Y, JANG H S and SUNG D K: Hourly solar irradiance prediction based on support vector machine and its error analysis, *IEEE Transactions on Power Systems*, Vol.32(2017) No.2, p.935-945.
- [2] CECI M, CORIZZO R and FUMAROLA F: Predictive modeling of PV energy production:how to set up the learning task for a better prediction? , *IEEE Transactions on Industrial Informatics*, Vol. 13 (2017) No.3, p. 956-966.
- [3] OSPINA J,NEWAZ A and FARUQUE M O: Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model, *IET Renewable Power Generation*, Vol. 13 (2019) No.7, p. 1087-1095.

- [4] ZHAO Bing, W ANG Zengping, JI Weijia, GAO Xin and LI Xiaobing: A Short-term Power Load Forecasting Method Based on Attention Mechanism of CNN-GRU, *Power System Technology*, Vol. 43 (2019) No.12, p. 4370-4376 .
- [5] Jicheng Liu and Yinghuan Li: Study on environment-concerned short-term load forecasting model for wind power based on feature extraction and tree regression, *Journal of Cleaner Production*, Vol. 264 (2020), p. 305-307.
- [6] XIE Xiaoyu, ZHOU Junhuang, ZHANG Yongjun, WANG Jiang and SU Jieying: W-BiLSTM Based Ultra-short-term Generation Power Prediction Method of Renewable Energy, *Automation of Electric Power Systems*, Vol. 45 (2021) No.8, p. 175-184.
- [7] LI Zhengming, LIANG Caixia and WANG Manshang: Short-term power generation output prediction based on a PSO-DBN neural network, *Power System Protection and Control*, Vol. 48 (2018) No.8, p. 149-154.
- [8] WANG Su, JIANG Xin, ZENG Liang and CHANG Yufang: Ultra-short-term Photovoltaic Power Prediction Based on VMD-DESN-MSGP Model, *Power System Technology*, Vol. 44 (2020) No.3, p. 917-926.
- [9] REN Liqiang, ZHANG Limin, WANG Haipeng and GUO Qiang: Short-term power load interval forecasting based on improved particles warm optimization and Gaussian process regression, *Computer Engineering and Design*, Vol. 40(2019) No.10, p. 3002-3008.
- [10] WANG Dongfeng and MENG Li: Performance Analysis and Parameter Selection of PSO Algorithms, *Acta Automatica Sinica*, Vol. 42(2016) No.10, p. 1552-1561.
- [11] ALFI Alireza: PSO with Adaptive Mutation and Inertia Weight and Its Application in Parameter Estimation of Dynamic Systems, *Acta Automatica Sinica*, Vol. 37(2011) No.5, p. 541-549.
- [12] Jiang Yun-liang, Zhao Kang, Cao Jun-jie, Fan Jing and Liu Yong: Asynchronous Parallel Hyperparameter Search with Population Evolution, *Control and Decision*, Vol.37(2020)No.5,p.1-10.
- [13] Zhiqiang Geng, Ju Bai, Deyang Jiang and Yongming Han: Energy structure analysis and energy saving of complex chemical industries: A novel fuzzy interpretative structural model, *Applied Thermal Engineering*, Vol. 142(2018), p. 433-443.