EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

Distributed Photovoltaic Power Energy Generation Prediction Based on Improved Multi-objective Particle Algorithm

Yuanzheng Xiao[1*], Huawei Hong[1], Feifei Chen[1], Xiaorui Qian[1], Ming Xu[1], Hanbin Ma[2]

[1]State Grid Fujian Marketing Service Center (Metering Center and Integrated Capital Center), Fuzhou 350013, Fujian Province, China
[2]State Grid Info-telecom Great Power Science and Technology Co.,Ltd., Fuzhou 350003, China

Abstract

Accurate prediction of distributed photovoltaic (DPV) power generation is crucial for stable grid operation, yet existing methods struggle with the non-linear, intermittent nature of solar power, and traditional machine learning models face hyperparameter selection and overfitting challenges. This study developed a highly accurate DPV power prediction method by optimizing a Long Short-Term Memory (LSTM) network's hyperparameters using an improved Multi-Objective Particle Swarm Optimization (MO-PSO) algorithm. A hybrid LSTM-PSO model was created, where the LSTM network served as the core prediction model, and the improved MO-PSO algorithm optimized its hyperparameters, enhancing generalization and avoiding overfitting. The LSTM-PSO model significantly improved prediction accuracy compared to traditional methods. Key results from two power stations included a maximum deviation of 6.2 MW at Power Station A, a peak time deviation of less than 0.1 MW at Power Station B, and a prediction interval error controlled below 30 MW at an 80% confidence level. The optimized LSTM-PSO model effectively captures DPV power generation dynamics, and the superior performance metrics demonstrate its potential for intelligent grid management. However, limitations include prediction accuracy under extreme weather and computational efficiency for large datasets. Future work will focus on broader applicability and more efficient algorithm variants.

Keywords: Particle algorithm, Distributed photovoltaic power generation, Power prediction, Long short-term memory network, Intelligent power grid.

Received on 20 December 2024, accepted on 09 March 2025, published on 13 March 2025

Copyright © 2025 Y. Xiao et al., licensed to EAI. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/ ew.8901

1. Introduction

With the transformation of the global energy structure and the increasing awareness of environmental protection, the utilization of renewable energy, especially solar energy, has received unprecedented attention [1]. As a clean and pollution-free energy source, the development and utilization of solar energy are of great significance for reducing greenhouse gas emissions and alleviating energy crises [2-3]. Distributed Photovoltaic Power Generation (DPPG) systems play an important role in solar energy utilization due to their flexibility and reliability [4]. However, the intermittency and instability of Solar Power Generation (SPG) pose challenges to the stable operation of the power grid. Accurately predicting photovoltaic power generation has become the key to achieving effective grid connection and optimizing scheduling. In traditional research, photovoltaic power prediction methods mainly include physical models, statistical models, and Machine Learning (ML) methods [5]. Physical models rely on meteorological parameters such as solar radiation and temperature, but are often limited by parameter simplification and the complexity of actual meteorological conditions. Statistical methods such as Auto-Regressive Integrated Moving Average (ARIMA) models, although less dependent on historical data, are difficult to

*Corresponding author. Email: huangf0305@163.com

capture complex nonlinear relationships. ML, especially neural networks, performs well in handling nonlinear problems, but traditional algorithms still have shortcomings in hyperparameter selection and overfitting problems [6]. In response to the above shortcomings, this study proposes an innovative prediction method, namely a hybrid model (LSTM-PSO) based on Long Short-Term Memory Network (LSTM) and improved Multi-objective Particle Swarm Optimization (MO-PSO). The innovation lies in combining advanced deep learning techniques with intelligent optimization algorithms to achieve breakthroughs in the accuracy of photovoltaic power prediction.

2. Literature Review

Many experts have conducted research on the improved MO-PSO algorithm. Hu et al. developed an improved algorithm called FCAN-MOPSO, which significantly improves the performance of complex network graph clustering through instance frequency weighting and MO-PSO [7]. Suresh M et al. utilized an enhanced MO-PSO algorithm to optimize hybrid energy systems, aiming to reduce energy costs and improve the reliability of household applications in remote areas [8]. Kahhal et al. combined response surface methodology and MO-PSO to optimize the friction stir welding process of AH12 1050 aluminum alloy, and found that low speed and high feed rate help improve the mechanical properties of the material [9]. Rahimi et al. explored the evolutionary algorithms of single objective PSO and MO-PSO under constrained conditions through literature review and econometric analysis methods. Compared with single objective optimization, constraint handling techniques in Multi-Objective Optimization (MOO) have received less attention. The most promising algorithms included genetic algorithm, differential evolution algorithm, and particle swarm intelligence. It was expected that future research would further increase in engineering, computer science, and mathematics [10].

With the increasing demand for renewable energy, the prediction and optimization technology of solar photovoltaic power generation has become crucial. Ge L et al. reviewed the virtual acquisition technology of Distributed Photovoltaic Systems (DPVS) and proposed application methods such as similarity analysis, providing development references for the DPVS industry [11]. Iheanetu K J focused on SPG prediction technology, especially the analysis of data-driven methods, laying the foundation for improving model accuracy [12]. Akhter et al. used the RNN-LSTM model in deep learning to predict photovoltaic power plants with different technologies and found that it outperformed other methods in accuracy and robustness [13]. AlKandari et al. designed a fusion model combining ML and statistical methods, which improves the accuracy of SPG prediction through structural and data diversity techniques, as well as different ensemble methods, and performs well on real datasets. These studies provided new perspectives and methods for efficient prediction and management of SPG [14].

In summary, many experts have conducted research on complex network graph clustering, hybrid energy system optimization, MOO of mechanical performance, constrained optimization algorithms, DPVS virtual acquisition technology, SPG prediction technology, and SPG prediction hybrid models. However, current research still has shortcomings such as limited model generalization ability, low algorithm computational efficiency, weak adaptability to practical application scenarios, and the need to improve prediction accuracy and stability. Therefore, this study further proposes optimization algorithms to enhance model robustness, improve prediction accuracy, and explore more practical application scenarios.

3. Application of Improved PSO in DPPG System

3.1 DPPG Energy System Based on PSO

DPPG is a technology that combines photovoltaic systems with power systems to convert solar energy into electrical energy. It converts solar energy into direct current by setting up an inverter between the photovoltaic array and the power grid, and then the inverter converts it into alternating current that matches the power grid, which is used by the power grid or directly supplied to users [15]. Compared with traditional centralized photovoltaic power generation systems, DPPG systems have higher energy utilization efficiency and power quality, while improving the reliability of power supply [16]. Its working principle is based on the photovoltaic effect, where solar radiation acts on silicon solar panels to generate current, as shown in Figure 1.

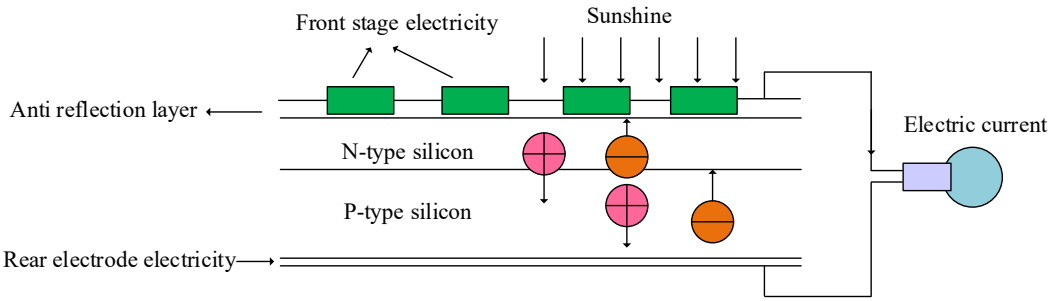


Figure 1 Schematic diagram of DPPG system operation

The DPPG system mainly consists of solar photovoltaic module arrays and distributed inverters. The photovoltaic module array converts solar energy into direct current, while the inverter converts direct current into alternating current for user use or transmission to the power grid. The system automatically runs when there is sufficient sunlight, and stops immediately when there is insufficient solar radiation or abnormal power grid to ensure safety and stability. The performance of the system is affected by internal and external factors such as photovoltaic materials, construction management, and solar radiation intensity, as shown in Figure 2.

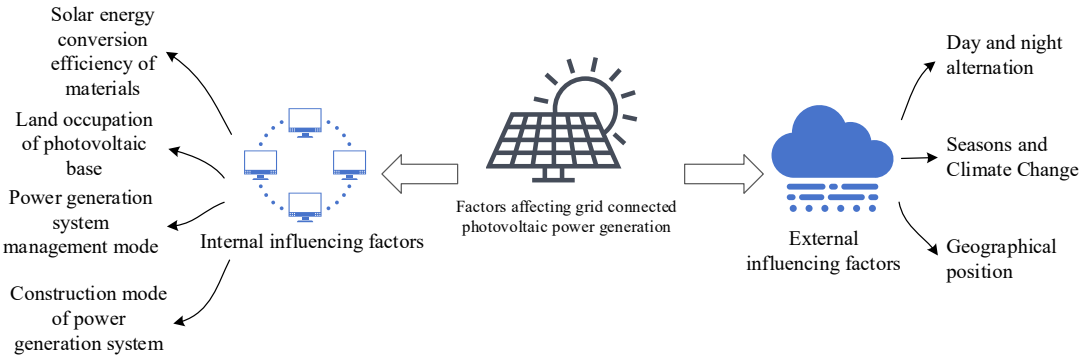


Figure 2 Factors affecting grid-connected photovoltaic power generation

The photoelectric conversion efficiency of the DPPG system under ideal conditions is approximately 15%. Although there is room for improvement in this efficiency compared to traditional power generation methods, the advancement of technology and deepening of research and development indicate that breakthroughs in photovoltaic conversion efficiency are expected in the future. This not only means the improvement of energy utilization efficiency, but also represents the optimization of energy structure and the diversification of energy supply. The development of DPPG system is not only related to energy acquisition, but also to the sustainable utilization of energy. The MO-PSO simulates the hunting behavior of bird flocks. It can balance the search for local and global optimal solutions when solving problems, and the parameter settings of the algorithm are relatively simple. In the DPPG system, the idea of MO-PSO can be borrowed to optimize system design and operation. In the DPPG system, each photovoltaic module can be regarded as a "particle" that performs a "search" in an n-dimensional parameter space to find the optimal working state, and has a certain motion speed that can be adjusted based on flight experience. Assuming there is a particle swarm in the search space of dimension S, which contains m particles. Particle i specifically represents a S-dimensional vector, and its flight velocity is also an S-dimensional vector $\vec{V}_i = (V_{i1}, V_{i2}, ..., V_{iS})$. Among them, $i = 1, 2, ..., m$, the position of particles also represents the potential solution of the optimization problem, and the fitness value is calculated by substituting $\vec{x}_i$ into the objective function. Assuming that the current optimal solution or position searched by the particle swarm is vector $\vec{P}_{gS} = (P_{g1}, P_{g2}, ..., P_{gS})$, and the current optimal solution of a single particle i is $\vec{P}_{iS} = (P_{i1}, P_{i2}, ..., P_{iS})$. The objective function that PSO algorithm should minimize is $f(x)$, and the formula for the current optimal solution $p_i(t+1)$ of particle i can be obtained as shown in equation (1).

$$p_i(t+1) = \begin{cases} p_i(t) & f(x_i(t+1)) \geq f(p_i(t)) \\ X_i(t+1) & f(x_i(t+1)) < f(p_i(t)) \end{cases} \quad (1)$$

In equation (1), the value range of i is $[1, m]$, and

the current velocity $v_{is}(t+1)$ of particle i is equation (2).

$$v_{is}(t+1)=v_{is}(t)+c_1r_{1s}(t)[p_{is}(t)-x_{is}(t)] +c_2r_{2s}(t)[p_{gs}(t)-x_{is}(t)] \quad (2)$$

In equation (2), the range of values for s is $[1,S]$. r_1 and r_2 are continuous variables that follow a uniform distribution of $[0,1]$. c_1 and c_2 are both non negative constants. The range of values for velocity v_{is} is $[-v_{\max},v_{\max}]$, where $v_{\max}$ is the maximum value of particle velocity. The position $x_{is}(t+1)$ of particle i is updated as shown in equation (3).

$$x_{is}(t+1)=x_{is}(t)+v_{is}(t+1) \quad (3)$$

If the search space is within the range of $[-x_{\max},x_{\max}]$, there is $v_{\max}=kx_{\max}$, where $k\in[0.1,1.0]$ represents the elastic modulus. In equation (4), a control factor is introduced to limit the current speed from being affected by the speed of the previous moment.

$$v_{is}(t+1)=\omega v_{is}(t)+c_1r_{1s}(t)[p_{is}(t)-x_{is}(t)] +c_2r_{2s}(t)[p_{gs}(t)-x_{is}(t)] \quad (4)$$

In equation (4), ω is a non negative coefficient. When the value is high, the current speed is more affected by the speed of the previous moment, which enhances the global search ability of the algorithm. On the contrary, when the value is low, the current speed is less affected by the speed of the previous moment, thereby enhancing the local search ability of the algorithm. Therefore, this study introduces the MO-PSO, and Figure 3 shows the calculation process.

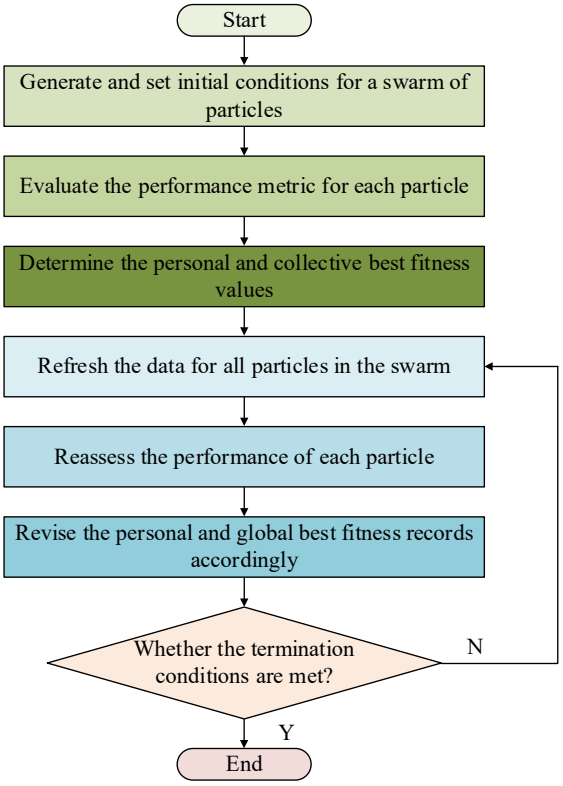


Figure 3 Calculation flow chart of PSO algorithm

Based on Figure 3, the process for calculating PSO is as follows: the first is to create an initial particle swarm containing particles m, each with a random initial position and velocity. Next is to calculate the fitness of each particle, which is the key to evaluating its performance. Then, the historical optimal position p_{is} of each particle is compared and updated, and if the current fitness is better, it is replaced with the new optimal position p_{is}. Subsequently, according to the algorithm formula, the positions and velocities of all particles are determined, and the fitness is re-evaluated, while determining the optimal fitness values for individuals and populations. Finally, to determine whether the termination condition is met. If it is met, to output the optima and stop. Otherwise, to return to the fitness calculation step to continue iterating. Algorithm initialization includes setting the number of particles m, generating corresponding positions x_{is} and corresponding velocities v_{is} for all i、s within the range of $[-x_{\max},x_{\max}]$ and $[-v_{\max},v_{\max}]$ according to a uniform distribution rule. The third step is to set rule $y_i=x_i$, where $i=1,2,...,m$.

3.2 DPPG Prediction Technology Based on LSTM and MO-PSO Algorithms

With economic growth, energy consumption has intensified environmental pressure, prompting the country to promote clean energy, among which photovoltaic power generation is becoming increasingly important due to its clean characteristics [17]. Although DPPG systems face challenges in power prediction due to variables such as weather and holidays, especially in densely populated residential areas, the application of LSTM models can improve the accuracy and efficiency of predictions by solving the gradient problem of long sequence data [18]. Figure 2 shows the unit structure of LSTM.

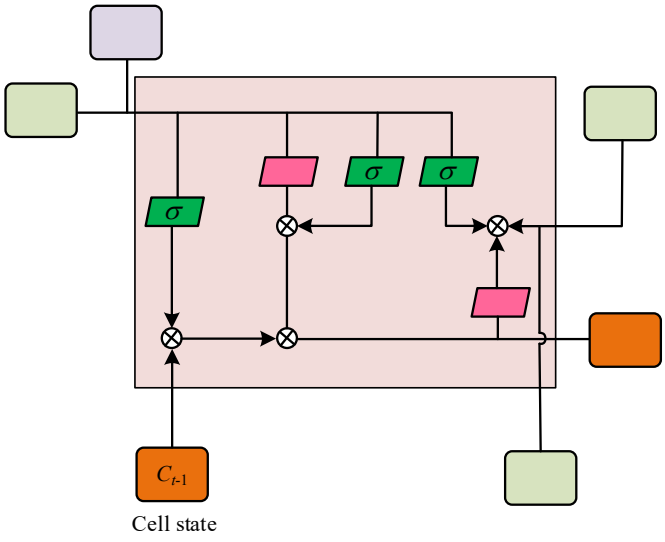


Figure 4 LSTM unit structure

In Figure 4, at a given moment, the input information of LSTM includes cell state, hidden layer output, and input sequence. In LSTM, input, forget, and output gates can be obtained through concatenation. The output value of the input gate i_t is calculated as shown in equation (5).

$$i_t = \sigma\left(W_i\left[x_t, h_{t-1}\right]^T + b_i\right) \quad (5)$$

In equation (5), σ is the activation function. W represents the input gate Weight Matrix (WM). x_t is the input sequence at time t. h is the output of the hidden layer. T is the transpose matrix. b_i is the input gate Bias Matrix (BM). The output value of the forget gate f_t is calculated as equation (6).

$$f_t = \sigma\left(W_f\left[x_t, h_{t-1}\right]^T + b_f\right) \quad (6)$$

In equation (6), W_f and b_f are the forgetting gate's WM and BM. The output value of the output gate o_t is calculated is equation (7).

$$o_t = \sigma\left(W_o\left[x_t, h_{t-1}\right]^T + b_o\right) \quad (7)$$

In equation (7), W_o and b_o are the output gate's WM and BM. To optimize the hyperparameters of the LSTM algorithm for power prediction, this study uses the MO-PSO algorithm for hyperparameter optimization to improve the accuracy of the algorithm. When optimizing, a multidimensional space is established and the positions and velocities of particles in the space are represented to extract the individual optimal positions of particles, as shown in equation (8).

$$H_i^k = \left(h_{i1}, h_{i2}, ..., h_{iD}\right)\ i = 1, 2, ..., N \quad (8)$$

In equation (8), H_i^k is the individual optimal position of the i-th particle. h_{iD} is the individual best position parameter. The expression for extracting the global optimal position of particles is equation (9).

$$G_i^k = \left(g_1, g_2, ..., g_D\right) \quad (9)$$

In equation (9), G_i^k is the global optimum location of the i-th particle. g_D is the global optimal position parameter. The velocity update calculation of particles is equation (10).

$$V_i^{k+1} = \omega V_i^k + c_1 r_1\left(H_i^k - X_i^k\right) + c_2 r_2\left(G^k - X_i^k\right) \quad (10)$$

In equation (10), V_i^{k+1} and V_i^k are the particle velocities after and before the update. ω is the inertia factor. c is the acceleration factor. r is a random number between 0 and 1. k is the current iteration count. The position update calculation of particles is equation (11).

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (11)$$

In equation (11), X_i^k and X_i^{k+1} are the particle positions before and after the update. Through continuous iteration and optimal fitness calculation, the global optimal position and fitness are obtained. After data preprocessing, an LSTM prediction model is constructed. Through experiments, it is determined that the model uses 40 historical data points per day, each with 39 features, resulting in an input layer dimension of 40*39. The output layer is single dimensional and corresponds to the predicted power generation. Appropriate activation functions are selected to avoid gradient vanishing, and backpropagation is used to adjust parameters during training to optimize model performance. The loss function design can reflect the error between the predicted and real data during training in real time, and the relevant expression is equation (12).

$$\theta^* = \arg\min_{\theta} \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i, \theta) + \lambda\Phi(\theta) \quad (12)$$

In equation (12), the loss function is L. The empirical risk function is set to θ, the predicted value is

$f(x)$, and the true value is y. N is a quantity, and $\Phi(\square)$ is a function. The Root Mean Square Error (RMSE) is taken as the assessment index for prediction performance, as shown in equation (13).

$$\varepsilon_{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\frac{\hat{P}(i)-p(i)}{Cap_i}]^2} \quad (13)$$

In equation (13), ε_{RMSE} represents the training performance of the model. $p(i)$ and $\hat{P}(i)$ are the actual and predicted values of power generation. n is the number of predicted validation data. i is the predicted point sequence number. The maximum value of the output sequence of $\hat{P}(i)$, and Cap_i is the maximum value of the predicted sequence power. This study will name the optimized algorithm LSTM-PSO algorithm. The functions of the constructed photovoltaic power generation prediction platform are shown in Figure 5.

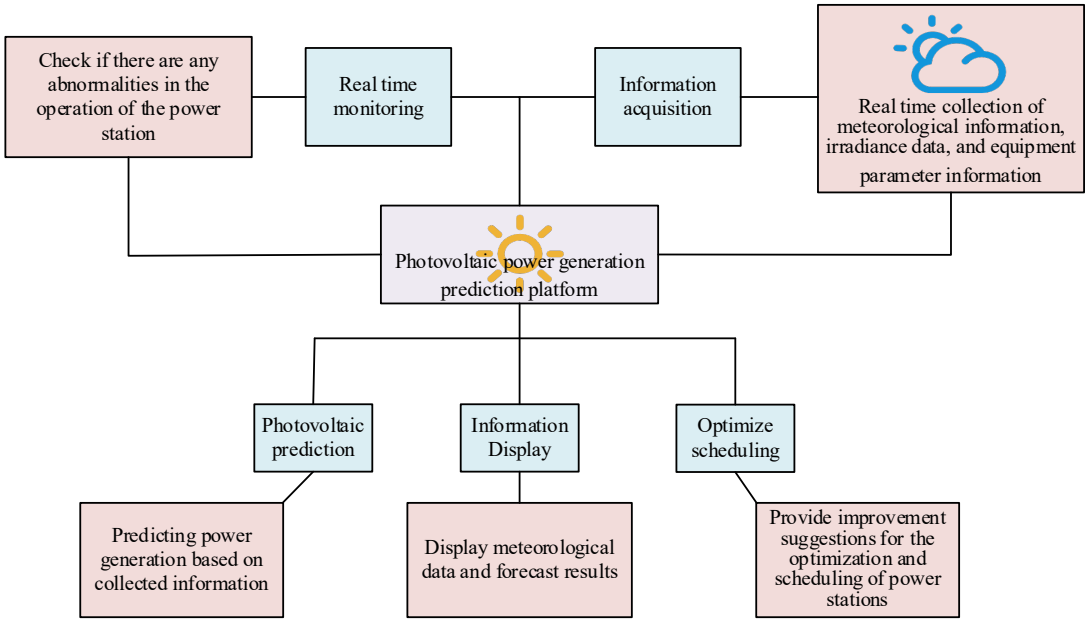


Figure 5 Functions of photovoltaic prediction platform

Figure 5 shows an integrated platform with five core functions: real-time monitoring, data collection, load forecasting, information display, and optimized scheduling. Among them, load forecasting, as the core component of the platform, uses the LSTM deep learning framework to process the received data and construct a model for prediction. The predicted results are then stored in a database for other applications to access and use.

4. Algorithm Validation of Photovoltaic Prediction and Positioning Accuracy

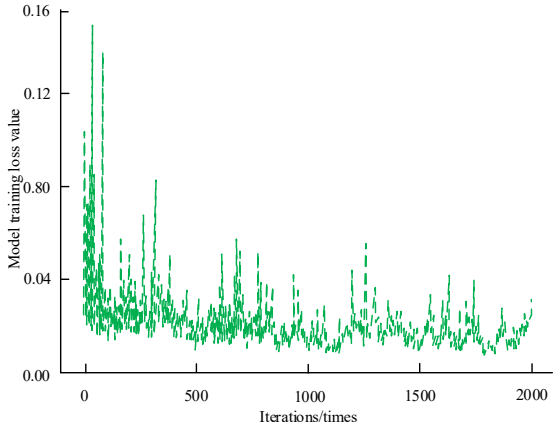
A detailed analysis is conducted on the parameters of the constructed LSTM model, especially the setting of the hidden layer. By selecting different numbers of hidden layers, namely 9, 10, and 11 layers, the impact of the hidden layers on the predictive performance of the model are studied. Table 1 summarizes these analysis results and demonstrates the specific impact of layer changes on prediction performance.

Table 1 The Influence of Hidden Layers on Training Results

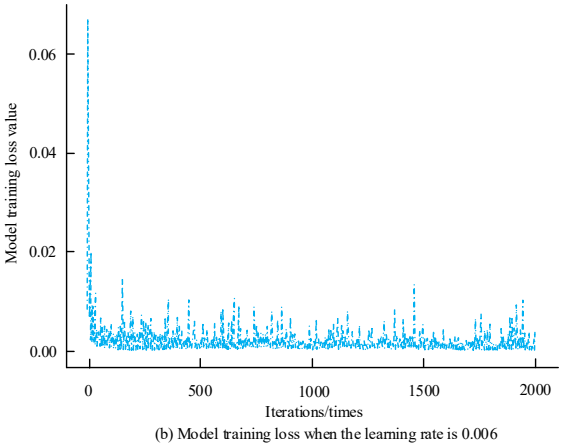
Number of hidden layers	Iteration number	Learning rate	Step	Training block size	RMSE

9					0.011-0.014
10	2000	0.0006	40	50	0.0011-0.0040
11					0.013-0.021

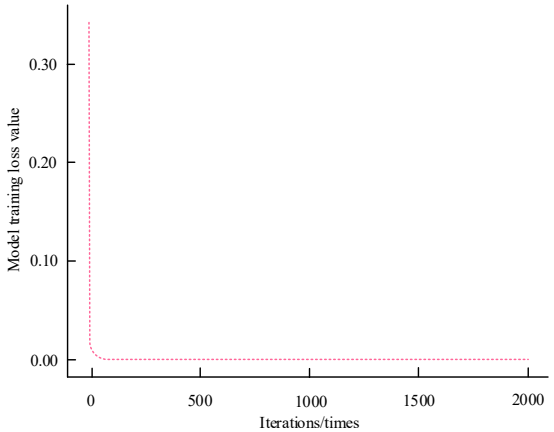
In Table 1, all models run under fixed training parameters, with the only variable being the number of hidden layers. When the hidden layer is 10, the model reaches the lowest RMSE, thus determining 10 layers as the optimal hidden layer. The variation of model training loss under different learning rates is shown in Figure 6.



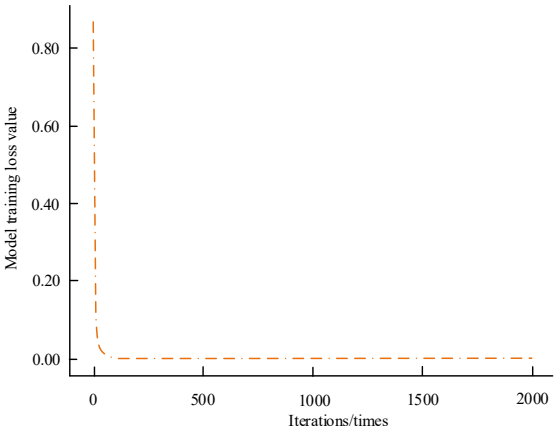
(a) Model training loss when the learning rate is 0.06



(b) Model training loss when the learning rate is 0.006



(c) Model training loss when the learning rate is 0.0006



(d) Model training loss when the learning rate is 0.00006

Figure 6 Model training loss values at different learning rates

In Figure 6 (a), when the learning rate is 0.06, the model has a high loss value after 2000 iterations and significant oscillations during the training process, indicating that the training effect is not ideal. In Figure 6 (b), the learning rate decreases to 0.0061, and the oscillation amplitude decreases after the model loss value decreased. Although it still does not fully converge after 2000 iterations, the error is reduced to some extent. In Figure 6 (c), the learning rate is further reduces to 0.0006, and the model quickly converges after 45 iterations, demonstrating good training performance. In Figure 6 (d), when the learning rate drops to 0.00006, the model converges faster, but requires more iterations to reach a stable state, with a convergence iteration of 75, which is 30 more than Figure 6 (c). This indicates that higher learning leads to poor model training performance, while lower learning rates result in better training performance. Considering training efficiency, a learning rate of 0.0006 is more suitable as the final model parameter. Node distribution is carried out within a 100×100 area, with 70 target nodes, 30 Anchor Nodes (ANs), and a communication radius of 30. The experiment is simulated

using 40 populations. Orthogonal Projection Algorithm (OPA) is used as the benchmark for comparison. Three PSO algorithms are tested, namely LSTM-PSO, Backtracking Control PSO (BCS-PSO), and ordinary PSO algorithm. The impact of different anchor node numbers on the number of nodes that can be located and the Relative Positioning Error (RPE) of each algorithm is displayed in Figure 7.

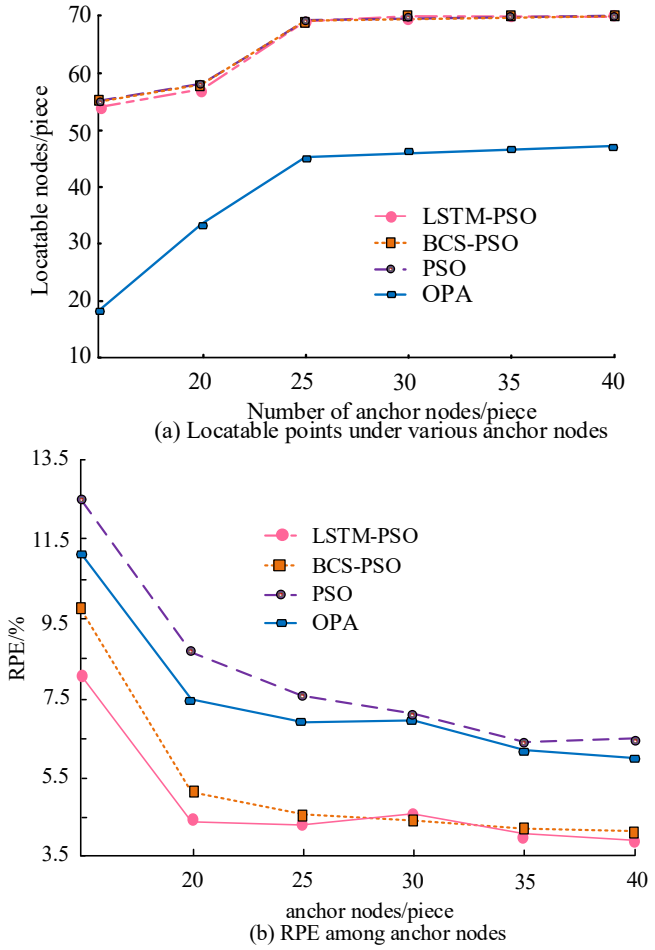


Figure 7 The locatability and positioning accuracy of various algorithms with varying numbers of Ans

In Figure 7 (a), as the ANs increase, the locatable nodes of the three PSO algorithms initially increase slowly, then accelerate, and eventually stabilize, showing a high degree of consistency. Meanwhile, the locatable nodes in OPA tend to stabilize after increasing and remain lower than that in PSO. Specifically, when the anchor node is 25, the locatable node in LSTM-PSO is 68, which is 23 superior than OPA, and the maximum amount of locatable nodes in LSTM-PSO reaches 70, significantly higher than OPA algorithm. Figure 7 (b) shows that as the ANs grows, the RPEs of the four algorithms gradually decrease. Under the same quantity of ANs, PSO has the highest RPE, followed by OPA, while LSTM-PSO has the lowest error. Especially when the anchor node reaches 40, the RPE of LSTM-PSO is minimized to 3.4%. The results of predicting the daily photovoltaic power are shown in Figure 8.

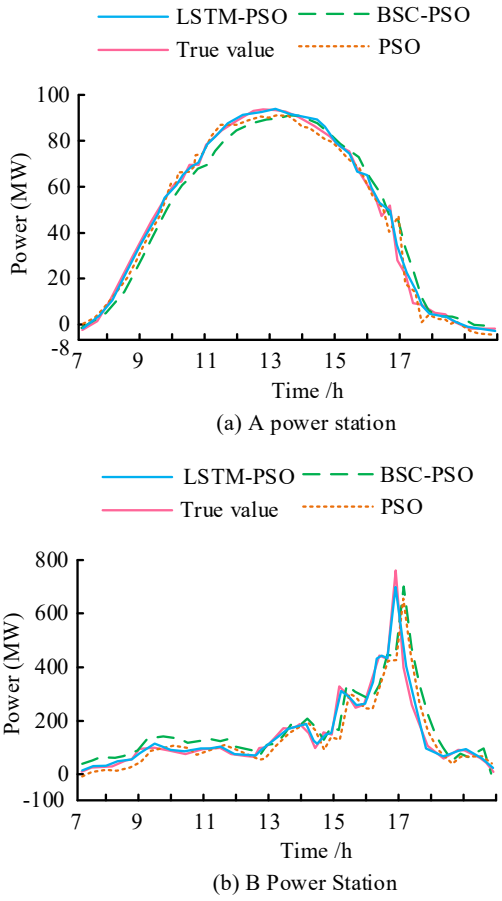


Figure 8 Daily photovoltaic power prediction results

Figure 8 shows the performance of different methods in predicting daily photovoltaic power. In Figure 8 (a), at Power Station A, the predicted value of BSC-PSO is significantly lower than the actual value from 7:00 to 14:00, with a maximum deviation of 10.1 MW. The predicted values of PSO are generally lower than the actual values, with a maximum deviation of 19.8 MW. In contrast, the predicted values of LSTM-PSO are more next to the actual values, with a maximum deviation controlled within 6.2 MW. At power station B (Figure 8b), the predicted values of BSC-PSO are mostly higher than the actual values, especially at peak times where significant deviations occur, with a maximum deviation of 200.3 MW. The predicted value of PSO is close to the actual value most of the time, but there is a significant deviation around 17:00, with a maximum deviation of 299.96MW. The predicted values of the LSTM-PSO algorithm have a small deviation at the peak time, staying within 0.1MW. The comparison of the analysis of prediction residuals for different methods is shown in Figure 9.

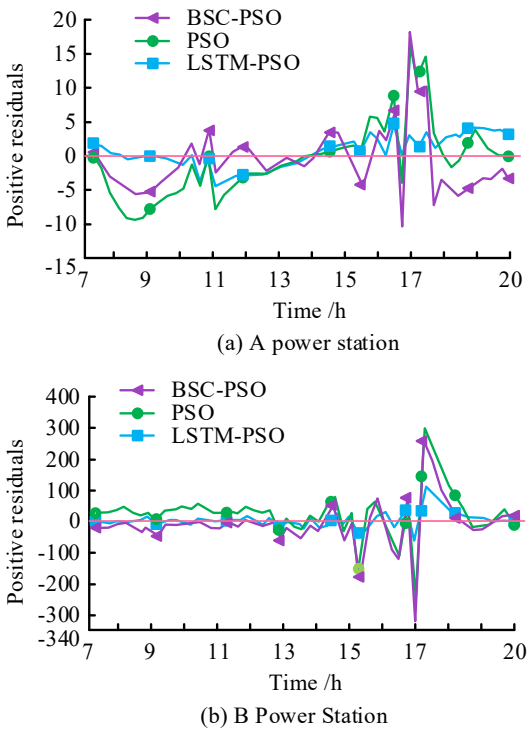


Figure 9 Residual fluctuations of different prediction methods at power station A and power station B

In Figure 9 (a), at power station A, BSC-PSO records the maximum positive residual (18.1MW) at 17:00 and the maximum negative residual (-10.1MW) at 16.7. LSTM-PSO records the maximum positive residual (5.0MW) at time 16.4 and the maximum negative residual (-4.8MW) at time 11.2, showing small fluctuations. In Figure 9 (b), at power station B, the maximum positive residual of BSC-PSO at time 17.2 reaches 259.5 MW, while the maximum negative residual at time 17.0 is -317.2 MW. The maximum positive residual of PSO at time 17.5 is 296.2MW, and the maximum negative residual at time 17.0 is -236.4MW. The maximum positive residual of LSTM-PSO at time 17.5 is 108MW, and the maximum negative residual at time 17.0:00 is -56.2MW, with relatively small fluctuations compared to other methods. Figure 10 analyzes the prediction intervals of different methods at 80% confidence level.

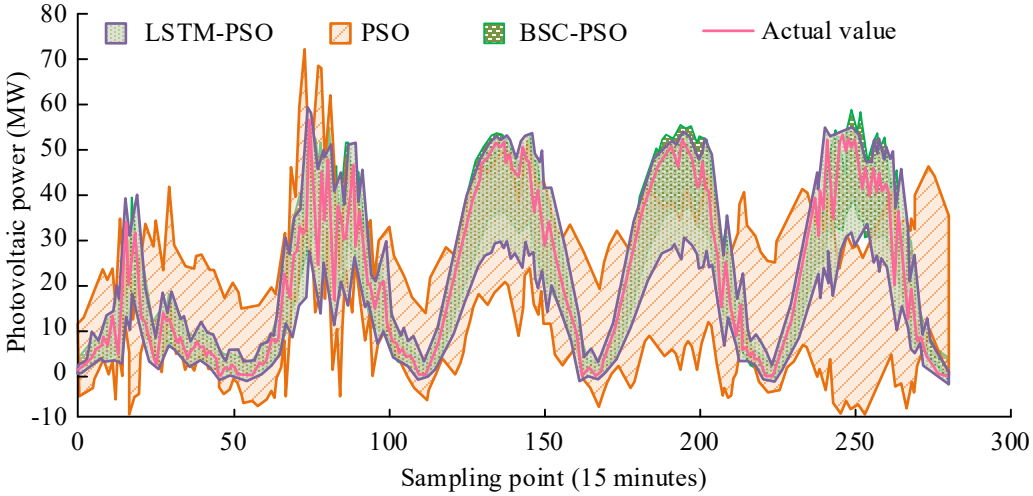


Figure 10 Interval fluctuations of different prediction methods

Figure 10 shows the interval fluctuations of different prediction methods at 300 15-minute sampling points. Although BSC-PSO has small fluctuations in the predicted interval, it fails to cover the actual values at more than 30 sampling points. The prediction interval of PSO is relatively wide, with a maximum difference of 50MW between its upper/lower limits, and the actual values exceed the prediction interval at more than 50 sampling points. In contrast, the difference between the upper/lower limits of the prediction interval of LSTM-PSO is controlled within 30MW, and there is no situation where the actual value exceeds the prediction interval. The model predicted the photovoltaic power generation on weekdays and non-working days, as shown in Figure 11.

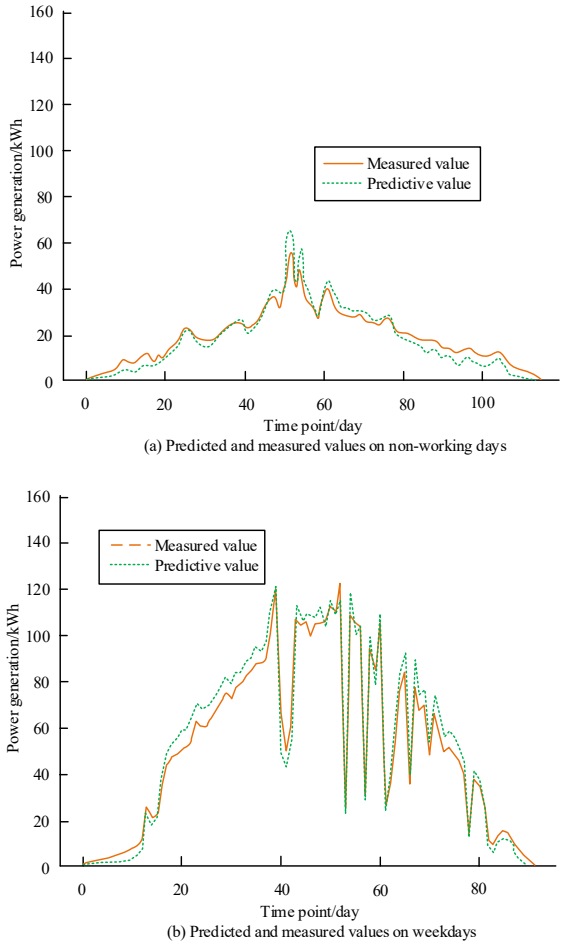


Figure 11 Comparison between model prediction results and actual values

In Figure 11 (a), on non-working days, the model's predicted values closely match the actual value curve, and in some time periods, they are consistent with the actual values. For example, on the 20th day, the predicted power generation is slightly lower than the actual value of 1.17kWh. On the 26th day, both are completely consistent at 22.78 kWh. The predicted value on the 70th day is slightly higher, exceeding the actual value by 1.54kWh. In Figure 11 (b), the predicted curve remains close to the actual value and the overall trend is consistent. On the 37th day, the predicted power generation exceeds the actual value by 9.15 kWh. On the 20th day, the predicted value is 7.42kWh higher than the actual value. The predicted value on the 40th day decreases by 10.61kWh compared to the 20th day.

5. Conclusions

This study proposed a prediction method that combines LSTM and an improved MO-PSO algorithm to handle the challenges of DPPG power prediction. By optimizing the LSTM hyperparameters, this study significantly improved the prediction accuracy. The results indicated that the proposed LSTM-PSO algorithm significantly outperformed other traditional algorithms in terms of prediction accuracy. The specific values showed that the maximum deviation of A power station was controlled within 6.2MW, the predicted deviation at the peak time of B power station remained within 0.1MW, and the prediction interval error was controlled within 30MW at 80% confidence level. Despite significant research achievements, there are still some shortcomings in this study. For example, the predictive ability of the current model for extreme weather conditions still needs to be further improved, and the computational efficiency of the algorithm needs to be optimized for application on large-scale datasets. Future research will extend the model to a wider range of regions and different types of photovoltaic power generation systems, while exploring more efficient algorithm variants to meet larger scale data processing needs.

Acknowledgements

This work was supported by Research and industrialization of key technologies of digital twin based on source-grid-load-storage integration (3502Z20231006).

References

[1] Y. Wang, W. Fu, X. Zhang, Z. Zhen, and F. Wang, "Dynamic directed graph convolution network based ultra-short-term forecasting method of distributed photovoltaic power to enhance the resilience and flexibility of distribution network," IET Generation, Transmission & Distribution, vol. 18, no. 2, pp. 337-352, Sep. 2024. DOI: 10.1049/gtd2.12963.

[2] W. C. Tsai, C. S. Tu, C. M. Hong, and W. Lin, "A review of state-of-the-art and short-term forecasting models for solar PV power generation," Energies, vol. 16, no. 14, pp. 5436-5466, Jul. 2023. DOI: 10.3390/en16145436.

[3] E. M. Al-Ali, Y. Hajji, Y. Said, et al., "Solar energy production forecasting based on a hybrid CNN-LSTM-transformer model," Mathematics, vol. 11, no. 3, pp. 676-698, Jan. 2023. DOI: 10.3390/math11030676.

[4] J. L. J. Pereira, G. A. Oliver, and M. B. Francisco, "A review of multi-objective optimization: methods and algorithms in mechanical engineering problems," Archives of Computational Methods in Engineering, vol. 29, no. 4, pp. 2285-2308, Jun. 2022. DOI: 10.1007/s11831-021-09663-x.

[5] B. Abdollahzadeh and F. S. Gharehchopogh, "A multi-objective optimization algorithm for feature selection problems," Engineering with Computers, vol. 38, suppl. 3, pp. 1845-1863, Aug. 2022. DOI: 10.1007/s00366-021-01369-9.

[6] Y. Xu, C. Xu, H. Zhang, L. Huang, Y. Liu, Y. Nojima, et al., "A multi-population multi-objective evolutionary algorithm based on the contribution of decision variables to objectives for large-scale multi/many-objective optimization," IEEE Transactions on Cybernetics, vol. 53, no. 11, pp. 6998-7007, Nov. 2022. DOI: 10.1109/TCYB.2022.3180214.

[7] L. Hu, Y. Yang, Z. Tang, Y. He, and X. Luo, "FCAN-MOPSO: an improved fuzzy-based graph clustering algorithm for complex networks with multiobjective particle swarm

optimization," IEEE Transactions on Fuzzy Systems, vol. 31, no. 10, pp. 3470-3484, Mar. 2023. DOI: 10.1109/TFUZZ.2023.3259726.

[8] M. Suresh, R. Meenakumari, H. Panchal, V. Priya, E. Agouz, and M. Israr, "An enhanced multiobjective particle swarm optimisation algorithm for optimum utilisation of hybrid renewable energy systems," International Journal of Ambient Energy, vol. 43, no. 1, pp. 2540-2548, Apr. 2022. DOI: 10.1080/01430750.2020.1737837.

[9] P. Kahhal, M. Ghasemi, M. Kashfi, G. Hossein, and J. Kim, "A multi-objective optimization using response surface model coupled with particle swarm algorithm on FSW process parameters," Scientific Reports, vol. 12, no. 1, pp. 2837-2857, Feb. 2022. DOI: 10.1038/s41598-022-06652-3.

[10] I. Rahimi, A. H. Gandomi, F. Chen, and E. Montes, "A review on constraint handling techniques for population-based algorithms: from single-objective to multi-objective optimization," Archives of Computational Methods in Engineering, vol. 30, no. 3, pp. 2181-2209, Apr. 2023. DOI: 10.1007/s11831-022-09859-9.

[11] L. Ge, T. Du, C. Li, Y. Li, J. Yan, and M. Rafiq, "Virtual collection for distributed photovoltaic data: Challenges, methodologies, and applications," Energies, vol. 15, no. 23, pp. 8783-8807, Nov. 2022. DOI: 10.3390/en15238783.

[12] K. J. Iheanetu, "Solar photovoltaic power forecasting: A review," Sustainability, vol. 14, no. 24, pp. 17005-17036, Dec. 2022. DOI: 10.3390/su142417005.

[13] M. N. Akhter, S. Mekhilef, H. Mokhlis, Z. Almohaimeed, A. Muhammad, A. Khairuddin, et al., "An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants," Energies, vol. 15, no. 6, pp. 2243-2264, Mar. 2022. DOI: 10.3390/en15062243.

[14] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," Applied Computing and Informatics, vol. 20, no. 3/4, pp. 231-250, Jun. 2024. DOI: 10.1016/j.aci.2019.11.002.

[15] W. Zhang, X. Chen, and K. He, "Semi-asynchronous personalized federated learning for short-term photovoltaic power forecasting," Digital Communications and Networks, vol. 9, no. 5, pp. 1221-1229, Oct. 2023. DOI: 10.1016/j.dcan.2022.03.022.

[16] S. M. Miraftabzadeh, C. G. Colombo, and M. Longo, "A day-ahead photovoltaic power prediction via transfer learning and deep neural networks," Forecasting, vol. 5, no. 1, pp. 213-228, Feb. 2023. DOI: 10.3390/forecast5010012.

[17] N. Khodadadi, F. S. Gharehchopogh, and S. Mirjalili, "MOAVOA: a new multi-objective artificial vultures optimization algorithm," Neural Computing and Applications, vol. 34, no. 23, pp. 20791-20829, Dec. 2022. DOI: 10.1007/s00521-022-07557-y.

[18] P. Jangir, H. Buch, S. Mirjalili, and P. Manoharan, "MOMPA: Multi-objective marine predator algorithm for solving multi-objective optimization problems," Evolutionary Intelligence, vol. 16, no. 1, pp. 169-195, Feb. 2023. DOI: 10.1007/s12065-021-00649-z.

Graphic abstract: