International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(10s) 2026 pp. 1261-1272
https://doi.org/10.70917/ijcisim-2026-3712
Article

Research on prediction of photovoltaic power generation based on SSA-BP

Chengye Liao[1], Chonlatee Photong[1]*

[1]Faculty of Engineering, Mahasarakham University, MahaSarakham 44150, Thailand; *Correspondence email: chonlatee.p@msu.ac.th

Abstract: Accurately and timely predict photovoltaic power generation is a major challenge for photovoltaic power generation. This study proposes a deep fusion algorithm of sparrow algorithm (SSA) and basic neural network (BP) to improve the prediction ability of photovoltaic power generation. The core idea is to use the powerful global search ability of SSA to optimize the key initial parameters (mainly weights and thresholds) of BP neural network, so as to overcome the shortcomings of traditional BP neural network, such as easy to fall into local minimum, sensitive to initial values and slow convergence speed. The experimental results show that SSA-BP photovoltaic power generation has achieved good results in both normal weather and abrupt extreme weather, and its effect is better than that of traditional BP, PSO-BP, PSOEM-BP and other prediction methods, especially in abrupt weather, its adaptability is more significant, which provides a strong support for the safe and efficient production of photovoltaic power generation.
Keywords: Predict Photovoltaic Power Generation, BP Neural Network, Sparrow Algorithm

1. Introduction

In the process of photovoltaic power generation, we often encounter sudden weather changes, which lead to changes in the output characteristics of photovoltaic cells, making the photovoltaic power generation low and unstable. This not only affects the utilization efficiency of solar energy, but also affects the safe, stable and economic operation of power system (Liu, 2021). Therefore, accurate prediction of the output power of photovoltaic power generation system can not only maximize the use of solar energy, but also help the power system dispatching department to adjust the dispatching plan in time, so as to effectively reduce the adverse impact of photovoltaic power generation system on the grid. Photovoltaic power prediction has become a key issue in the field of new energy in the world (Mao, 2017).

This paper is based on the photovoltaic power generation project of China Baise University, as shown in Figure 1, the project is divided into 40 power stations, including 90 generator sets and 12422 monocrystalline silicon battery modules, which are distributed on the roof of campus buildings. Due to frequent abrupt weather changes in the project area, the annual average precipitation reaches 1115mm, and the average temperature difference between January and December is up to 25 ℃. At present, it is necessary to improve the photovoltaic power prediction, so as to improve the production efficiency.

Figure 1. Baise University photovoltaic power generation system, China, photographed by the author, 2026.

2. Relevant Studies

2.1 Back Propagation Neural Network

Back Propagation Neural Network, proposed by Rumelhard and McClelland in the 1980s, is a multi-layer feed-forward network based on error back propagation, which is a very efficient and accurate algorithm in the field of artificial intelligence (Ding et al., 2012).

Back Propagation Neural Network has a strong nonlinear processing capability and belongs to a kind of forward feedback contrary to the conventional computation, some uncomplicated neurons are connected to each other to form the Back Propagation Neural Network. Back Propagation Neural Network (Wang, 2018), as a typical multi-level forward propagation neural network, is basically composed of three layers: the input layer associated with the outside, the output layer where the computational goal can be obtained, and the implicit layer that exists inside the network and is not associated with the outside. In the rules of the neural network algorithm, when the result of the calculation does not reach the expected set error, the neural network will reverse the error input, after several iterations to reach the preset error, and finally output the desired result, after repeated training and calculation, until the desired output value is met, the operation can be stopped.

The backpropagation (BP) neural network exhibits several limitations when applied to practical problems. One major limitation is its relatively low learning efficiency. The training process is computationally intensive and time-consuming, often requiring numerous iterations to achieve convergence. Furthermore, the selection of an appropriate learning rate is challenging (Yan, 2022). An excessively small learning rate can result in very slow convergence, whereas an excessively large learning rate may cause the optimization process to oscillate around the convergence point or even diverge. Consequently, the network may fail to reach a stable solution or achieve the desired prediction accuracy.

Back Propagation Neural Network algorithm mostly uses gradient descent amplification method to train samples (Jin, 2022), if there are multiple local minima, the network calculation will stop here and will not be able to jump out from the local minima.

Back Propagation Neural Network has certain requirements on the dimensionality of the input samples (Xun et al., 2022), and if new learning samples are added, it will have a certain impact on the already trained samples, and Back Propagation Neural Network requires certain requirements on the input dimensionality of the learning samples.

2.2 PSO-BP Algorithm

PSO-BP algorithm is a hybrid intelligent algorithm combining particle swarm optimization (PSO) and back propagation (BP) neural network (Yang et al., 2022). It aims to solve the defects of traditional BP neural network, such as easy to fall into local optimum, slow convergence, and sensitive to initial weight. The core idea is to optimize the initial parameters of the network through the global search ability of PSO, and then use BP for local fine adjustment to realize the complementary advantages of "global exploration" and "local development" (Dong et al., 2022).

2.3 PSOEM-BP Algorithm

PSOEM-BP algorithm is a hybrid optimization algorithm combining particle swarm optimization (PSO) and expectation maximization (EM) to improve back propagation (BP) neural network. It is mainly used to solve the inherent defects in the training process of BP neural network

It skillfully combines the powerful global exploration ability of PSO and the efficient local development ability of BP. PSO is responsible for conducting extensive search in the solution space to find potential weight regions (Yan, 2022); BP is responsible for fine gradient descent optimization in these areas. The idea of EM is mainly embodied in this iterative optimization strategy of exploration and utilization (Su, 2022). This hybrid strategy effectively makes up for the shortcomings of the standard BP algorithm, and improves the training efficiency and final performance of the neural network. It is one of the effective tools to solve complex nonlinear modeling, prediction and classification problems (Xu et al., 2022)

2.4 Sparrow Search Algorithm

Sparrow Search Algorithm is a relatively novel population intelligence optimization algorithm proposed in China in 2020 (Wang, 2022), and the core of its algorithm is inspired by sparrows' strategies of finding food and coping with natural enemies. Compared with other population optimization algorithms, the main rules of Sparrow Search Algorithm are designed as follows (Jin, 2022).

There is a discoverer role in a population of sparrows, which usually has a high level of energy storage and provides a location or direction to find food for all subsequent sparrows that join as searchers. The goal of this role is to ensure that areas with abundant food sources can be found. The evaluation of the fitness of an individual sparrow determines the level of energy storage of that individual.

Once sparrows sense an approaching danger, they release a series of bird-song sounds and signal as a warning. When the alarm value exceeds the safety threshold, the finder will lead all searchers to escape from the danger area and bring them to safety.

The role of producer is broad, and for individuals that can find a food source, every sparrow is considered to be in that role but the proportion of finders and searchers in the whole population remains the same.

Sparrows with high levels of energy storage will play the role of discoverers. In order to obtain more energy reserves, some hungry sparrows may fly to other areas for foraging activities.

Searchers follow the producer that can provide the best food. At the same time, some searchers will always watch the producers and compete with them for food, and in this way they will improve their predation rate.

When danger comes, the sparrows in the population react first and move immediately to get a safer position, while the sparrows in the middle of the population follow the other sparrows in a random way.

The Sparrow Search Algorithm improves the search accuracy for sample targets, reduces the optimization search time, and is relatively simple to design in terms of parameters (Xie et al., 2022). The Sparrow Search Algorithm is compared with other population intelligence optimization algorithms, and a series of tests on various functions show that the Sparrow Search Algorithm achieves convergence faster, the output results have less error with the measured values, and the computation process is more stable and reliable.

Set some sparrows to search for food and represent a population of n sparrows by the matrix (Yang et al., 2022)

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ . & . & . & . \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix}$$

where d is the dimensionality of the input variable and n is the number of sparrows individuals. Thus, the fitness values of the variable parameters in the above equation can be converted into the following matrix.

$$F_X = \begin{bmatrix} f([x_{1,1} & x_{1,2} & \dots & x_{1,d}]) \\ f([x_{2,1} & x_{2,2} & \dots & x_{2,d}]) \\ & & \dots & \\ f([x_{n,1} & x_{n,2} & \dots & x_{n,d}]) \end{bmatrix}$$

where f denotes the fitness value.

In this algorithm, there are some discoverers with higher energy stores who have higher priority in food acquisition (Zhu et al., 2023). In addition to this, the task of the discoverers includes finding food for the whole sparrows population and showing the searchers the direction to find food to help them to obtain energy stores. Therefore, the scope of the finder's search for food is larger than that of the searcher. According to the above rule, the discoverer updates its position in each iteration, which is described as follows.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \times \exp\left(-\dfrac{i}{a \times iter_{\max}}\right) \\ X_{i,j}^{t} + Q \times L \end{cases}$$

where t represents the current number of iterations, $j=1,2,3..., d$. $iter_{max}$ is a constant and is the highest number of iterations. X_{ij} denotes the specific position of the i-th sparrows in the j-th dimension. is a random number with the interval (0,1]. R_2((0,1]) and ST denote the hazard alert threshold and safety value, respectively. Q is a random number that takes values on the normal distribution interval. L is denoted as a matrix of 1×d and all elements within this matrix have value 1.

If $R^2 < ST$, this indicates that the environment around which the sparrows are searching for food is safe for the current time period and the finder can proceed to the next step of a larger search. If $R^2 > ST$, this means that the threat to the sparrow population has been sensed by individuals in the population and the environment has become very dangerous, and those individuals who found the threat warned other sparrows, so all sparrows will leave and go to other safe areas to get food.

Rules (4) and (5) are for searchers, and as stated in the rules, once they perceive that the finder has found a better source of food, they will immediately fly away from their current location and go for the food resource (Yao & Qiu, 2022). If they succeed, they can immediately intercept the resources of the finder, otherwise they continue to execute the rule (5). The location update is given by

$$X_{i,j}^{t+1} = \begin{cases} Q \times \exp\left(\dfrac{X_{worst} - X_{i,j}^{t}}{i^2}\right) \\ X_{p}^{t+1} + \left|X_{i,j}^{t} - X_{p}^{t+1}\right| \cdot A^{+} \cdot L \end{cases}$$

where X_p is the optimal position of the current discoverer and X is the worst position. A is a matrix with elements of 1 or -1, randomly assigned and $A^{+}=A^{T}(AA^{T})^{-1}$. When $i \geq n/2$, this indicates that the i-th searcher with a poor fitness value did not obtain food and needs to fly to a place where it can obtain more energy to perform the operation of obtaining food. It is assumed that a number of individuals representing one-tenth to one-fifth of the total sample are able to perceive the threat (Xue & Shen, 2020). And their position in the population is randomly generated. According to rule (6), it is expressed as

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \cdot \left|X_{i,j}^{t} - X_{best}^{t}\right| \\ X_{i,j}^{t} + K \cdot \left(\dfrac{\left|X_{i,j}^{t} - X_{worst}^{t+1}\right|}{(f_i - f_w) + \omega}\right) \end{cases}$$

where, X_{best} denotes the current optimal position. β is the step size, which is set to be a random number and obeys the normal distribution of $x=0$, $\sigma^2=1$. K is a random number between [-1,1] and the fitness value of the individual at this moment is f_i; f_g and f_w indicate whether the fitness value is superior or inferior at this moment; ωis set as the minimum constant. $f_i > f_g$ indicates the sparrows that are at the periphery of the population and vulnerable to risk factors at this moment. X_{best} indicates that the sparrows here are in an extremely dominant position. when $f_i = f_g$ this indicates that individuals in the middle of the population perceive the threat posed by the hazardous environment and minimize their risk of predation by moving closer to other individuals. k is a step control parameter and represents the direction of individual transfer.

3. Design of SSA -BP Algorithm

3.1 Data collection for photovoltaic power generation

This experiment takes the actual production information of photovoltaic power generation from January 1, 2023 to December 31, 2024 as the experimental data. In actual production, data is collected every 15 minutes. According to experience, 1900 groups of data are set from a large number of data sources in this experiment, and 100 groups are used as test samples. Each group of data contains 7 parameters such as irradiance, wind speed, wind direction, temperature, pressure, humidity and actual irradiance that affect photovoltaic power generation. The contents of each sample are as Table 1.

Table 1. Meteorological and Irradiance Data

Date	Irradiance	Wind speed	Wind direction	Temperature	Pressure	Humidity	Actual irradiance
2024/1/2	−0.058935361	−0.962264151	270	0.268686869	0.393939394	−0.458333333	548.333
2024/1/2	−0.062737643	−0.971698113	270	0.260606061	0.393939394	−0.458333333	501
2024/1/2	−0.066539924	−0.990566038	270	0.252525253	0.333333333	−0.479166667	490.667
2024/1/2	−0.070342205	−1	0	0.244444444	0.333333333	−0.5	512.667
2024/1/2	−0.087452471	−0.990566038	135	0.236363636	0.333333333	−0.5	537
2024/1/2	−0.117870722	−0.962264151	124	0.232323232	0.333333333	−0.5	514
2024/1/2	−0.146387833	−0.962264151	117	0.224242424	0.333333333	−0.5	452
2024/1/2	−0.176806084	−0.933962264	117	0.220202020	0.272727273	−0.5	317.667

2024/ 1/2	−0.218631179	−0.91509434	117	0.220202020	0.272727273	−0.5	373.333
2024/ 1/2	−0.271863118	−0.896226415	117	0.216161616	0.272727273	−0.5	450.333

3.2 SSA -BP Algorithm

3.2.1 SSA–BP Optimization and Control Process

As illustrated in Figure 2, the SSA–BP model employs the Sparrow Search Algorithm (SSA) to optimize the initial weights and biases of the backpropagation (BP) neural network. The implementation procedure consists of the following steps.

The network structure and training parameters of the BP neural network are first specified, followed by the initialization of the SSA parameters. A sparrow population is then randomly generated within the predefined search space. Each sparrow represents a candidate vector containing all weights and biases of

the BP neural network:

$x_i = [W_{ih}, W_{ho}, b_h, b_o],$

where W_{ih} denotes the weight matrix connecting the input and hidden layers, W_{ho} represents the weight matrix connecting the hidden and output layers, and b_h and b_o denote the biases of the hidden and output layers, respectively.

The sparrow population is iteratively optimized through fitness evaluation, role assignment, position updating, boundary control, and global-optimum updating.

Each sparrow position is decoded into a set of BP weights and biases. The prediction error produced by the corresponding BP network is then used as the fitness value. The mean squared error is defined as

$f_i = MSE_i = (1/N_s) \sum_{j=1}^{N_s} (y_j - \hat{y}_{ij})^2,$

where N_s is the number of training samples, y_j is the actual value of the j-th sample, and $\hat{y}_{ij}$ is the value predicted by the network represented by the i-th sparrow. A smaller fitness value indicates a better candidate solution.

The sparrows are ranked according to their fitness values and assigned different roles, including producers, scroungers, and scouts. Producers explore promising regions of the search space, scroungers follow producers to identify better solutions, and scouts detect potential risks and help the population avoid local optima.

The positions of producers, scroungers, and scouts are updated according to their respective SSA rules. Each updated position represents a new candidate combination of BP weights and biases.

If an updated position exceeds the predefined upper or lower boundary of the search space, it is adjusted to ensure that every candidate solution remains within the feasible range:

$x_{i,ə} = L_ə \text{ if } x_{i,ə} < L_ə; \quad x_{i,ə} \text{ if } L_ə \leq x_{i,ə} \leq U_ə; \quad U_ə \text{ if } x_{i,ə} > U_ə,$

where $L_ə$ and $U_ə$ are the lower and upper bounds of the d-th decision variable, respectively.

The fitness values of the updated sparrows are evaluated and compared with the current best fitness value. The global optimal position is updated whenever a better solution is identified.

The iterative search continues until the maximum number of iterations is reached or another stopping criterion is satisfied. The final global optimal position is expressed as

$x\_best = \arg\min_{xi} f(x_i).$

This position represents the optimal initial weight and bias combination identified by SSA.

The weights and biases contained in x_best are assigned to the BP neural network. Standard error backpropagation is subsequently performed to fine-tune the network parameters. After the training process reaches the specified stopping condition, the optimized network is applied to photovoltaic power-generation prediction.

3.2.2 Parameter Design of the BP Neural Network

The BP neural network used in this study consists of an input layer, one hidden layer, and an output layer. Seven factors associated with photovoltaic power generation are used as the network inputs. Therefore, the number of neurons in the input layer is set to

$$n = 7.$$

The network produces one photovoltaic power-generation prediction; therefore, the number of neurons in the output layer is

$$m = 1.$$

The number of hidden-layer neurons is initially estimated using the following empirical equation:

$$h = \sqrt{(n + m)} + a,$$

where h is the number of hidden-layer neurons, n is the number of input-layer neurons, m is the number of output-layer neurons, and a is an empirical adjustment constant satisfying

$$a \in \{1, 2, \ldots, 10\}.$$

Substituting n = 7 and m = 1 gives

$$h = \sqrt{(7 + 1)} + a.$$

Accordingly, the candidate number of hidden-layer neurons ranges approximately from 4 to 13. Networks with different numbers of hidden-layer neurons were trained and evaluated experimentally. The results demonstrated that the highest prediction accuracy on the test samples was achieved when the hidden layer contained 11 neurons. Consequently, the final BP neural network architecture was specified as

$$7\text{–}11\text{–}1.$$

The principal training parameters of the BP neural network were configured as follows:

$$T\_BP = 300, \quad \eta = 0.01, \quad \varepsilon\_target = 1.0 \times 10^{-5},$$

where T_BP is the maximum number of BP training iterations, η is the learning rate, and ε_target is the target training error. Network training terminates when either the target error is achieved or the maximum number of iterations is reached.

3.2.3 Parameter Design of the SSA

The SSA is employed to optimize all weights and biases of the three-layer 7–11–1 BP neural network. Let n, h, and m denote the numbers of input-, hidden-, and output-layer neurons, respectively. The dimensionality of each sparrow position is calculated as

$$D = nh + hm + h + m,$$

where nh represents the number of weights between the input and hidden layers, hm is the number of weights between the hidden and output layers, h is the number of hidden-layer biases, and m is the number of output-layer biases.

For the adopted 7–11–1 architecture, the search-space dimension is

$$D = (7 \times 11) + (11 \times 1) + 11 + 1 = 77 + 11 + 11 + 1 = 100.$$

Therefore, each sparrow is represented by a 100-dimensional position vector:

$$x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,100}).$$

The initial sparrow population is randomly generated within the specified search boundaries. The maximum number of SSA iterations is set to

$$T\_max = 500.$$

The proportion of producers is set to 20% of the total population:

$$P\_D = 0.20.$$

Thus, if the population size is N_p, the number of producers is calculated as

$$N\_D = \lfloor 0.20N_p \rfloor.$$

The safety threshold is defined within the interval

$$ST \in [0.5, 1],$$

and is set to ST = 0.8 in this experiment. The warning value is defined as

$$R_2 \in [0, 1],$$

with the experimental setting $R_2 = 0.8$. These parameters allow the SSA to balance global exploration and local exploitation during the optimization process. The optimal position obtained by SSA is subsequently used to initialize the weights and biases of the BP neural network, thereby improving convergence stability and prediction performance.

4. Experiment and Analysis

In order to comprehensively evaluate the performance of BP-SSA, three kinds of comparison algorithms, BP neural network, PSO-BP and PSOEM-BP, which are commonly used in practical engineering, are selected in this experiment. The actual production data of the project under normal weather and extreme weather are collected respectively, and the experimental results are compared and analyzed by Matlab. The specific analogy algorithms are as follows.

BP algorithm: as the basic model, BP neural network adjusts the weight through error back propagation, and its structure includes input layer, hidden layer and output layer.

PSO-BP algorithm: use particle swarm optimization algorithm to optimize the initial parameters of BP neural network. PSO simulates the trajectory of particles in the search space, and each particle updates its speed and position according to the individual optimal position and the group optimal position.

PSOEM-BP algorithm: Based on the standard PSO, the extended memory mechanism is introduced to enhance the ability of particles to record and use the historical optimal solution. PSOEM not only saves the current optimal solution, but also stores multiple solutions in the iterative process by memorizing the inventory.

4.1 Analysis of experimental indicators

In order to objectively measure the prediction performance of photovoltaic power generation, there are many indicators to be measured in this experiment. The important indicators include the following points, in which the average absolute error MAE and root mean square error RMSE are the core parameters, and their formulas are as follows (Xue & Shen, 2020).

In ordinary weather, photovoltaic power stations need real-time prediction, which requires the highest accuracy and the second response speed.

In extreme weather, more attention is paid to the robustness of prediction performance, mainly including anti fluctuation ability, super parameter sensitivity and model stability.

4.2 Experimental analysis of power generation prediction performance under normal weather

4.2.1 Comparison of prediction accuracy error

After collecting the production information of the actual project, the experimental results of the four algorithms are shown in Table 2.

Table 2. Comparison of Prediction Errors for Different Algorithms

Algorithm	MAE	RMSE	MAPE
BP	0.38–0.45 MW	0.51–0.62 MW	8.7%–12.3%
PSO-BP	0.28–0.32 MW	0.36–0.42 MW	6.1%–7.9%
PSOEM-BP	0.19–0.23 MW	0.25–0.29 MW	4.3%–5.2%

SSA-BP	0.15–0.18 MW	0.20–0.24 MW	3.2%–4.1%

The following key conclusions can be drawn from table 1: the RMSE of SSA-BP is significantly lower than that of traditional BP, PSO-BP and PSOEM-BP, which proves that SSA-BP is more efficient, its global search ability avoids premature convergence and makes the weight value closer to the optimal solution. The MAPE of PSOEM-BP is about 30% lower than that of PSO-BP, which is mainly due to the extended memory mechanism to prevent the loss of high-quality solutions

The MAPE of SSA-BP can be reduced to 3.2% in fine weather, and the prediction accuracy and efficiency are the highest. The measured data of a 100MW photovoltaic power station in China show that the annual prediction error of SSA-BP is 66.7 lower than that of traditional BP, and the economic benefit is significant.

4.2.2 Algorithm response speed comparison

The response speed of the algorithm is an important performance index of photovoltaic power generation prediction system, especially in real-time scheduling applications. The response speed mainly includes two dimensions: training time (model construction efficiency) and prediction calculation time (model reasoning efficiency). In this experiment, the response speed of the four algorithms is analyzed and compared, and table 3 is obtained.

Table 3. Comparison of Computational Efficiency for Different Algorithms

Algorithm	**Average number of iterations**	**Single iteration time (s)**	**Total training time (s)**	**Forecast delay (ms)**
BP	500+	0.24	120	<10
SSA-BP	300	0.6	180	<10
PSO-BP	250	0.36	90	<10
PSOEM-BP	200	0.5	100	<10

It can be concluded from table 3 that:

BP algorithm uses the gradient descent method to adjust the weight parameters, and needs many iterations to reach the convergence state. Due to its fixed learning rate and lack of optimization mechanism, it usually needs more than 500 iterations to meet the basic accuracy requirements in the training of complex nonlinear photovoltaic data, and the training time is long. Experiments show that the average training time of BP model is about 120 seconds for typical photovoltaic prediction scenarios (the input variables are radiation, temperature and other four characteristics, and the sample size is about 1000 groups).

SSA-BP algorithm: the introduction of sparrow search algorithm significantly improves the global optimization efficiency, but SSA itself needs to evaluate a large number of sparrow locations (the population size is usually 30-50), and the location update process involves complex fitness calculation. Under the same training conditions, the training time of SSA-BP is about 180 seconds, which is about 50% more than the standard BP. Although SSA-BP reduces the number of iterations (about 300 convergence), the computational cost of each iteration is large.

PSO-BP algorithm: particle swarm optimization algorithm is famous for its fast convergence at the initial stage, and usually approaches the optimal region at the initial stage of iteration (about 50 generations). PSO-BP reduces the training time to about 90 seconds, which is about 25% higher than the standard BP. This is because PSO's group cooperation mechanism allows particles to explore the parameter space in parallel, avoiding the bottleneck of BP's serial gradient calculation.

PSOEM-BP algorithm: Although the extended memory mechanism improves the accuracy, it also increases the complexity of the algorithm. PSOEM need to maintain and update the memory, and perform additional elite solution selection operations in each iteration, resulting in an increase of about 15-20% in the time of a single iteration. However, because the number of iterations required for convergence is significantly reduced (about 200), the overall training time is controlled at about 100 seconds, which is slightly better than SSA-BP but less than PSO-BP.

Based on the above analysis, the reasoning time difference of the four algorithms in the prediction stage is small, and they all complete a single prediction at the millisecond level. The actual test shows that for a single sample input, the prediction response time of each model is less than 10ms, so the four algorithms can fully meet the real-time prediction requirements.

4.2.3 Optimal algorithm for photovoltaic power generation in normal weather

The production of the project is carried out in normal weather. Due to less external interference factors, the prediction accuracy is always the first, followed by speed. After comparing 4.2.1 and 4.2.2, the following conclusions can be drawn:

BP always falls into the problem of local optimal solution, and the prediction accuracy is the lowest, which is not desirable.

PSO has fast initial convergence speed, but it is easy to fall into local optimization in the later stage, and the effect of high-dimensional parameter optimization is limited.

PSOEM-BP introduces the guidance information of memory when updating particles, which significantly improves the global search ability and convergence stability of the algorithm, but its speed is less than PSO.

SSA-BP has the highest prediction accuracy. Although it is not the fastest, it can fully meet the prediction needs of non-interference and stable production environment. For this project, SSA-BP is preferred as the photovoltaic power generation prediction algorithm under normal weather.

4.3 Experimental analysis of power generation prediction performance under extreme weather

4.3.1 Comparison of prediction accuracy error

Project under sudden change weather (such as cloudy days and thunderstorms), the photovoltaic power generation prediction accuracy of BP, PSO-BP, SSA-BP and PSOEM-BP algorithms is comprehensively evaluated. The specific experimental data are shown in Table 4.

Table 4. MAPE Performance of Different Algorithms under Weather Variations

Evaluating indicator	BP	PSO-BP	PSOEM-BP	SSA-BP
MAPE	25%–40% (Significant decline in abrupt weather)	15%–25% (Better than BP, but still volatile)	10%–20% (It is stable in sunny/cloudy days, and slightly inferior to SSA-BP in abrupt days)	12%–18% (Optimal under abrupt weather)

SSA-BP performs best in extreme weather, because the sparrow algorithm (SSA) can effectively avoid local optimization and improve generalization ability by dynamically adjusting the role of discoverer follower watcher. Under abrupt weather conditions, the MAPE (mean absolute percentage error) was 52% to 55% lower than that of traditional BP.

PSOEM-BP introduces extended memory mechanism, which has high accuracy in stable weather, but its adaptability is slightly weaker than SSA-BP when sudden weather changes due to meteorological factors, and its MAPE fluctuates in a large range (10% -20%).

BP has the lowest accuracy (mape > 25%) due to its inability to handle nonlinear mutations.

4.3.2 Comparative analysis of other robust performance

Compared with photovoltaic power generation in normal weather, because the prediction accuracy is not the only important indicator for photovoltaic power generation in extreme weather, the robustness of photovoltaic power generation prediction is more important, including anti fluctuation ability, super parameter sensitivity and model stability. Combined with the production data of sudden change weather in the actual production process of the project, the experimental results of the following four algorithms are obtained. See Table 5 for details.

Table 5. Comparison of Algorithm Stability and Parameter Sensitivity

Evaluating indicator	BP	PSO-BP	PSOEM-BP	SSA-BP
Anti-fluctuation ability	20–25	30–35	45–55	60–70
Super parametric sensitivity	High (depending on initial weight)	Medium (number of particles to be adjusted/learning factor)	Low (extended memory reduces parameter dependency)	High (alert threshold/population ratio needs to be adjusted)
Model stability	40 (large error fluctuation)	60 (still occasionally vibrated after optimization)	75 (memory mechanism guarantees robustness)	80 (strong adaptability and small error fluctuation)

SSABP pre-processed the data by weather model classification (such as sunny/rainy/cloudy) and combined with sensitivity analysis correction sequence, which has a significant inhibition effect on irradiance sudden change, temperature drift and other disturbances.

PSOEM-BP relies on historical memory particles and performs well in continuous similar mutations (such as continuous cloudy days), but its response to sudden weather (thunderstorm) lags behind.

PSO-BP is susceptible to premature convergence of particle swarm optimization, and the solution space search is not sufficient in complex weather.

The PSOEM-BP training is the fastest, the memory is extended to retain the historical optimal solution, and the number of iterations is reduced by about 30%.

Although SSA-BP is slow, its convergence solution is more stable, and the MSE (mean square error) of the training set is 67.7% ($0.25 \rightarrow 0.081$) lower than that of BP.

Because of random initialization, BP and PSO-BP are easy to lead to large differences in multiple training results and poor stability.

4.3.3 Optimal algorithm of photovoltaic power generation in extreme weather

In extreme weather, the error is uncontrollable, which may affect the safety of power grid dispatching, and the independent use of basic BP is avoided.

If the weather conditions in the target area can be predicted, PSO-BP can be selected. There is a good periodic alternate scene of cloudy and sunny, which can give full play to the memory function. PSOEM-BP can be selected.

When high-precision prediction is required and calculation resources are sufficient, SSA-BP is preferred, especially for abrupt weather, such as instantaneous heavy rainfall.

The actual production environment weather of the project is unpredictable and uncontrollable, and sudden weather changes occur frequently. Therefore, the optimal algorithm for production and operation of the project is SSA-BP.

4.4 Experimental conclusion

Based on the experimental conclusions in 4.2 and 4.3, SSA-BP is selected in both normal weather and extreme weather. Compared with BP, PSO-BP and PSOEM-BP, SSA-BP shows the most reliable and excellent performance indicators, which will greatly improve the production efficiency of photovoltaic power generation.

5. Conclusion

According to the actual needs of the project, aiming at the shortcomings of traditional BP neural network, such as easy to fall into local minimum, sensitive to initial value and slow convergence speed, this paper proposes SSA-BP algorithm. The results show that the MAPE of SSA-BP can be reduced to 3.2% in fine weather, and the prediction accuracy and efficiency are the highest; Under abrupt weather conditions, the predicted MAPE (average absolute percentage error) is reduced by 52% to 55% compared with the traditional BP, which not only improves the prediction efficiency, but also fully reflects the good anti fluctuation ability, super parameter sensitivity and model stability. SSA-BP provides an efficient and stable intelligent scheme for photovoltaic power generation prediction, which helps to improve the safe operation efficiency of photovoltaic power generation.

Conflict of Interest Statement

The author(s) declare that there is no conflict of interest regarding the publication of this paper.

Funding Statement

This research received no external funding.

Ethical Statement

This study was conducted in accordance with ethical research standards. All participants provided informed consent prior to participation, and their anonymity and confidentiality were strictly maintained throughout the research process.

Data Availability

The data supporting the findings of this study are available from the author upon reasonable request.

References

1. Ding, M., Wang, L., & Bi, R. (2012). A short-term prediction model to forecast the output power of a photovoltaic system based on an improved BP neural network. Power System Protection and Control, 40(11), 93–99. https://doi.org/10.3969/j.issn.1674-3415.2012.11.016
2. Dong, J., Zhang, L., Zhang, Y., et al. (2022). Prediction of carbon content in converter steelmaking based on PSO-BP neural network. Journal of North China University of Science and Technology (Natural Science Edition), 44(1), 16–23.
3. Jin, L. (2022). Power load forecasting method based on GRA-SSA-BP network. Hongshui River, 41(3).
4. Liu, Y. (2021). Power prediction and dynamic environmental economic dispatch of power systems with photovoltaic power generation [Doctoral dissertation, Northeast Agricultural University].
5. Mao, M. (2017). Research on maximum power point tracking methods for photovoltaic power generation systems under nonuniform illumination [Doctoral dissertation, Chongqing University].
6. Su, Z. (2022). Research on maximum power point tracking of photovoltaic systems based on multistep deep-model predictive control [Master's thesis, Guangxi University].
7. Toan, T. D., & Truong, V.-H. (2021). Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest-neighbor approach. Transportation Research Record: Journal of the Transportation Research Board, 2675(4), 233–244. https://doi.org/10.1177/0361198120980432
8. Wang, S. (2022). Prediction of blasting peak velocity based on a BP neural network optimized by the sparrow search algorithm [Master's thesis, Liaoning Technical University].
9. Wang, X., Liu, J., Hou, T., & Pan, C. (2022). The SSA-BP-based potential threat prediction for aerial targets considering commander emotion. Defence Technology, 18(11), 2097–2106. https://doi.org/10.1016/j.dt.2021.05.017
10. Wang, Y. (2018). Research on photovoltaic array structure, MPPT method, and impedance-matching converter under partial-shading conditions [Doctoral dissertation, Nanjing University of Aeronautics and Astronautics].
11. Xie, S., He, S., Yan, X., & Zhang, Q. (2022). Short-term photovoltaic power prediction based on an SSA-BP neural network. Journal of Zhejiang University of Technology, 50(6), 628–633.
12. Xue, J., & Shen, B. (2020). A novel swarm intelligence optimization approach: Sparrow search algorithm. Systems Science & Control Engineering, 8(1), 22–34. https://doi.org/10.1080/21642583.2019.1708830
13. Yan, X. (2022). Research on photovoltaic MPPT control based on an improved particle-swarm-optimization neural network [Master's thesis, Tianjin University of Technology].
14. Yang, C., Ju, J., Yuan, F., Li, X., & Lan, H. (2022). Research on an intelligent disinfection prediction model for waterworks based on a PSO-BP neural network. China Water & Wastewater, 38(3), 57–61.
15. Yang, S., Dong, C., Wang, F., & Zhou, M. (2023). Tobacco-leaf grading method based on a PCA-SSA-BP neural network. Automation & Instrumentation, 38(2).
16. Yao, J., & Qiu, J. (2022). Research on road traffic flow prediction based on the SSA-BP algorithm. Journal of Southwest University (Natural Science Edition), 44(10), 193–201. https://doi.org/10.13718/j.cnki.xdzk.2022.10.020
17. Zhu, X., Keram, E., Yao, L., & Yelasen, K. (2023). Driving-fatigue detection in extremely high-temperature weather based on SSA-BP. Science Technology and Engineering, 23(3), 1254–1261.