Prediction of Photovoltaic Power Based on SSA-BP Algorithm with Chaotic Mapping

Zihan Li*

Shanghai DianJi University, university of business studies, Shanghai, China

Abstract: Photovoltaic (PV) power generation is of great significance to the energy reform in China. In order to better grid-connect PV, maximize the utilization of light resources, and reduce the adverse effects on the power grid due to the uncertainty and randomness of light energy, we need to study the power generation of PV. This paper discusses the factors affecting the efficiency of PV power generation and the related prediction algorithms in the light of current literature and further uses software to optimize the algorithm for PV prediction.

1.Introduction

In the context of China's commitment to the national "dual-carbon" goal, solar energy has entered a new phase of development, showcasing advancements in key technologies and promising application prospects. Solar energy, as a form of renewable energy, offers a wide range of applications beyond photovoltaic power generation and solar thermal power generation. It can be harnessed through photochemical reactions, light-induced processes, and photobiological conversions. Currently, photovoltaic power generation accounts for 99% of total solar power generation, making it a major solar energy technology. Distributed photovoltaic power generation, akin to traditional energy sources, is an energy-saving, environmentally friendly, and clean energy solution. It is being actively pursued and implemented in rural areas, pastures, unused roofs of public buildings, industrial parks, and open plateaus. On March 9, 2022, the National Energy Administration provided an update on the progress of PV power generation projects in 2021. The data revealed that China's newly installed PV capacity reached 54.8800 GW in 2021, with centralized PV accounting for 25.6007 GW and decentralized PV accounting for 29.2790 GW. This trend of decentralized PV is expected to continue growing in the future¹.

Solar energy, as a clean, safe, efficient, and sustainable new energy source, has garnered global attention. However, due to its inherent characteristics, such as intermittency, fluctuations, and randomness, integrating solar energy into the power grid can significantly impact its stability. Therefore, accurately predicting PV power generation is crucial to ensure the safety, stability, and efficient scheduling of the power grid.

1.1. Analysis of photovoltaic power influencing factors

Factors influencing photovoltaic power generation can be broadly categorized into two groups²: those affecting energy transfer and those affecting energy conversion. The orientation of the sun plays a role, as the angle of sunlight varies throughout the day, with the highest energy levels reached at noon when the sun shines perpendicularly. Cloud cover also affects PV power generation, as higher clouds result in greater shading on the ground. Atmospheric quality is another factor, with poorer air quality leading to increased energy loss due to carbon dioxide and other particles. Ambient temperature influences photovoltaic modules, with power generation efficiency increasing within a certain temperature range. However, beyond that range, efficiency tends to saturate and may even decline. The temperature of the PV panels themselves also impacts energy conversion efficiency, with higher temperatures leading to lower efficiency. Additionally, the accumulation of dirt on PV panels can produce corrosive substances when combined with water vapor in the air, damaging the photovoltaic devices and reducing power generation efficiency.

1.2.Forecasting methods

Depending on the type of prediction³, it can be categorized into deterministic and probabilistic prediction. The former refers to the prediction of the PV power at a certain point in time, while the latter is an estimate of the PV power, which not only estimates the range of power fluctuations, but also provides an effective prediction interval.

In the realm of solar energy prediction, two main approaches can be employed: physical prediction and statistical prediction. The physical prediction model relies

*Corresponding author: 1580641907@qq.com

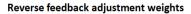
enhance its accuracy and reliability. However, this approach may not be suitable for newly established power plants that lack extensive historical data⁴. Specific classification can be divided into: (1)

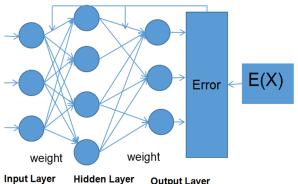
modeling method⁵: including regression analysis, time series method, Kalman filter method, wavelet analysis method. This type of prediction is characterized by easy calculation, fast convergence, and requires less historical data. (2) Artificial Intelligence Method: It includes neural network prediction method, gray system prediction method, fuzzy logic method, expert system prediction method. The characteristics of this type of method: high prediction accuracy and great adaptive ability. (3) Comprehensive method: It is also known as the combination method, which is mainly to synthesize the modeling method and artificial intelligence method and absorb their advantages in order to improve the final prediction legree, and the key to its application lies in which method is used for the synthesis. The optimal combination forecasting method is a method of synthesizing various forecasting methods according to the weights of the various forecasting methods in the forecasting results.

Optimal combination forecasting contains two concepts: one is a forecasting method that selects appropriate weights for weighted average of the forecast results obtained by several forecasting methods; the other is a comparison between several forecasting methods, and selects the forecasting model with the best fit or the smallest standard deviation as the optimal model to make forecasts⁶.

2.Basic principles of BP neural networks

The BP neural network algorithm is an algorithm modeled after biological neurons. The algorithm has three layers including input layer, hidden layer and output layer. The topology is shown in Fig 1:





nput Layer Hidden Layer Output Layer Fig.1. The Topological Structure of Neural Networks

The specific algorithmic process of the BP neural network can be described as follows in a nature language style: The algorithm starts with forward propagation and backward propagation. In forward propagation, a series of inputs X in the input layer are multiplied by different weights W and added with bias b. These values are then passed to the hidden layer, where they undergo a transformation using the activation function f(x). The transformed values are then passed to the output layer, where they are again multiplied by weights W and passed through the activation function f to obtain the final output result. The output result is compared with the expected value to evaluate the fitting effect. If satisfactory results are obtained, the model is tested and applied. However, if the results are not satisfactory, the algorithm proceeds with back propagation. This involves using the gradient descent method and performing multiple iterations to minimize the "error value = (output value - predicted value) 2 ". The weights W are adjusted accordingly, and the process of forward propagation is repeated. The implementation of the BP neural network algorithm in MATLAB can be roughly divided into the following programming steps⁷:

1. Data reading: Randomly generate a series of arrays and divide them into a test set and a training set.

- 2. Build the structure of the BP neural network.
- 3. Train the network using the training set.

4. Use the trained network structure to simulate and obtain predicted values using the test set.

5. Calculate the error, which is the difference between the predicted values and the true values.

Due to the highly nonlinear relationship involved in PV prediction, the choice of the generating unit and the climate environment significantly impacts its accuracy. Therefore, the use of a BP neural network for PV power prediction is considered more appropriate. However, relying solely on the BP neural network prediction method has some noticeable drawbacks. When faced with complex problems, this method becomes computationally inefficient⁸. This inefficiency arises from the computational approach of gradient descent, which involves sampling the inputs and outputs and transforming them into a nonlinear form. The solution is then obtained by continuously minimizing the error through weight updates. To achieve this, the network needs to learn from

the samples and minimize the error as much as possible. Another reason for the algorithm's inefficiency is the complex distribution of the optimal solution surface. The BP neural network algorithm essentially adjusts the weights using the gradient descent method, making it challenging to quickly reach the minimum value. In severe cases, the algorithm may converge to a local optimal solution rather than the global optimal solution. This occurs because the error between the predicted and true values may converge around a certain value. The convergence and divergence of the BP neural network are influenced by the number of hidden layers and nodes in its structure. Increasing the number of hidden layers enhances its ability for nonlinear mapping. However, this also leads to a larger number of weight calculations, reducing computational efficiency and ultimately diminishing utility. On the other hand, if the number of nodes is too small, it can decrease computation accuracy, resulting in unreliable final results.

3.Fundamentals of Sparrow Search Algorithm

Sparrow search algorithm is a new population intelligence optimization algorithm. Compared with other population intelligence algorithms, SSA is a new population intelligence optimization algorithm superior to Gray Wolf Optimization Algorithm (GWO), Particle Swarm Optimization (PSO), Gravity Search Algorithm (GSA) and other algorithms. The central idea of population intelligence optimization algorithm is to find the optimal solution of the solution space distributed in a certain range by simulating the movement and behavioral laws of certain things or organisms in nature

3.1.Bionicist principle

There is an initial population of sparrows, which is then divided into finders and predators according to their position in the population and their energy level in the "social division of labor". The former are the ones that need to find food, and the latter are the ones that catch prey based on the traces of food found by the former and some biological information they leave behind. The higherenergy ones are used as discoverers, and they are able to fly higher and find food more easily. And the remaining ones with lower energy follow the discoverers as predators. In addition to this, the sparrow population is supposed to have the ability to detect natural enemies in order to better survive in nature, so another percentage of the sparrow population has to act as perceivers to sense natural enemies. When it is threatened by a natural enemy or realizes the danger, it will alert the whole population. When the warning value is greater than the safety value, the producer leads the population to migrate to other feasible safe areas, and the sparrows at the edge of the population move faster towards the safe area, while the sparrows in the middle area of the population walk randomly to follow the population. The mathematical model of the algorithm is as follows:

3.2. Mathematical Modeling of Algorithms

Discovery's location was updated as follows

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^t \cdot \exp(\frac{-i}{\alpha \cdot iter_{\max}}) & R_2 < ST \\ x_{i,j}^t + Q \cdot L & R_2 \ge ST \end{cases}$$
(1)

where Xti,j is the position of the Ith sparrow in dimension j at the Tth iteration; α is a random number between (0,1]; Q is a random number obeying a normal distribution; and L is a 1 × d matrix with element 1.

Predator locations were updated as follows

$$s_{i,j}^{t+1} = \begin{cases} s_b^{t+1} + |s_{i,j}^t - s_b^{t+1}| \cdot A^+ \cdot L & i \le \frac{n}{2} \\ Q \cdot \exp(\frac{s_w^t - s_{i,j}^t}{i^2}) & i > \frac{n}{2} \end{cases}$$
(2)

A 10% to 20% proportion of perceivers was randomly selected in the initial population. When natural enemies and danger are sensed, sparrows at the edge of the population fly toward the most adapted spotter, and sparrows in the middle of the population move closer to others to avoid predation. Updated position of perceivers:

$$s_{i,j}^{t+1} = \begin{cases} s_b^t + w | s_{i,j}^t - s_b^t | & F_i > F_b \\ s_{i,j}^t + p \cdot (\frac{s_{i,j}^t - s_w^t}{F_i - F_w + \varepsilon}) & F_i = F_b \end{cases}$$
(3)

3.3.Basic program steps

the first step initializes the population by randomly distributing individual sparrows over the search interval. Then record the historical sparrow positions and new sparrow positions. Immediately after that, the optimal fitness and optimal solution of the initial population are calculated. Once the preparation is done, the population evolution phase begins, first the population fitness is sorted in order to have the high fitness as discoverers and the others as predators, and then the positions are updated by the code of the three positional formulas above.

4. Chaotic mapping algorithm

Chaotic mappings typically consist of one or more nonlinear equations that capture the evolutionary dynamics of a system. Through iterative application of these equations, the system undergoes state changes and displays a seemingly unpredictable behavior. Chaotic mappings find extensive applications across various domains, such as cryptography, random number generation, data encryption, communications, and chaos control. They are employed for generating pseudo-random number sequences, ensuring information security, simulating the behavior of intricate systems, and more. Additionally, chaotic mapping serves as a vital instrument for studying chaos theory and the dynamics of complex systems.

The purpose of incorporating random numbers into

various algorithms is to enhance the search space and improve the efficiency of algorithmic optimization. In conventional algorithms, random numbers are typically generated using random number generators, which produce numbers with statistical characteristics. However, strictly speaking, these numbers are generated by deterministic algorithms and are not truly random. Consequently, they can easily lead to convergence on local optima during the search for the optimal solution. To address this limitation, one can leverage chaotic mapping to generate random numbers. By utilizing chaotic mapping in operations such as population initialization, selection, crossover, and mutation, the overall algorithmic process can be influenced, often yielding superior outcomes compared to pseudo-random numbers.

Common chaotic mappings are logistic mapping, tent mapping, sine mapping, etc. The key is the control parameter, when the control parameter is relatively small, the solution tends to an immovable point, no longer change, which is a regular and orderly movement. When the control parameter is very large, the number of stabilized immovable points in this orderly motion will increase continuously. The increase in the number can still be regarded as orderly, and this increasing number of immobile points is called the "bifurcation phenomenon". With the increasing number of bifurcation immobile points, the ordered motion is gradually out of balance and enters a chaotic state.

5.Principles of SSA-BP optimization algorithms for chaotic mapping

From the fundamentals of the above two algorithms, it is clear that SSA algorithms require the original population, which is the basis for the evolution of subsequent populations. In general, the progress and capability of the algorithms are closely related to the original population. In order to improve the performance of the algorithms, chaos is often applied to optimization search problems. Among them, Tent model and Logistic model are the most commonly used chaotic models, but both of them are chaotic models with a limited number of mapping folds. Sin chaotic model is a model with an infinite number of mapping folds, and it is proved that Sine chaotic mapping is a chaotic mapping model based on sinusoidal function. It has the following advantages:

- Non-Linear Properties: Sine chaotic mapping is nonlinear, which means it can produce complex and unpredictable dynamic behavior. This nonlinear property makes it useful for a wide range of applications in fields such as cryptography, random number generation and data encryption.
- 2) Wide parameter space: the parameter space of Sine chaotic mapping is very wide, and the behavior of chaotic mapping can be changed by adjusting the parameters. This makes it very flexible and can be adapted to different application requirements.
- 3) Highly Sensitive: Sine chaotic mapping is very sensitive to small changes in initial conditions and parameters, which means that small input changes can lead to large changes in the output. This high

sensitivity makes it an important application in the field of encryption and communication, where it can be used to generate keys, encrypt data and protect communications.

4) Randomness: Sine Chaos Mapping produces highquality sequences of pseudo-random numbers. These pseudo-random numbers have good statistical properties and can be used in applications such as modeling random events, generating random numbers and performing Monte Carlo simulations.

Therefore, in this paper, Sin chaos is used to initialize the population of SSA algorithm, and the Sin chaos 1dimensional self-mapping expression is as follows:

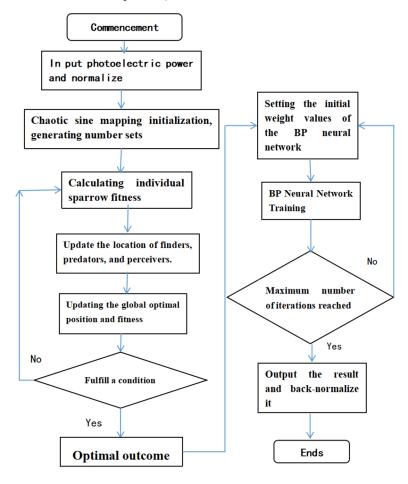
$$\begin{cases} x_{n+1} = \sin(2 / x_n) & n = 0, 1, ..., N \\ -1 \le x_n \le 1 & x_n \ne 0 \end{cases}$$
(4)

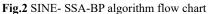
It produces a richer original population than the one generated using the rand random function⁹. Therefore, in this paper, Sine chaotic mapping is used to initialize the sparrow population, and the sparrow individuals with better adaptation and higher energy are selected as the original population. According to the program running process of BP network, it can be seen that the weights and thresholds have a great influence on the model, and the discoverer is close to the global optimal solution from the beginning of iteration, which leads to insufficient search range and easy to fall into the local extreme value space, resulting in insufficient search accuracy, this paper introduces the previous generation of global optimal solution in the updating formula of the discoverer's position, so that the position of the discoverer is affected by the position of the discoverer of the previous generation, and at the same time by the global optimal solution of the previous generation, thus, this paper adopts Sine chaos mapping to initialize the sparrow population. global optimal solution, which can effectively prevent the algorithm from falling into local optimality. In addition, we borrow the idea of inertia weight, and continue to introduce the dynamic weight factor ω into the finder position updating equation, so that it has a larger value at the beginning of the iteration, which is better for global exploration, and then adaptively decreases at the end of the iteration, which is better for local searching, and improves the speed of convergence at the same time. The formula of the weighting factor ω and the modified discoverer position updating formula are as follows:

$$\omega = \frac{e^{2(1-t/iter_{\max})} - e^{-2(1-t/iter_{\max})}}{e^{2(1-t/iter_{\max})} + e^{-2(1-t/iter_{\max})}}$$
(5)

Firstly, the historical data of PV power is normalized, and the first 80% is used as the training set, and the second 20% is used as the prediction set. Set the parameters of the sparrow algorithm: population size, maximum number of iterations, and safety threshold. The input layer of the BP network is 2, the hidden layer is 10, and the output layer is 1. Construct the chaotic sine sparrow optimizer to optimize its initial population size and number of iterations, and then import the historical data of the PV from the training set to the Sine-SSA-BP training model immediately afterward to update the position of the sparrow continuously. The mean square error (MSE) is chosen as the fitness function of the algorithm, and the algorithm loops every time and compares the optimal fitness value of the current sparrow with that of the previous generation of sparrows, and if it is greater than the value of the previous generation then it updates its position, and if it is not then it does not update it, and finally, it iterates until it satisfies the conditions to obtain the global optimum and the best fitness.

The flowchart of the SINE-SSA-Bp algorithm is shown in Fig 2:





6.Simulation results and analysis

In order to verify the superiority of SINE-SSA-BP prediction model in terms of accuracy. It is necessary to compare the simulation results of BP, RBP, SVM prediction model¹⁰, by comparing the size of the error between the predicted value and the real value of the simulation results, so as to reflect the superiority of the algorithm. The learning process of RBP neural network and BP neural network is largely the same, both contain input layer, implied layer, output layer. The input signal, including component temperature, air pressure, humidity, etc., through the implied layer to get the output value, and finally the output value and the expected value to do the difference, to determine whether to meet the accuracy requirements, and if not, then reverse transmission, and the use of gradient descent method to adjust the weights of each layer. The biggest difference between these two methods is that the action functions used are not the same. The action function of BP is a Sigmoid function, which is a non-zero value on an infinite number of inputs, so it has

a strong global optimization seeking ability. However, the action function of RBP neural network is Gaussian function and has the characteristic of non-zero value, which has the ability of local optimization and can greatly improve the learning efficiency of the network, and it is easier to meet the practical applications. SVM machine learning algorithm is an algorithm based on the principle of statistics, which is often used to deal with nonlinear dynamic regression problems, and it can make the original variables mapped from the low latitude to the highdimensional space so as to make predictions of the relevant problems. Mapping from low latitude to high dimensional space to make prediction of related problems. Below is a partial screenshot of the historical data for PV, taking into account module temperature, ambient temperature, barometric pressure, humidity, total radiation, direct radiation, and scattered radiation. Historical data is shown in Table 1.

time	hardware temperature	environmental temperature	pneumatic	humidity level
2019-01-01 00:00:00	-17.7	-13.154	926.071	53.497
2019-01-01 00:15:00	-18.115	-13.864	926.07	55.999
2019-01-01 00:30:00	-18.73	-14.461	926.07	57.747
2019-01-01 00:45:00	-18.0775	-13.066	926.071	54.162
2019-01-01 01:00:00	-17.2825	-12.823	926.07	52.807
2019-01-01 01:15:00	-18.0025	-12.793	926.07	51.031
2019-01-01 01:30:00	-18.7025	-13.671	926.071	53.597
2019-01-01 01:45:00	-18.1775	-13.905	926.071	54.602
2019-01-01 02:00:00	-18.235	-13.27	926.071	53.981
2019-01-01 02:15:00	-18.265	-13.598	926.071	54.478

Table1. Historical data on PV power

The simulation results are shown below:

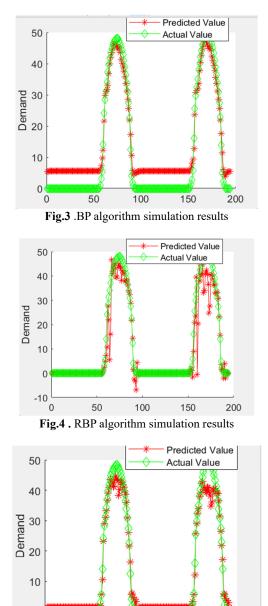


Fig.5 .SVM algorithm simulation results

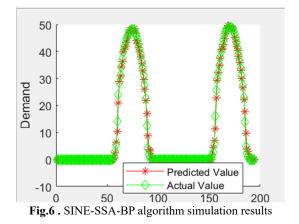
100

150

200

50

0



From the above Fig. 3, Fig. 4, Fig. 5 and Fig. 6, it can be seen that the SINE-SSA-BP neural network prediction model can meet the prediction of PV power, and the overall trend is closer to the actual value of PV power, and the relative error of the model is small and the prediction accuracy is high.

Further, in order to reflect the credibility of the simulation more intuitively, the root-mean-square error, mean-square error, and average absolute error of the four models are analyzed and compared, and if the smaller the error indicator is, the better the prediction effect of the model is.

ubles · comparison of Error maleators for Four models				
model	RMSE(Root	MSE(Mean	MAE(Mean	
	Mean	Squared	Absolute	
	Square	Error)	Error)	
	Error)			
BP	4.6261	21.4004	3.5356	
RBP	5.5095	30.3551	2.2974	
SVM	2.9448	8.6718	2.0671	
SINE-	1.2053	1.4527	0.81328	
SSA-BP				

Table2 . Comparison of Error Indicators for Four Models

From Table 2, it is clear that the SIEN-SSA-BP algorithm models are smaller and converge at essentially the same rate with respect to various error metrics relative to the other models.

7.Conclusion

In order to improve the prediction accuracy of photovoltaic power and ensure the safe, reliable and economic operation of the power grid. In this paper, the

traditional BP neural network has the disadvantage that the solution update rate is slow and it is very easy to enter the local optimum when encountering the situation of gentle gradient. The combined use of Sine chaotic mapping and sparrow algorithm is proposed, which can increase the stochastic and global search ability of BP neural network and improve the training effect and convergence speed of the network. These optimization methods can help BP neural networks better adapt to complex problems and improve their performance in tasks such as pattern recognition, prediction and optimization. The PV power prediction model using sine-SSA-BP neural network and the historical data of PV power of Guoneng Risin are tested by simulation. From the results of the simulation, the prediction accuracy of the SINE-SSA-BP neural network prediction model given in this paper is high and has good generalization performance.

References

- 1. WANG Changqing. Discussion on the technical characteristics and applications of solar thermal power generation[J]. China market,2019, 25(34):88-90.DOI:10.13939/j.cnki.zgsc.2019.34.088
- LI Zhonghong,HE Lesheng,WANG Jing, et al. Research on multi-scale photovoltaic power prediction based on full mining of meteorological information[J]. China Measurement & Test, 2022, 48(12):111-117.https://kns.cnki.net/kcms2/article/abstract?v=3u oqIhG8C44YLTIOAiTRKibYIV5Vjs7ioT0BO4yQ4 m mOgeS2ml3UDJaGIDTe3t-

C71O4xwaZkGMSpdmhWCVuoBKR9IMSpta&uni platform=NZKPT

- 3. DING Ming, WANG Lei, BI Rui. A short-term prediction model to forecast output power of photovoltaic system based on improved BP neural network[J].Power System Protection and Control, 2012, 40(11):93-99.https://kns.cnki.net/kcms2/article/abstract?v=3uo qIhG8C44YLTIOAiTRKgchrJ08w1e7fm4X_1ttJAk XBqTt_GkFylWsjGWC0bRipwxUOZKCtjun5Kou CTomKWAjrDHCPB13&uniplatform=NZKPT
- 4. DENG Weisi, MENG Zichao, WANG Haohuai,et al.Renewable energy power prediction characteristics analyses and accuracy improvementmeasures[J].Southern Power System Technology,2023,17(2):11-23.DOI:10.13648/j.cnki.issn1674-0629.2023.02.003
- 5. CHEN Changsong,DUAN Shanxu,YIN Jinjun. Design of photovoltaic array power generation prediction model based on neural network[J]. Transactions of China ElectrotechnicalSociety,2009,24(9):153-158.DOI:10.19595/j.cnki.1000-6753.tces.2009.09.023
- 6. ZHU Linlin, ZHONG Zhifeng, YAN Hai, et al. A new design of photovoltaic power generation forecasting mo -del[J]. Acta Energiae Solaris Sinica,

2016, 37(1): 63-68.

https://kns.cnki.net/kcms2/article/abstract?v=3uoqIh G8C44YLTIOAiTRKibYIV5Vjs7ijP0rjQD-AVm8oHBO0FTadhMtfHRhnUwfE8XAorofacVtGfl HSpAsyVZdO8Yr4GjY&uniplatform=NZKPT

- DING Ming,WANG Lei,BI Rui. A short-term prediction model to forecast output power of photovoltaic system based on improved BP neural network[J].Power System Protection and Control,2012,40(11):93-99.https://kns.cnki.net/kcms2/article/abstract?v=3uo qIhG8C44YLTIOAiTRKgchrJ08w1e7fm4X_1ttJAk XBqTt_GkFylWsjGWC0bRipwxU0ZKCtjun5KouC TomKWAjrDHCPB13&uniplatform=NZKPT
- HE Dong, LIU Ruiye. Ultra-short-term wind power prediction using ANN ensemble based on the principal components analysis[J]. Power System Protection and Control, 2013, 41(4): 50-54.https://kns.cnki.net/kcms2/article/abstract?v=3uo qIhG8C44YLTIOAiTRKgchrJ08w1e7xAZywCwkE EIyfIbD7Iew_BBuJY5Qjq9dIhNCt0GeiCganc4hmR pLugQbhLO1nbuN&uniplatform=NZKPT
- 9. WANG Xinpu, ZHOU Xiangling, XING Jie, et al.Prediction method of PV output power based on the combination of improved grey back propagation neural network[J]. Power System Protection and Control, 2016, 44(18): 81-87.

https://kns.cnki.net/kcms2/article/abstract?v=3uoqIh G8C44YLTIOAiTRKibYIV5Vjs7ijP0rjQD-AVm8oHBO0FTadpcWfbz2WCB7BT0cCxLvGdfNc Fdho2zcQlfk7xJNsf4b&uniplatform=NZKPT

 WEI X H. Short term photovoltaic power prediction based on grey model and machine learning[D]. Lanzhou : Lanzhou University , 2019.https://kns.cnki.net/kcms2/article/abstract?v=3 uoqIhG8C475KOm_zrgu4lQARvep2SAkOsSuGHv NoCRcTRpJSuXuqe2BzK-HIJJFZOgmR-RVlsMFQatP7gS4iGcvmSXIVZ9E&uniplatform=N ZKPT