Short-term forecast of photovoltaic power generation output based on improved PSO-Elman neural network

Yuhui Wu 1, Jingbo Liu 1, Jiaan Zhang 2,3 and Feng Hao 2

1 Power Science Research Institute of State Grid Ji Bei Electric Power Co., LTD; 2 Hebei University of Technology

3 Email: zhangjiaan@foxmail.com

Abstract. An improved particle swarm optimization (PSO) algorithm is proposed to solve the problems of low power prediction accuracy and high randomness of PV power generation. In the first place, by changing the inertia weight and learning factor and concept of the hybrid genetic algorithm is introduced into the particle swarm optimization algorithm, determine the initial weight of the Elman neural network, establish Elman neural network, then using day feature vector selection and forecasting meteorological features high similarity of date, determine the main meteorological factors affecting the photovoltaic output, meteorological data and electricity generation are used as training sets of Elman neural network, prediction model is established. Finally, the simulation results show that the model is superior to the Without optimization Elman neural network, has higher prediction accuracy, and shows good stability and generalization ability.

1. Introduction

In recent years, the photovoltaic power generation industry has developed rapidly, but the power of photovoltaic power generation is affected by many factors, and its instability and volatility bring a severe test to the prediction of photovoltaic power generation [1,2]. Therefore, accurate prediction of photovoltaic power is an important basis for grid security dispatching and is of great significance to grid security dispatching [3].

Power system power prediction is the prediction of power in the next hour or day. There are many factors affecting photovoltaic power generation, but the results predicted by traditional prediction methods are often not satisfactory. Therefore, the current photovoltaic power prediction mainly includes support vector machine [4], neural network [5-6] and time series method [7]. Among them, back propagation network (BP) is the most common, but the learning efficiency of this algorithm is low, and it is difficult to achieve a satisfactory result. Elman network is a kind of feedback neural network. On the basis of BP network, a continuation layer is added to make the system have the ability to adapt to time-varying characteristics, improve the global stability of the network, and have higher calculation accuracy and generalization [8-9]. However, the convergence speed is slow and it is easy to fall into local optimum, which affects the prediction accuracy.

Therefore, based on the correlation of photovoltaic output, this paper analyzes the influence of different factors on photovoltaic output. The proposed improved particle swarm optimization algorithm optimizes Elman neural network and introduces the concept of hybridization to improve the diversity of particle swarm and avoid premature phenomenon. So as to improve the prediction accuracy.
2. Selection and analysis of similar days
Solar radiation intensity is closely related to weather types, and photovoltaic system output varies greatly under different weather types [10]. The weather types selected in this paper are mainly irradiance, temperature and air humidity.

The prediction accuracy can be improved to some extent by searching the historical days through these meteorological factors.

1) Matching coefficient of meteorological factors
The matching coefficient of meteorological factors refers to the meteorological similarity between the predicted day and the historical day. In this paper, the meteorological factors that have the greatest influence on photovoltaic power generation are obtained through mutual information. The characteristic vector of meteorological factors is constructed, and the matching coefficient of meteorological factors is calculated through gray correlation analysis, as shown in equation (1).

\[
\xi_i(k) = \frac{\min_{i,k} |x_i(k) - x_j(k)| + \rho \max_{i,k} |x_i(k) - x_j(k)|}{|x_i(k) - x_j(k)| + \rho \max_{i,k} |x_i(k) - x_j(k)|}
\]  
(1)

In Equation (1), \(\rho\) is the value of the resolution coefficient in the range (0,1) and is usually \(\rho\) equal to 0.5.

2) Matching coefficient of week factor
The matching coefficient of week factor indicates the similarity between the predicted day and the historical day in the type of week, and the greater the similarity of the type of week, the greater the matching coefficient of week factor. At the same time, the week type is quantified, that is, Monday is mapped to 0.1, Tuesday to Thursday is mapped to 0.2, Friday is mapped to 0.3, Saturday is mapped to 0.7, and Sunday is mapped to 1.

Equation (2) is used to calculate the matching coefficient of weekly factor:

\[
\beta_i = 1 - |f(X_i) - f(X_0)|
\]  
(2)

Where, \(X_i\) is the week type of the \(i^{th}\) historical day, and \(X_0\) is the week type of the forecast day \(f(X_i)\) and \(f(X_0)\) are respectively the values after \(X_i\) and \(X_0\) mapping.

3) Comprehensive matching coefficient
To sum up, meteorological factor matching coefficient and week factor matching coefficient are obtained, and the two matching coefficients are multiplied to obtain the total similarity, that is, the comprehensive matching coefficient. The larger the comprehensive matching coefficient is, the closer the characteristics of the selected similarity day and the predicted day are.

3. Elman neural network
Elman Neural Network (ENN) was proposed by scholar Elman in 1990. It is a typical local regression network, which can be regarded as a forward neural network with memory units.

Compared with the traditional BP neural network, an additional continuation layer can be used as a delay operator to realize the memory function. Therefore, the Elman neural network can deal with nonlinear problems well, thus ensuring the global stability.

The Elman neural network consists of a four-layer structure: the input layer for receiving external information, the hidden layer for receiving signal transmission and transformation processing, and the continuity layer for saving and remembering the output value. Finally, the output layer that outputs the results.

Literature [11] describes the mathematical expression of this neural network as the following formula.

\[
y(k) = g(w_yx(k))
\]  
(3)

\[
x(k) = f(w_1x_1(k) + w_2(u(k-1)) + w_3x(k-1))
\]  
(4)

\[
x_1(k) = x(k-1)
\]  
(5)
Among them, $w_1$ is the connection weight between the input layer and the hidden layer; $w_2$ is the connection weight between the adjacent layer and the hidden layer; $w_3$ is the connection weight between the hidden layer and the output layer; $u$ is the r-dimension input vector; $y$ is the output node vector of m-dimension; $x$ is n dimensional intermediate node unit vector; $x_c$ is the n-dimensional feedback state vector; $g$ is the transfer function of the output layer, and $f$ is the transfer function of the hidden layer. The Elman neural network is characterized by connecting the output of the hidden layer to the input of the hidden layer through the delay and storage of the adjacent layer. This method improves the capability of dynamic information processing.

4. Improved particle swarm optimization Elman neural network

4.1. Particle swarm optimization

When the Elman neural network model is established, the initial weight of Elman network structure is obtained randomly. If the model is improperly selected, the convergence time may become longer or it falls into the local optimal solution. Therefore, this paper adopts an improved particle swarm optimization algorithm to optimize the initial weight of Elman neural network.

Particle swarm optimization algorithm was first inspired by the regularity of bird swarm activity, and then used swarm intelligence to build a simplified model. Particle swarm optimization (PSO) first initializes a set of random solutions, and then searches for the optimal solution in the process of continuous iteration. During each iteration, particles keep pressing Equations (6) and (7) to update the velocity and position of the next step.

$$\omega = \omega_{\text{max}} - \frac{t(\omega_{\text{max}} - \omega_{\text{min}})}{N}$$

where $\omega_{\text{max}}$ and $\omega_{\text{min}}$ are the minimum and maximum of inertia weights respectively. They are updated according to the results of each iteration. $N$ is the maximum number of iterations, and $t$ is the current number of iterations. The improved schemes for dynamic updating of learning factors are shown in Equation (9) and Equation (10). Compared with fixed learning factors, the response speed is accelerated.

$$c_1 = \frac{c_{\text{max}} - c_{\text{min}}}{N} + c_{\text{min}}$$

$$c_2 = 4 - c_1$$
In Equations (9) and (10), $c_{\text{max}}$ and $c_{\text{min}}$ are the initial and final values of $c_1$. In this paper, $c_{\text{max}} = 2.5$, $c_{\text{min}} = 0.5$, $t$ is the number of current iterations, and $N$ is the maximum number of iterations. At the beginning of the optimization, the global search of the entire search space is promoted, while at the end of the search, particles are encouraged to converge to the global optimization.

Then the hybrid concept of genetic algorithm is introduced into PSO to improve the diversity of PSO and avoid local convergence. The specific step is to give a random hybridization probability and select a certain number of particles into the hybridization pool according to the hybridization probability. The particles in the pool hybridize 2-2, so they will produce the same number of offspring, not the particles of their parents, so the whole remains the same. According to formula (11-14), the position and velocity of the new particle can be calculated to generate the progeny.

\begin{align*}
\text{Chi}_1(X) &= p*\text{Par}_1X + (1 - p)*\text{Par}_2X \\
\text{Chi}_2(X) &= (1 - p)*\text{Par}_1X + p*\text{Par}_2X \\
\text{Chi}_1(V) &= \frac{\text{Par}_1V + \text{Par}_2V}{|\text{Par}_1V + \text{Par}_2V|} |\text{Par}_1V| \\
\text{Chi}_2(V) &= \frac{\text{Par}_1V + \text{Par}_2V}{|\text{Par}_1V + \text{Par}_2V|} |\text{Par}_2V|
\end{align*}

In the formula, $\text{Chi}(V)$ and $\text{Par}V$ represent the velocity of offspring and parental generation respectively, while $\text{Chi}(X)$ and $\text{Par}X$ represent the position of offspring and parental generation respectively. Where $P$ is a random number between 0 and 1. The optimization method can improve the diversity of the population and prevent it from falling into local optimum.

4.3. Improved PSO optimizes Elman network implementation

The specific steps to optimize Elman neural network with improved particle swarm optimization algorithm are shown in Figure 1.

Step 1. Initialize Elman neural network, that is, the number of neurons in each layer, and conduct data preprocessing.

Step 2. Initialize each parameter in particle swarm, namely inertia weight $\text{min}$ and $\text{Max}$, as well as learning factor.

Step 3. Update inertia weight and particle velocity and position.

Step 4. Evaluate the historical optimal position of each particle.

Step 5. Update the global best position of the whole.

Step 6. Select a certain number of particles according to the hybridization probability to enter the hybridization pond for hybridization. And the velocity and position of the particle after hybridization are obtained through Equation (11-14).

Step 7. Judge whether the iteration is terminated. When the highest number of iterations is reached, the iteration is terminated and the optimal solution is output. Otherwise, step 3 is returned.

Step 8. Take the optimal value after PSO optimization as the initial weight of Elman neural network. The Elman neural network was pre-trained and the modeling of Elman neural network was finally completed.

5. Simulation and result analysis

In this paper, MATLAB is used to complete the simulation of photovoltaic power generation output power prediction. In order to verify the effectiveness of the prediction model, the meteorological data and output power of photovoltaic power station are respectively compared with the Elman neural network which is not optimized.
5.1. Model evaluation index

In this paper, mean absolute percentage error (MAPE) and determination coefficient ($R^2$) are adopted as the evaluation criteria for the prediction effect of the model. The formulas of the two evaluation indexes are as follows

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_{ai} - \bar{Y}}{\bar{Y}} \right|$$

(15)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{ai} - \bar{Y})^2}{\sum_{i=1}^{n} (y_{ai} - \bar{Y})^2}$$

(16)

In Equations (15) and (16), $N$ represents the number of photovoltaic power generation periods on the day to be tested, $y_{ai}$ represents the actual value of output power at time $i$, $\bar{Y}$ represents the predicted value of output power at time $i$, and $\bar{Y}$ represents the average value of actual output power.

**Figure 1.** Steps to optimize Elman Neural Network by improving particle swarm optimization Algorithm.
5.2. Case analysis
In this paper, one year's data of a photovoltaic power station is selected for training. The measured power value is collected every two minutes, namely 480 points from 5:00 to 21:00 every day. The improved particle swarm optimization algorithm is used to get the optimal weight of Elman neural network, and then the training sample is used to train Elman neural network to get the photovoltaic power prediction model. A day in February was randomly selected as the prediction day, and the date with the highest similarity to the prediction day was selected as the similarity day. Meteorological factors are the input variables of the prediction model, namely weather type, irradiance, temperature and humidity. And it is verified with the unoptimized PSO-Elman neural network, and the results are shown in Figure 2.

The prediction indexes of the two prediction methods were analyzed, and the results were shown in Table 1.

Table 1. Prediction error index analysis.

| Error indicators | PSO-Elman | Improved PSO-Elman |
|------------------|-----------|--------------------|
| MAPE             | 11.17     | 8.09               |
| R²               | 0.81      | 0.97               |

As can be seen from the above Table 1, the MAPE value and R² value of the prediction method in this paper are 8.09% and 0.97 respectively, while the MAPE value and R² value of the PSO-Elman model without optimization are 11.17% and 0.81 respectively. It can be seen that the prediction results of the prediction model in this paper are better than those of the PSO-Elman neural network model without optimization, and the accuracy is higher, therefore. The validity and accuracy of the prediction model are verified.

![Figure 2. Photovoltaic power prediction results.](image)

6. Conclusions
In this paper, the test date and the optimal similarity date are obtained by correlation analysis. And by changing the inertia weight and learning factor in particle swarm optimization (PSO) algorithm, then the hybrid concept is introduced into particle swarm optimization to optimize the initial weights of Elman neural network, An improved PSO-Elman neural network prediction model is proposed, and compared with standard particle swarm optimization algorithm to optimize Elman neural network is used in the comparison, the experimental results show that this algorithm has better prediction effect prediction accuracy and high stability.
Acknowledgments
This work was supported by the project “(5100-202036012A-0-0-00) Study of the new energy terminal active support ability evaluation online and online simulation modeling technology research and demonstration” State Grid Corporation headquarters science and technology project.

References
[1] Liu Weiliang, Liu Changliang, Lin Yongjun, et al. 2015 Ultra-short-term power prediction of photovoltaic power generation with factors affecting haze [J] Chinese Journal of Electrical Engineering 38(14) 4086-4095
[2] Cui Yang, Chen Zheng hong, Sun Pengjie 2008 Comparative Analysis of short-term PHOTOVOLTAIC power generation power forecast in different latitudes under the condition of optical abandoning and power limiting [J] Journal of Solar Energy 39(6) 1610-1618
[3] Liao N-Hwan, HU Zhihong, Ma Ying-ying, et al. 2011 Summary of short-term load forecasting methods for power systems [J] Power System Protection and Control 39(1) 147-152
[4] Ji Ling, Niu Dongxiao, Wang Peng 2015 Research on photovoltaic Power Prediction based on similar daily clustering and Bayesian Neural Network [J] Chinese management science 23(3) 118122
[5] Guo Yifei, Gao Houlei, Tian Jia 2016 Photovoltaic output Modeling with cluster Analysis and its Application in Reliability Evaluation [J] Power system automation 40(23) 93100
[6] Yao Zhongmin, Pan Fei, Shen Yuhui, et al. 2015 Short-term Prediction of photovoltaic power station output based on GA-BP and POS-BP neural networks [J] Power system protection and control 43(20) 8389
[7] Zhang Jianyi, Zhou Buxiang, Lin Nan, et al. 2013 The grid of grey Elman neural network is negative for a long time Charge prediction [J] Journal of power systems and automation 25(4) 145149
[8] Wang Zhizhong, Han Maolin, Hu Hai 2019 Research on Photovoltaic Power Premeasurement Method based on Sunniness Coefficient and Multi-level Matching [J] Electrical measurement and instrumentation 56(8) 4550
[9] Yang MAO, Yang Yu 2019 Short-term Photovoltaic output Power Prediction Research based on Wavelet Packet and LSSVM [J] Renewable Energy 37(11) 1595-1602
[10] Liang Caixia, GAO Zhaliang 2019 Short-term Power Prediction of PHOTOVOLTAIC power Generation based on Similar Day and GA-DBN Neural Network [J] Electrical Applications 3 97-102
[11] Chen Yanping, MAO Yi, Chen Ping, et al. 2016 Based on EEMD- sample entropy and Elman neural network Short-term power load forecasting [J] Journal of power systems and automation 28(3) 59-64