

Photovoltaic Output Power Prediction Based on Weather Type

Nan Shi ^{1a}, Xianhui Zhu ^{2b}, Pengtao Yuan ^{1c}, Jiaoxia Hao ^{2d}, Qinnan Wang ^{2e}

¹ The Engineering Training and Basic Experimental Center, Heilongjiang University of Science and Technology, Harbin, China

² School of Electrical and Control Engineering, Heilongjiang University of Science and Technology, Harbin, China

^a snhit@sina.cn, ^b zhu_xianhui@sina.com, ^c YPTuanengao@sina.cn,

^d 1193762804@qq.com, ^e 3220796445@qq.com

Abstract. Precise prediction of photovoltaic (PV) output power can effectively improve the safe and stable operation of power grids. In this paper, the influence of weather types on the output of PV output power is analyzed. Based on extreme learning machine (ELM) neural network, a prediction model of PV power generation output taking into account weather types is established. The grey correlation analysis method is used to process the weather types, and the established ELM neural network prediction model is trained by using the prediction results of the improved equal-dimension grey (IEDG) model, and the trained model is used to predict the output of PV output power. Compared results show that the prediction model presented in this paper can predict the output of each day in different weather types. The results show that the proposed model based weather type is of high accuracy and reliability in predicting the output power of PV.

1. Introduction

With the increasing utilization of fossil energy, photovoltaic (PV) and wind power have developed rapidly. By the end of 2017, the accumulative installed capacity of the global PV power generation has reached 405 GW, and the accumulative installed capacity of China is 130 GW. PV power generation has a random fluctuation characteristic because it is affected by factors such as radiation intensity and ambient temperature. On the other hand, the large-scale distributed PV power connected with the grid poses great challenges to the planning, operation and control of the power grid. In order to reduce the impact of PV output instability on the grid, it is often necessary to use traditional loads for rotating backups, and the forecast of PV output power is important for the grid.

At present, PV output power prediction modeling includes physical methods and statistical methods. Physical methods are based on energy conversion devices in PV power generation and mathematical models of various parts of control systems. The effectiveness of the prediction depends on the control degree of the object structure and the precision of the model parameters in the process of PV power generation. The design links are many, the process is complex, and the parameters are difficult to solve [2]. The representative statistical methods are ARMA [3] artificial neural network [4], support vector machine (SVM) [5], and Markov chain [6].



The traditional PV output power generation forecasting methods have some limitations, which need to be improved. In this paper, an improved equal-dimension grey(IEDG)-extreme learning machine (ELM) PV power generation prediction method based on weather type is proposed. Through the analysis and statistics of historical data, the PV output power is predicted with the weather information. The results show that the model has accurate prediction ability under various weather types.

2. Select similar days based on weather types

In this paper, the weather is divided into four types: sunny day, cloudy day, rainy day, and overcast day. According to the meteorological parameters of these four types, historical days similar to the predicted day are selected. Meteorological factors have a significant influence on PV power generation, ambient temperature and relative humidity are selected from all meteorological factors, and form the feature vector of meteorological factors, and then the grey correlation analysis method are used to calculate the similarity [7].

In this paper, the feature vectors of weather factors for the forecast day and the i -th historical day are $W(0)=[W_0(1),...,W_0(m)]$, $W_i=[W_i(1),..., W_i(m)]$, m is the number of meteorological factors. The difference between the k -th meteorological factor eigenvectors on the forecast date and the i -th historical date is

$$\Delta_i(k) = |W_i(k) - W_0(k)| \quad (1)$$

The normalized calculated value of $\Delta_i(k)$ is

$$\Delta_{li}(k) = \frac{\Delta_i(k) - \min_i \Delta_i(k)}{\max_i \Delta_i(k) - \min_i \Delta_i(k)} \quad (2)$$

The correlation coefficient of the k -th meteorological factor eigenvector between the forecast date and the i -th historical day is

$$\begin{cases} \gamma_i(k) = \frac{\Delta(\min) + \rho\Delta(\max)}{\Delta_{li}(k) + \rho\Delta(\max)} \\ \Delta(\min) = \min_i \min_k \Delta_{li}(k) \\ \Delta(\max) = \max_i \max_k \Delta_{li}(k) \end{cases} \quad (3)$$

Where $\Delta(\min)$ is the smallest characteristic value of the difference in all historical days, $\Delta(\max)$ is the largest characteristic value of the meteorological factors in all historical days, ρ is the resolution coefficient that is a constant between 0 and 1, which is usually 0.5.

With the correlation coefficients of the feature vectors of various meteorological factors, the similarity between the forecast date and the i -th historical day is

$$\alpha_i(k) = \frac{1}{m} \sum_{k=1}^m \gamma_i(k) \quad (4)$$

In prediction, the similarity between the nearest historical day and the forecast day is calculated at first, then the similarity between each historical day and the forecast day is calculated forward day by day, and the last N day ($r_i > 0.8$) is selected as the similar day of the prediction day [8].

PV output power generation data at the same time on the five days that are similar to the forecast day will be used as input, and the ant colony gray model will be used to predict.

3. Improved equal-dimension grey (IEDG) model

3.1. GM(1, 1) Model

For the problem that the prediction accuracy is not high with a few information systems, the traditional gray model GM(1, 1) generates a sequence with a regular pattern by accumulating the data of the

original time series once and fits it with a typical curve., and effective prediction of a small amount of data can be achieve . Specific steps are as follows:

- 1) Assumin that the time sequence of the original power is $P^{(0)}=[P^{(0)}(1), P^{(0)}(2), \dots, P^{(0)}(t)]$ ($t=1, 2, \dots, n$), a new sequence $P^{(1)}$ is generated by first order accumulation of the original power sequence.
- 2) According to the new, sequence $P^{(1)}$, the differential equation of whitening form is established.

$$\frac{dP^{(1)}}{dt} + aP^{(1)} = b \quad (5)$$

Where a and b are the model parameters determined by using the least square method.

The parameters a and b of the GM(1,1) can be determined by using the least square method as follows:

$$\begin{bmatrix} a \\ b \end{bmatrix} = [\beta^T \beta]^{-1} \beta^T Y \quad (6)$$

$$\text{Where } Y = \begin{bmatrix} P^{(0)}(2) \\ P^{(0)}(3) \\ \vdots \\ P^{(0)}(n) \end{bmatrix}, \quad \beta = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(2)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(3)}(n) & 1 \end{bmatrix}.$$

Where

$$Z^{(1)}(t) = \omega * P^{(1)}(t-1) + (1-\omega) * P^{(1)}(t) \quad (7)$$

- 3) The Accumulated Generating Operation (AGO) Grey prediction can be obtained:

$$\hat{P}^{(1)}(k) = [P^{(0)}(1) - \frac{b}{a}]e^{-a(k-1)} + \frac{b}{a} \quad (8)$$

- 4) With the inverse accumulation the predicted power is obtained.

$$\hat{P}^{(0)}(t+1) = \hat{P}^{(1)}(t+1) - \hat{P}^{(1)}(t) = (1-e^a)[P^{(0)}(1) - \frac{b}{a}]e^{-at}, t=1, 2, \dots, n \quad (9)$$

3.2. IEDG model

The GM (1, 1) model adopts the historical data at $t=n$ when modeling. However, during the development of any grey system, some random disturbances will continue to enter the system and affect the development trend of the system as time goes on. When using the traditional GM (1,1) model to predict, the higher precision is only a few data closer to the prediction moment, and the point farther from the prediction moment has less influence on the actual prediction result. In order to reflect the influence of stochastic disturbances on the grey system and improve the prediction accuracy, the traditional GM (1, 1) model is improved and the equal dimension GM (1, 1) model is obtained. The process is as follows.

$P^{(0)}(n+1)$ is added to the original data sequence to delete the oldest information $P^{(0)}(1)$, and the new data sequence is: $P^{(0)}=[P^{(0)}(2), P^{(0)}(3), \dots, P^{(0)}(n+1)]$, the initial value modified GM(1, 1) model that is the improved equal-dimension information grey model is established according to the above steps. The improvement forecast data is

$$\hat{P}^{(0)}(n+2) = \hat{P}^{(1)}(n+1) - \hat{P}^{(1)}(n) = [P^{(0)}(2) - \frac{b}{a}](1-e^a)e^{-at} + \frac{b}{a} \quad (10)$$

In the grey model GM(1, 1), the value of w in equation (4) is determined by the values of a and b , and the values of a and b directly affect the prediction accuracy of PV power generation output. The following methods can be used to improve the prediction accuracy. The absolute error between the PV output power generation predicted value $\hat{P}^{(0)}(t+1)$ and the measured PV output power generation

output $P^{(0)}(t+1)$ is $\Delta(t+1)$, and the first t moments prediction data absolute error of PV power generation can be obtained by using historical data. The choice of an appropriate ω value in equation (7) has an important influence on the accuracy of prediction of PV power generation output. Therefore, the ant colony optimization algorithm is used in this paper to determine the value of ω so that the squared sum function F of the PV output prediction deviation is the minimum. The F function is

$$F = \sum_{t=0}^i \Delta(t)^2 = \sum_{t=0}^i \left| \hat{P}^{(0)}(t) - P^{(0)}(t) \right|^2 \quad (11)$$

The value of ω is obtained by using the minimum square sum of the output deviation of the PV power generation, and then the values of a and b are determined and the output is predicted to achieve better prediction accuracy. The ant colony optimization algorithm is used to optimize the value of ω with the goal of minimizing F [9].

4. ELM Prediction Model

The ELM algorithm has a simple structure, fast learning speed, good global search ability and excellent generalization performance. Because ELM needs a lot of data training, this paper uses the ant colony grey prediction model provided in the paper to use the four similar power generation power as ant colony at every time. The input of improved equal-dimension gray model is used to predict the power generation at each time of the forecast day, and the result is taken as the input of ELM and is used to train ELM, the predicted value can be got.

The process of ELM is as follows[10]:

- 1) Input: a given set of training sets $\{x_i, t_i\}_{i=1}^N \in \mathbb{R}$, test sample set, activation function $v(x)$, the number of hidden layer nodes L , random generation of hidden layer node parameters (c_i, d_i) , $i=1, 2, \dots, L$, where c_i is the input weight of hidden layer units, and d_i is the deviation of hidden layer elements.
- 2) The output matrix H of the hidden layer unit is calculated

$$\left\{ \begin{array}{l} H(c_1, L, c_L, d_1, L, d_L, x_1, L, x_N) = \begin{bmatrix} v(c_1, d_1, x_1) & v(c_L, d_L, x_1) \\ \vdots & \vdots \\ v(c_1, d_1, x_N) & v(c_L, d_L, x_N) \end{bmatrix}_{N \times L} \\ H^+ = (H^T H)^{-1} H^T \\ \beta = H^+ T \\ \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \end{array} \right. \quad (12)$$

Where β is the output weight, H^+ is the Moore-Penrose generalized inverse matrix of the hidden matrix unit output matrix H

- 3) Output: according to the test sample set $\{y_i\}_{i=1}^M$ and the output weight β , the predicted values corresponding to the data in the test sample set are calculated.

For good prediction accuracy, ELM model needs a lot of data to train. In this paper, the prediction results of the IEDG model in this paper as the input of the limit learning machine, and then ELM is used to train and predict. The amount of data can be reduce and the accuracy of prediction can be improved.

5. Example analysis

Taking the power of a PV power station in Heilongjiang as the original data, first of all, according to the data of different weather types, the environmental temperature and relative humidity are analyzed by grey correlation analysis in the two seasons of summer and autumn in 2016, and 4 types of day

(sunny, cloudy, rainy and overcast) are selected and predicted. The 20 days of the similar degree greater than 0.8 are selected with different weather types. With the data of 15 days (6-18 points) are used to train and the data of the other 5 days as the test data, the prediction results obtained by the only ELM model using the same historical data are compared with the results of IEDG-ELM model proposed in this paper under the conditions sunny days, cloudy days, rainy days and overcast days. The results are shown in figure 1-4. In figure 1-4, line 1 and line 2 represent the actual measured values and the predicted values based on IEDG-ELM model, and line 3 represents the predicted values based on ELM model.

The relative errors of the prediction results based on IEDG-ELM model and ELM model are 12.72% and 16.08 in sunny day from figure 1. The relative errors of the prediction results based on IEDG-ELM model and ELM model are 14.08% and 15.41% in rainy day from figure 2. The relative errors of the prediction results based on IEDG-ELM model and ELM model are 13.24% and 17.31% in cloudy day from figure 3. The relative errors of the prediction results based on IEDG-ELM and ELM are 18.85%, and 23.8% in cloudy day from figure 4. It can be seen that the accuracy is higher based on IEDG-ELM model than ELM model. IEDG-ELM model can effectively predict the PV output based different weather types.

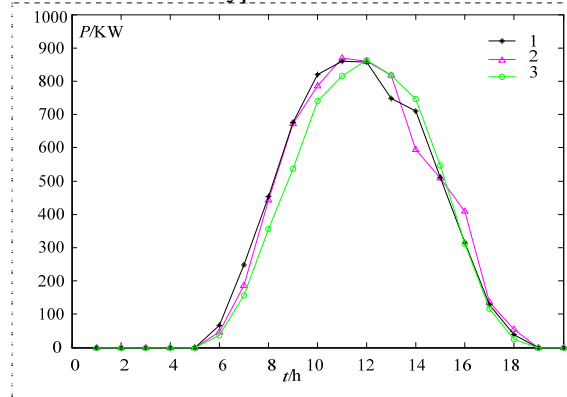


Figure 1. Forecasting results of a sunny day

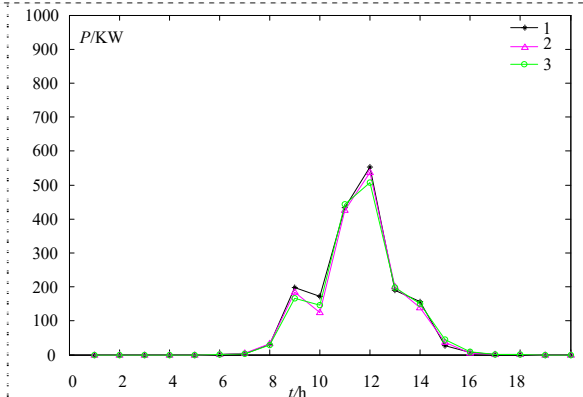


Figure 2. Forecasting results of a rainy day

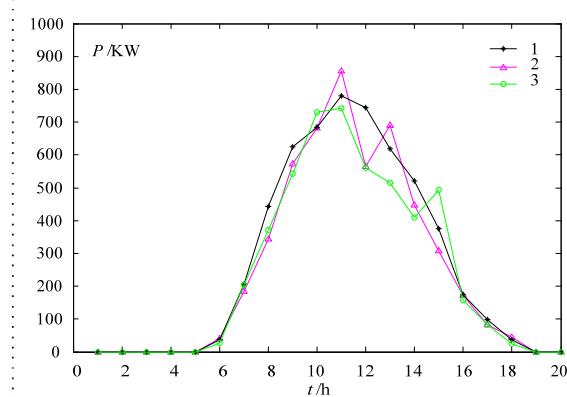


Figure 3. Forecasting result of a cloudy day

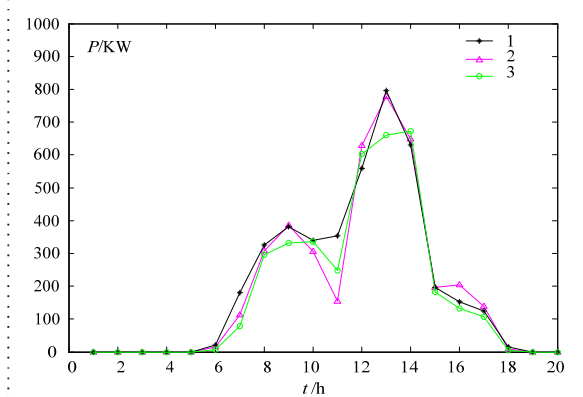


Figure 4. Forecasting result of an overcast day

6. Conclusion

In this paper, based on the influence of environmental factors on the output power of PV power generation, based on the limit learning machine, a forecasting model of PV power output is set up with weather types, and the historical power data of the PV power station is reasonably divided according to the different weather types, and the weather type is further classified by the grey correlation analysis method. The output results of IEDG model are used as the input of the ELM model, and ELM model is trained and tested. The prediction results are analyzed to get the following conclusions.

1) IEDG-ELM proposed in this paper can effectively predict the output power of PV systems under various weather conditions.

2) IEDG-ELM models proposed in this paper has a relatively accurate prediction ability and strong applicability.

Acknowledgments

This work was supported by University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (UNPYSCT-2017144) and National Natural Science Foundation of China (51677057).

References

- [1] Shengwei M, Jianquan Z 2013 *Acta Automatica Sinica*. **39** 119
- [2] Fei W, Zengqiang M, Zhao Z, Guang, Y and Haiming Z 2013 *Proceedings of the CESS* **33** 75
- [3] Hua L, Zhimin L and Yang Z 2011 *Electrical Measurement and Instrumentation*. **48** 31
- [4] Shouxiang W and Na Z 2012 *Automation of Electric Power Systems*. **36** 1
- [5] Mei ping F, Hong wei M and Jianrong M 2012 *Power System Protection and control*. **40** 65
- [6] Ding, M., and Xu N 2011 *Power System Technology*. **35** 152
- [7] Si feng L, Tianbang G, Yaoguo D 1999 *Theory and Application of Grey System* (Beijing: Science Press) pp 35-42
- [8] Meng Y, Lu J, Sun H, Pan X, Gao D and Liao Y 2010 *Power System Technology*. **34** 163
- [9] Nan S, Sunquan Z and Yidan L 2015 *Journal of Electric Power Science & Technology*. **30** 61
- [10] Hongbin D, Jinwei P, Qilong H 2015 *computer science* .**42** 78