Photovoltaic fault diagnosis model based on dynamic switching

Nanzhou Chen¹, Shan Hu², Wenhao Zhu³*, Fei Wang⁴

¹ School of Computer Engineering and Science, Shanghai University, Shanghai, Province, 200000, China
² School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, Province, 200000, China
³ School of Computer Engineering and Science, Shanghai University, Shanghai, Province, 200000, China
⁴ School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, Province, 200000, China

¹ e-mail: nzhchen@shu.edu.cn
* Corresponding author’s e-mail: whzhu@shu.edu.cn

Abstract: PV module fault diagnosis is mostly based on the prediction of photovoltaic module's power generation in the short term, and the fault of the module can be judged by comparing the difference between the actual power generation and the predicted power generation. However, the photoelectric conversion efficiency of photovoltaic modules varies significantly with seasons and weather, single prediction model is lack of stability, the season in a significant and a change in the weather forecast results will appear large deviation, this will lead to frequent false alarms of fault diagnosis, make power plant maintenance and repair costs unnecessarily. According to the characteristics that the photoelectric conversion efficiency of photovoltaic power generation is greatly affected by weather fluctuations, a fault diagnosis method for photovoltaic power generation based on dynamic model switching is proposed. When meet preset conditions, using the different quarters of historical data, combined with recent data to update the model, improved the fitness of model for the current weather and season. Compared with a single fixed model, the results show that the proposed method has better adaptability to different seasons and different weather.

1. Introduction

Adaptive model prediction method has been proved to be an effective method to improve the accuracy of wind power generation[1]. In this paper, a fault diagnosis method for photovoltaic power generation based on dynamic model switching is proposed based on the characteristics that the photoelectric conversion efficiency of photovoltaic power generation is greatly affected by weather fluctuations. For the first time, the concept of Weather fluctuation rate and Restart Land Mark was proposed, the variation of meteorological environment around photovoltaic power station was statistically analyzed, the fluctuation rate of weather conditions around modules was obtained, and the model was updated appropriately according to the weather fluctuation. The aim is to make the model grasp the change of photoelectric efficiency of the component in time. Adaptive model prediction method has been proved...
to be an effective method to improve the accuracy of wind power generation. Compared with the traditional prediction method[2], this method improves the adaptability of the model to different weather and different seasons, and the overall detection accuracy is significantly improved. On the other hand, the data demand of the model is reduced, which makes the realization cost of the fault detection model lower.

2. Model introduction

2.1. Process of the model
Dynamic switching component fault diagnosis model of generating principle is to influence the efficiency of the power components power statistical and testing of a variety of weather conditions that can direct response of the fluctuation in the weather volatility, once the weather more than a specified threshold, the volatility model automatically search the latest data from the database, use the new data set to update the model. The automatic grid pattern parameter can ensure that the model automatically looks for the optimal model structure during the updating process, and the latest data sets can enable the model to learn the generation rules of the generation components related to the current weather and meteorological conditions. The whole process is completed automatically by the program without human intervention.

In order to ensure the applicability of the model in practical application and to avoid the over-fitting of the model caused by the small amount of data and the high degree of data concentration, we extracted a certain amount of data from the data set of the power station in the past year according to different months to build the basic training set of the model. Of choice of meteorological parameters to monitor at the same time, calculating the recent volatility of weather, when there is a qualified replacement nodes, the model automatically from the data platform for data collection in the latest week, mix with the construction of the initial data sets, update to the latest data set as a model, and then use the network style and super parameter of the model for automatic debugging. The best model is selected to update the current model in order to summarize the influence of recent weather fluctuations on the photoelectric conversion efficiency of photovoltaic modules.

2.2. weather fluctuation rate
Weather Fluctuation Rate[3] is a calculation method used to measure weather fluctuation. By calculating the relative offset of various weather parameters in the high-dimensional space in the recent period, and combining the weight of weather fluctuation situation before different time span at the current moment, the important parameters reflecting the recent weather fluctuation are calculated. The reset landmark is an important symbol for updating the whole prediction model. The program calculates the recent weather volatility around the photovoltaic power station in real time. When the weather volatility is determined to exceed the specified threshold, the model used to predict the real-time power generation is updated. To ensure that the model can grasp the change of photoelectric conversion efficiency due to weather changes in time. Hourly weather data of the weather station nearest to the power station are used as the basis for calculating the volatility, and meteorological parameters that may affect the light-point conversion efficiency of the module are selected as the basis for calculating the weather volatility.

We arrange the weather conditions of the past week in chronological order, and for a certain moment we define as t, we calculate the offset of the meteorological feature vector in high dimensional space relative to t-∆t at the previous moment. In order to prevent the model deviation caused by the different units of different meteorological parameters, we use the relative percentage to calculate the weather offset. The calculation formula for the weather volatility at a certain time t is as follows:

$$r_t = \sqrt{\frac{1}{n} \left( \left( \frac{S_l^1 - S_{t-\Delta t}^1}{S_{t-\Delta t}^1} \right)^2 + \left( \frac{S_l^2 - S_{t-\Delta t}^2}{S_{t-\Delta t}^2} \right)^2 + \cdots + \left( \frac{S_l^n - S_{t-\Delta t}^n}{S_{t-\Delta t}^n} \right)^2 \right)}$$

(1)
Where S represents a number of meteorological parameters related to the photoelectric conversion efficiency of a photovoltaic power station, with a total of n; t represents the moment of current calculation; \( \Delta t \) represents the time accuracy used in our calculation.

We chase for the past week day weather volatility to carry on the summary, to draw the whole weather volatility, and presents the timing relationships, due to the change in the weather condition far away the weather volatility for the current moment under the influence of the components of the photoelectric conversion efficiency and precision of model is relatively lower, we can't in the past and present the data using the same weight, Therefore, the exponential function e-x is used to reduce the weight of the weather fluctuation in the period before the current moment. For the weather fluctuation in the day d before the current moment, we can express as follows:

\[
R_d = e^{-d} (r_t + r_{t-\Delta t} + \cdots r_{t-k\Delta t})
\]  
\( k = \frac{d}{\Delta t} \) 

For the overall volatility of meteorological conditions in the past week, it can be expressed as:

\[
R = R_d + R_{d-1} + \cdots + R_{d-7}
\]

Frequent model updating may reduce the prediction accuracy of the model and increase the system burden. Therefore, after calculating the weather volatility of the current date, we need to determine the threshold value of volatility and determine the optimal model update node. That is, the Reset Landmark we proposed above.

3. Experimental

3.1. the experimental setup

We input data from the past year into the simulation program and measure the accuracy of the model using the mean absolute percentage error (MAPE), which is calculated as follows:

\[
MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\text{actual}(t) - \text{forecast}(t)}{\text{actual}(t)} \right| \times 100\%
\]

In order to verify the improvement of the overall prediction accuracy of the model by the reset node and the effect of the improvement, two groups of comparative experiments were set up:

(1) Experiments with different precision: On the meteorological data precision of calculation is used to the weather volatility, the grouping experiment respectively using the time interval for one hour, two hours, four hours of timing formula of weather data input volatility, to calculate the corresponding volatility, and compared under the condition of no threshold, reset the node to ascend to the precision of the model, determined under different time intervals, Reset the maximum enhancement effect of the node on the specified model.

(2) Sub-model experiment: in order to ensure the accuracy of the experiment, we added reset nodes to several machine learning models for updating, so as to check whether the reset nodes generally improve the accuracy of different machine learning models and to what extent the accuracy is improved. We chose XGBoost[4], SVR[5] and BP neural network[6] as the reference model.

3.2. result

Table1: The minimum error percentage, average error percentage of each model at different time accuracy, as well as the accuracy of the model without adding the restart land mark.

| Accuracy | Model | Min MAPE | Average Of MAPE | MAPE without Restart Land Mark |
|----------|-------|----------|----------------|-----------------------------|
| One Hour | SVR   | 7.91     | 8.138          | 8.23                        |
|          | XGBoost | BP-NN  | SVR   |
|----------|---------|--------|-------|
| Two Hour | 7.15    | 9.87   | 7.72  |
|          | 7.387   | 10.034 | 8.119 |
|          | 7.56    | 10.01  | 8.23  |
| Four Hour| 7.12    | 9.84   | 7.95  |
|          | 7.397   | 10.028 | 8.147 |
|          | 7.56    | 10.01  | 8.23  |

4. Conclusion
From the experimental results, we can see that under the same accuracy, the resetting node has different lifting effects on different models. The promotion effect of XGBoost and SVR is obviously greater than that of BP-NN, and XGBoost has the best promotion effect among all models, mainly because XGBoost is a machine learning algorithm based on random forest, which can also achieve good results on the basis of small data. Under the premise of the same model, the time interval of selecting input to calculate weather volatility is different. The promotion effect of the model is also different. According to the current experimental results, on the premise that the time accuracy is selected as two hours, the resetting node has the best improvement effect on each model. On the whole, the current experimental results show that adding the reset node based on weather volatility calculation into the model has a certain effect on improving the prediction accuracy of the model.

Acknowledgments
This work is supported by National Natural Science Foundation of China (No. 61572434) and Shanghai Science and Technology Committee (No. 19DZ2204800).

References
[1] Lahour A, Slama J. (2017) Hour-ahead wind power forecast based on random forests. Renewable energy. 109: 529-541.
[2] Bracale A, Carpinelli G, De Falco P. (2016) A probabilistic competitive ensemble method for short-term photovoltaic power forecasting. IEEE Transactions on Sustainable Energy. 8(2): 551-560.
[3] Zhang Z. (2018) Influence of weather Fluctuation Rate on output of PV system. Conference Series: Earth and Environmental Science. 108-5.
[4] Chen T, Guestrin C. (2016) Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 785-7.
[5] Liu F T, Ting K M, Zhou Z H. (2008) Photovoltaic fault diagnosis model base on SVR. 2008 Eighth IEEE International Conference on Data Mining. IEEE. 413-422.
[6] G Graditi, S Ferlito, G Adinolfi. (2016) Comparison of Photovoltaic plant power production prediction methods using large measured dataset. Renewable Energy. 513-519