A fault diagnosis system for PV power station based on global partitioned gradually approximation method

S Wang, X N Zhang, D D Gao[1], H X Liu, J Ye and L R Li

School of Mechanical Engineering, Qinghai University, Xining, Qinghai Province, 810016, China

E-mail: gaodd@qhu.edu.cn, 1185725268@qq.com

Abstract. As the solar photovoltaic (PV) power is applied extensively, more attentions are paid to the maintenance and fault diagnosis of PV power plants. Based on analysis of the structure of PV power station, the global partitioned gradually approximation method is proposed as a fault diagnosis algorithm to determine and locate the fault of PV panels. The PV array is divided into 16×16 blocks and numbered. On the basis of modularly processing of the PV array, the current values of each block are analyzed. The mean current value of each block is used for calculating the fault weigh factor. The fault threshold is defined to determine the fault, and the shade is considered to reduce the probability of misjudgments. A fault diagnosis system is designed and implemented with LabVIEW. And it has some functions including the data real-time display, online check, statistics, real-time prediction and fault diagnosis. Through the data from PV plants, the algorithm is verified. The results show that the fault diagnosis results are accurate, and the system works well. The validity and the possibility of the system are verified by the results as well. The developed system will be benefit for the maintenance and management of large scale PV array.

1. Introduction

In the modern society of energy shortage and environmental pollution, the solar energy, an ideal type of new energy, is widely developed and utilized [1]. With the strong development of PV industry, the number of PV power generation system is also increasing. Over the past ten years, the PV industry has been developed rapidly in China. The PV installed capacity get to 43GW until 2015, which take up about 25% of the world. However, China is considered as the mostly world's largest PV manufacturer with unlocked its PV market potential. The country is planning to meet a goal of 20 GW of solar power installed by 2020, 5 GW of this by 2015 which is of course negligible considering its huge potential. The PV power plant's life enters a recession, more and more problem has cropped up. There are two main aspects need attention. Firstly, it is required to improve the efficiency of the system and reduce costs. Secondly, the failure of PV panels limits the normal run of the power station. The fault can be attributed to the electrical disturbances. So it is necessary to design a fault diagnosis method for PV array for improving the power generating efficiency.

At present, the fault detection method of the PV array is mainly divided into traditional diagnostic methods, intelligent algorithm and model algorithm [2-4]. The traditional diagnostics include online

[1]Address for correspondence: D D Gao, School of Mechanical Engineering, Qinghai University, Xining, Qinghai Province, 810016, China. E-mail: gaodd@qhu.edu.cn, 1185725268@qq.com

diagnostics and offline diagnostics [5]. Multi-sensor method is the most widely used for online diagnosis [6-7]. Offline diagnostics require downtime checking. The real-time online fault detecting and location can't be implemented, and diagnostic accuracy is limited by equipments [8]. Circuit model [9] and power model [10] are used to judge the fault, but the model algorithm is fail in fault location. BP neural network [11], data fusion theory [12-13], expert system, et al. are utilized in fault diagnosis to detect and locate the fault. However, the fault detection accuracy depends on the parameter settings and samples, which leads to low determinacy of the intelligent algorithm.

There are many methods about the design of the upper computer operating interface for the monitoring system, such as VB, VC++, web browsing, MCGS configuration software or LabVIEW development environment [14-19]. LabVIEW is widely used in industrial measurement and control platform due to the characteristic of powerful and simple programming.

In this paper, a new fault diagnosis algorithm, Global partitioned gradually approximation (GPGA) method, is proposed to realize fault detection and location for PV array. And a fault diagnosis monitoring simulation system for 20MW large-scale PV power station combined with the algorithm is implemented based on LabVIEW. The simulation system has the following functions: 1) data real-time display, 2) online check, 3) data statistic, 4) data save and print, 5) real-time power prediction, 6) fault diagnosis.

This paper is organized as follows: section 2 presents the algorithm for PV array. At first, the PV power station structure is analyzed. Then the fault diagnosis algorithm is designed according to the structure. In section 3, the simulation system is implemented based on LabVIEW, and the experiments are designed to verify the global partitioned gradually approximation method. In section 4, the simulation results and discussion are provided to illustrate the feasibility and validity of the fault diagnosis system. Finally, conclusions and future work are given in Section 5.

2. Algorithm of fault diagnosis for PV array

2.1. Photovoltaic power station structure analysis

In order to study the composition of PV power station system, a number of PV power stations are investigated in Golmud City, Qinghai Province. Figure 1 diagrams the structure of a 20MW PV power station. 20 single PV panels with the same specification (240W) in series form a branch. 16 branches go into a combiner box. 8 combiner boxes are connected to a single inverter, 2 inverters with the same specification (500kW) form an inverter square. There are 20 inverter squares in total. 4 inverter squares in parallel go into a box type transformer. 5 box type transformers go into 10kV bus to boost, and the main transformer feeds into the grid with 35kV high voltage output.

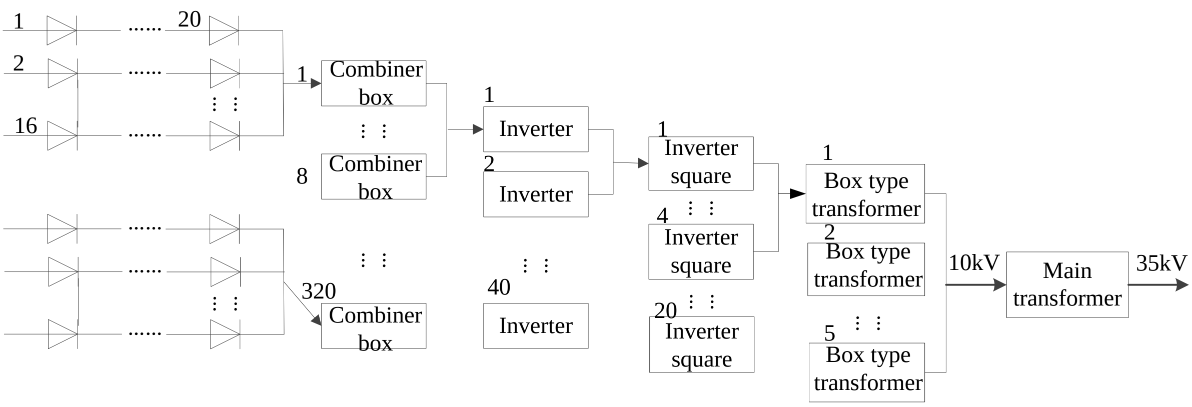


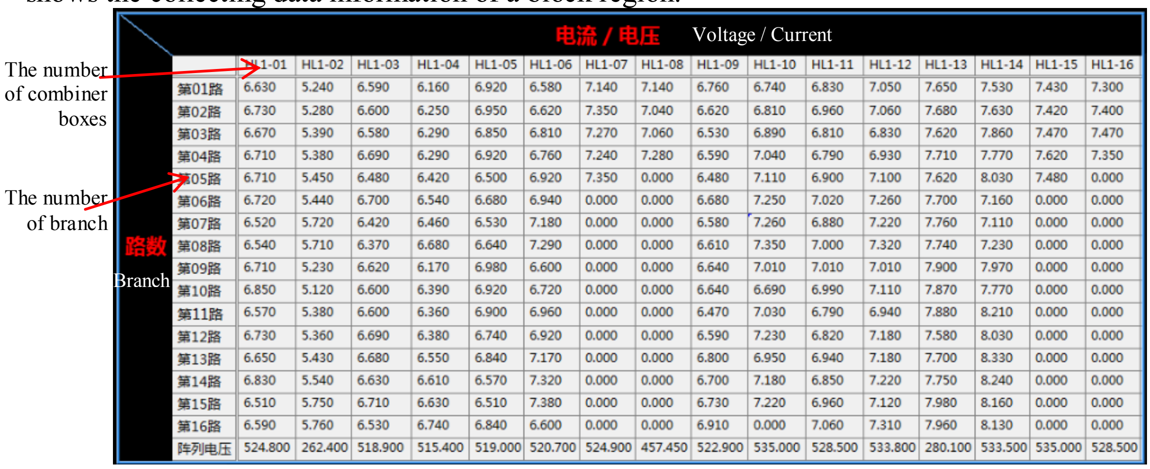
Figure 1. 20MW PV power station system structure

In the PV power station, the electric signal (including voltage and current) of combiner boxes and inverters can be collected and recorded. Through the analysis and processing of the history data, it's helpful for us to determine whether there is a fault in the PV plants.

2.2. *Global partitioned gradually approximation method*

When a fault occurs on PV panels, the collected voltage and current will become smaller than that of normal operation. Through the analysis of data, the disturbance of this electrical signal can provide some important information about the fault and its location [20]. And the fault diagnosis algorithm is implemented for large scale PV array according to the the disturbance.

By analyzing the basic structure of PV power station, the total number of PV modules in a 20MW PV power plant is over 10 million. When the data acquisition is completed and sent to the computer to determine the failure, all the data is processed at the same time, which will slow down the speed of the system and make the fault localization more difficult. Therefore, the matrix block processing method is employed to implement the modular processing of data. The data is divided into 20 modules with an inverter square as the unit. Each unit contains all the current data of the 16 combiner boxes. Figure 2 shows the collecting data information of a block region.



	HL1-01	HL1-02	HL1-03	HL1-04	HL1-05	HL1-06	HL1-07	HL1-08	HL1-09	HL1-10	HL1-11	HL1-12	HL1-13	HL1-14	HL1-15	HL1-16
第01路	6.630	5.240	6.590	6.160	6.920	6.580	7.140	7.140	6.760	6.740	6.830	7.050	7.650	7.530	7.430	7.300
第02路	6.730	5.280	6.600	6.250	6.950	6.620	7.350	7.040	6.620	6.810	6.960	7.060	7.680	7.630	7.420	7.400
第03路	6.670	5.390	6.580	6.290	6.850	6.810	7.270	7.060	6.530	6.890	6.810	6.830	7.620	7.860	7.470	7.470
第04路	6.710	5.380	6.690	6.290	6.920	6.760	7.240	7.280	6.590	7.040	6.790	6.930	7.710	7.770	7.620	7.350
第05路	6.710	5.450	6.480	6.420	6.500	6.920	7.350	0.000	6.480	7.110	6.900	7.100	7.620	8.030	7.480	0.000
第06路	6.720	5.440	6.700	6.540	6.680	6.940	0.000	0.000	6.680	7.250	7.020	7.260	7.700	7.160	0.000	0.000
第07路	6.520	5.720	6.420	6.460	6.530	7.180	0.000	0.000	6.580	7.260	6.880	7.220	7.760	7.110	0.000	0.000
第08路	6.540	5.710	6.370	6.680	6.640	7.290	0.000	0.000	6.610	7.350	7.000	7.320	7.740	7.230	0.000	0.000
第09路	6.710	5.230	6.620	6.170	6.980	6.600	0.000	0.000	6.640	7.010	7.010	7.010	7.900	7.970	0.000	0.000
第10路	6.850	5.120	6.600	6.390	6.920	6.720	0.000	0.000	6.640	6.690	6.990	7.110	7.870	7.770	0.000	0.000
第11路	6.570	5.380	6.600	6.360	6.900	6.960	0.000	0.000	6.470	7.030	6.790	6.940	7.880	8.210	0.000	0.000
第12路	6.730	5.360	6.690	6.380	6.740	6.920	0.000	0.000	6.590	7.230	6.820	7.180	7.580	8.030	0.000	0.000
第13路	6.650	5.430	6.680	6.550	6.840	7.170	0.000	0.000	6.800	6.950	6.940	7.180	7.700	8.330	0.000	0.000
第14路	6.830	5.540	6.630	6.610	6.570	7.320	0.000	0.000	6.700	7.180	6.850	7.220	7.750	8.240	0.000	0.000
第15路	6.510	5.750	6.710	6.630	6.510	7.380	0.000	0.000	6.730	7.220	6.960	7.120	7.980	8.160	0.000	0.000
第16路	6.590	5.760	6.530	6.740	6.840	6.600	0.000	0.000	6.910	0.000	7.060	7.310	7.960	8.130	0.000	0.000
阵列电压	524.800	262.400	518.900	515.400	519.000	520.700	524.900	457.450	522.900	535.000	528.500	533.800	280.100	533.500	535.000	528.500

Figure 2. Data information of a block region

According to the above analysis, all the collected data is treated as a M × N matrix A, and the data of each inverter is treated as 16×16 sub-matrix A_i (i=1, 2, 3,...... , 20). A and A_i are written as Equation (1) and (2).

$$A = \begin{bmatrix} [A_1]_{16\times16} & \cdots & [A_5]_{16\times16} \\ \cdots & \cdots & \cdots \\ [A_{11}]_{16\times16} & \cdots & [A_{20}]_{16\times16} \end{bmatrix}_{m\times n} \tag{1}$$

$$A_i = \begin{bmatrix} A_{pv11} & A_{pv12} & \cdots & A_{pv1c} \\ A_{pv21} & A_{pv12} & \cdots & A_{pv2c} \\ \cdots & \cdots & A_{pvrl} & \cdots \\ A_{pvb1} & A_{pvb2} & \cdots & A_{pvbc} \end{bmatrix}_{b\times c} \tag{2}$$

In Equation (2), $b=c=16$, $r=1\sim b$, $l=1\sim c$.

In this method, the relative error between the average value of the A_i and all data A_{pvrl} ($r=1,2,\ldots,b$, $l=1,2,\ldots,c$) in A_i is calculated in Equation (3). The relative error is the important parameter for judging whether there is a fault, which is called Fault weight factor δ.

$$\delta = \left| \frac{A_{\arg} - A_{pvrl}}{A_{\arg}} \right| \tag{3}$$

In Equation (3),

$A_{\arg} = A_{bc} / (b \times c)$;

$A_{bc} = \sum A_{pv11} + A_{pv12} + ... + A_{pv1c} + ... + A_{pvb1} + A_{pvb2} + ... + A_{pvbc}$.

The data A_{pvrl} has different properties for the different states. If there is no fault in the array, the branch current data A_{pvrl} is allowed to fluctuate within a small range. The small range is artificially set value, which is called fault threshold δ_0. If in a failure state, the deviation between the data A_{pvrl} in the fault branch and the average value A_{arg} is greater than fault threshold δ_0.

A fault function is f. $\delta \geq \delta_0$, which means the system has a fault, f=1. $\delta < \delta_0$, which means the system is normal, f=0. The conditions for determining the fault is written as Equation (4).

$$f = \begin{cases} 1 & \delta \geq \delta_0 \\ 0 & \delta < \delta_0 \end{cases} \tag{4}$$

The PV plant is divided into finite units, and the fault of each unit can be determined by Equation (4). The fault positions are located by searching all the collected data from the PV plant. The global partitioned gradually approximation algorithm is implemented with Matlab software, and its flow chart is shown in Figure 3.

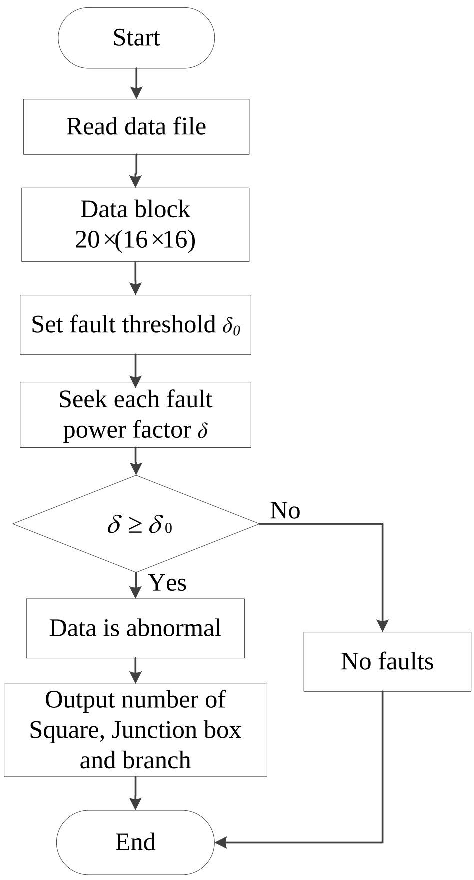
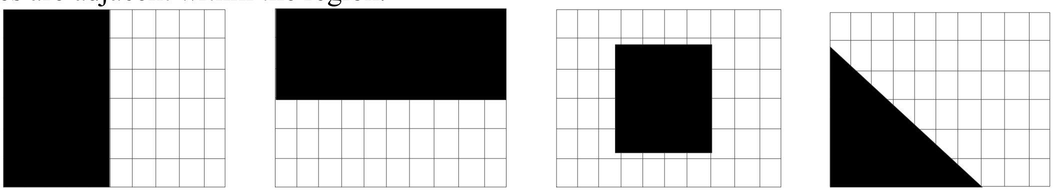


Figure 3. Flow chart of programmed algorithm

2.3. Shade determination

For large PV power plants, the shade is a very common phenomenon. The shade makes data measured of the shaded panels smaller than the normal state, which is similar to fault panels. But it is not a fault type [21]. Thus this phenomenon possibly leads to the mistake of fault detection with the proposed algorithm. So the shade situation should be considered for improving the validity of the GPGA algorithm.

As shown in Figure 4, four types of shade phenomenon are diagramed to illustrate the shade arrangement. Through the shade arrangement, the shade phenomenon has the following characteristics: (1) there are a larger number of shaded PV modules within the region; (2) most of the obscured PV modules are adjacent within the region.



(a) (b) (c) (d)

Figure 4. Shade situation analysis

The total number of PV units in an area is u, the number of shaded units is v, the number of adjacent shaded modules is w. The algorithm for determining the shaded panels is established as shown in Figure 5.

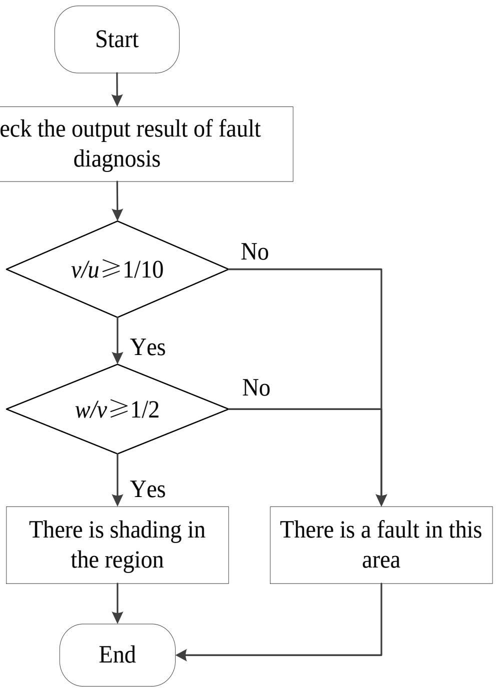


Figure 5. Flow chart of shading decision algorithm

The GPGA method combined with the shading decision algorithm is used to implement fault detecting and locating when the PV panels are shaded by the clouds. The shade-determining conditions are summarized as follows:

(1) $v/u \geqslant 1/10$ in the shading decision diagram.

(2) $w/v \geqslant 1/2$ in the shading decision diagram.

A shade function is y, when $v/u \geqslant 1/10 \cap w/v \geqslant 1/2$, the system has a shade, y=1. In other conditions, the system has no a fault, y=0. The shade-determining conditions are written as

$$y=\begin{cases}1 & v/u\geq 1/10\,\mathrm{I}\;\; w/v\geq 1/2\\ 0 & other\quad conditions\end{cases} \qquad (5)$$

2.4. *Determination algorithm of fault diagnosis*

The GPGA method is used to judge whether there is a fault, and the shading decision algorithm is employed to judge whether there is a shade. When f=1∩y=1, there is a shade in the region. When f=1∩y=0, there is a fault in the region.The flow chart of fault diagnosis algorithm is shown in Figure 6.

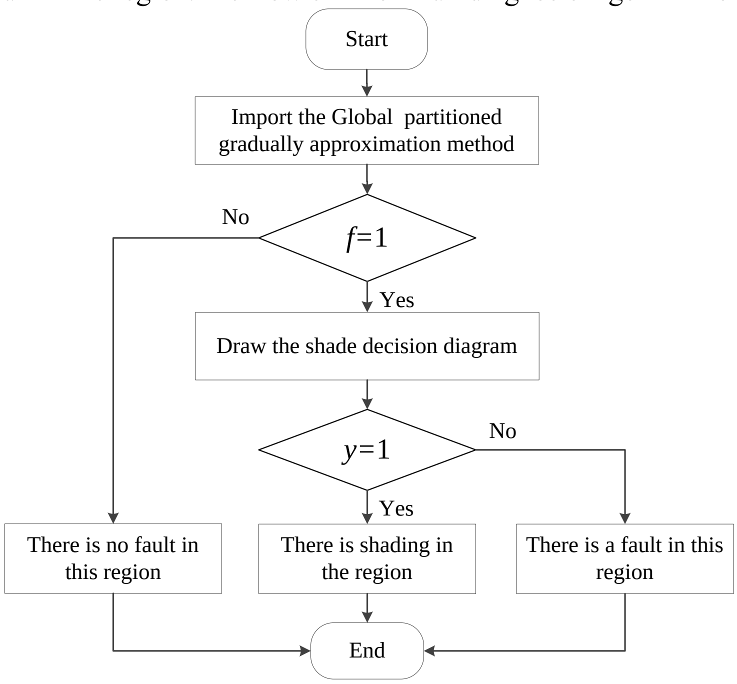


Figure 6. Flow chart of fault diagnosis algorithm

3. Simulation system design

3.1. *System function analysis for PV power station*

This simulation system is designed combing with the actual demand for PV power station. It has the following major functions.

1) Overall structure of PV power plant. The simulation system shows the overall structure block diagram of 20MW PV power station. The parameters, including the voltage, current, active power and the real-time power, are displayed with simulation system. The smallest unit in the structure block diagram is inverter square.

2) Real-time displaying for link data. The displayed data include voltage and current of each branch of the combiner box, basic data amount of the inverter power generation (the phase current, the phase voltage, total power, active power, reactive power, power factor, accumulated generated energy), the daily generated energy, temperature, solar irradiation, wind speed and so on.

3) History data viewing online. The daily power, daily accumulated energy, the daily inverter power, the environmental parameters are recorded and viewed in the simulation system.

4) Data statistics. This function is to analyze the data collected from the sensors, and plot the curves for the inverter and combiner box under different time periods.

5) Report generation. This function is to get the history data information. Through the choice of date, the data information of the inverter can be saved. Data font, font size, format, save type, save path and other basic settings also can be set in the function.

6) Power prediction. Based on the analysis of the history big data, daily power electricity is predicted. In order to view clearly, the system has the function of zoom in and out about power prediction curve, as well as the upper limit curve and lower limit curve.

7) Fault diagnosis. The global partitioned gradually approximation method is embedded into the simulation system, with which the fault detection and location of large-scale PV array is implemented.

The overall structure of the monitoring simulation system is shown in Figure 7.

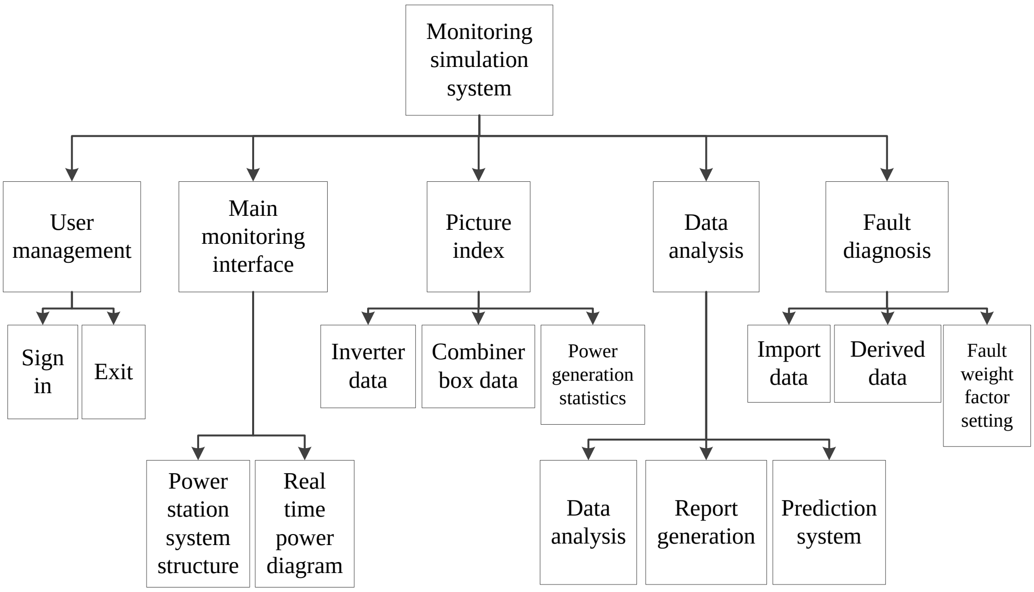


Figure 7. The overall structure of the monitoring simulation system

3.2. Simulation system design

The Qinghai-Tibet Plateau is one of the most abundant areas in the world. Because of the high altitude, the sunshine is sufficient. The amount of solar radiation in Qaidam Basin is more than $6800MJ/m^2$, in which the desertification land used for PV power station gets to100 thousand square kilometers. It is the best area of solar photovoltaic power generation in China, because its average annual sunshine gets to 3000 hours. The Golmud City is located in the hinterland of Qinghai-Tibet Plateau, and the city is located in Qaidam Basin at an altitude of 2780 meters. The actual data of a Golmud 20MW PV power station is used as the data source of the simulation system.

According to the requirement of the user and the function of the PV power station system, the development and design of the fault diagnosis simulation system based on GPGA method is carried out on LabVIEW. The simulation system mainly includes user management, data real-time display, storage management, data statistics, data prediction and fault diagnosis.

The following figures are a part of the application interface of fault diagnosis simulation system. Figure 8 shows the user login screen. The correct user name and password is necessary to enter the system main interface for system security.

Figure 8. The user login interface

The system main interface is shown in Figure 9. The main interface is composed of three parts including the menu switch control, the tree structure of PV power station and the real-time power display window. The system menu includes the main interface display, the interface index, data analysis and fault diagnosis. The data could be acquired in real-time, and the data includes the current, voltage and active power. The frequency of the acquitted data is $1.67 \times 10^{-3} s^{-1}$. The real-time data power curve is displayed in the display window, and the history data information can be observed at the same time. The data information in display window includes power, solar irradiation, ambient temperature and wind speed.

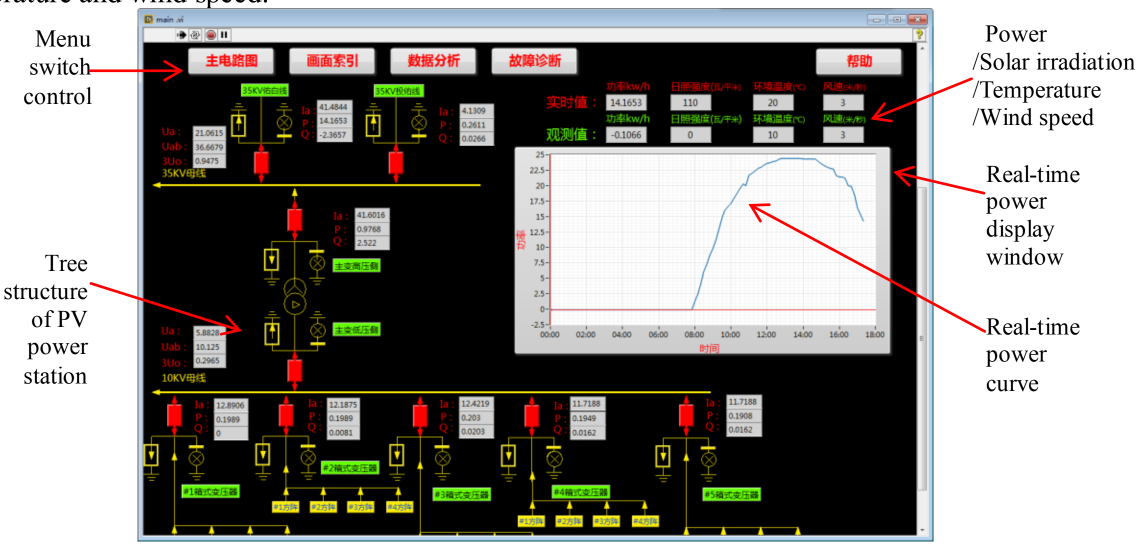


Figure 9. System main interface

Figure 10 gives the specific fault diagnosis interface. The interface has the following functions: 1) setting the fault threshold, 2) importing and exporting the data, 3) fault display. The data is calculated by LabVIEW program which is based on Global partitioned gradually approximation method. Its results show as “There is a fault in inverter No. ***, combiner box No. ***, branch No. ***.” .

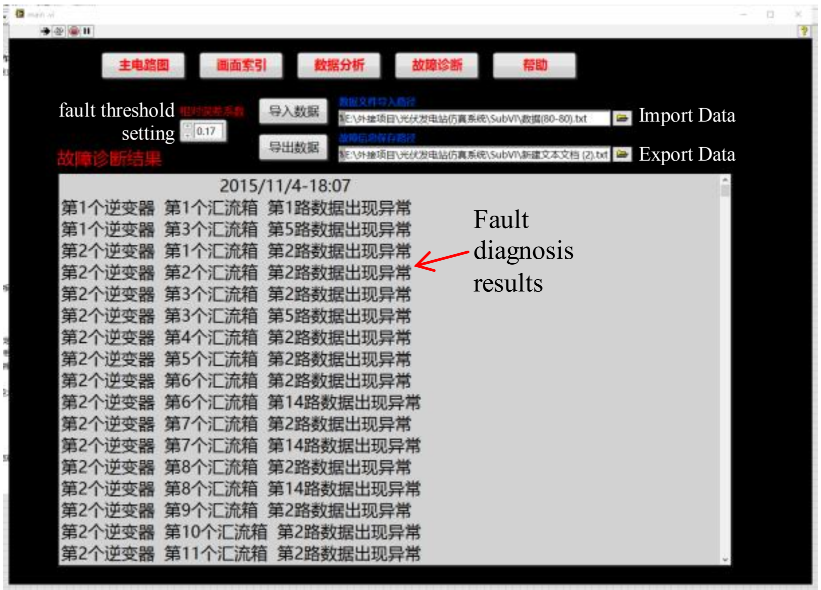


Figure 10. Fault diagnosis interface

4. Result and discussion

4.1. Simulation result of fault diagnosis algorithm

Some collected data is employed to validate the proposed fault diagnosis algorithm. The data is acquired from the PV power station in Glomud City. The data includes voltage and current of all the combiner boxes in July 11 to 13. And some data is modified as the faults to test and verify the proposed fault diagnosis algorithm. In this experiment, the fault threshold is set as 6 types, 5%, 7%, 10%, 12%, 15%, and 20%. Through calculating, the result of fault diagnosis is seen at the interface. Table 1 shows the accuracy of fault detection under different fault threshold.

Table 1. Fault detection accuracy rate under different fault threshold (δ_0)

	The original data fault points/a	Detect fault points/a	accuracy rate /%
δ_0=5%	100	210	---
δ_0=7%	100	203	---
δ_0=10%	100	100	100
δ_0=12%	100	100	100
δ_0=15%	100	94	94.0
δ_0=20%	100	88	88.0

The simulation results show that the number of abnormal data points is much larger than the actual fault points when the fault threshold is less than 10%. The normal data can be regarded as the fault easily. When the fault threshold is in 10% - 12%, the fault detection accuracy rate is 100%. When the fault threshold continues to increase, the accuracy rate of fault detection will decrease.

Finally the results of fault diagnosis are used for judging whether there is a fault occurred in some regions.

(1)If the abnormal data does exist in a region, f=1, and its number is more than 1/10 of the total number, shading decision algorithm is used for determining whether there is a shade or a fault. If more

IOP Conf. Series: Earth and Environmental Science **40** (2016) 012030 doi:10.1088/1755-1315/40/1/012030

than 1/2 of the abnormal data are adjacent, y=1, it can be determined that there is shade in the region, not a fault.

(2) If the abnormal data does exist in a region, f=1, and its number is less than 1/10 of the total number, and most of the abnormal data are nonadjacent, y=0, we can determine that there is fault in the region.

Taking the 2nd square inverter as an example, the shade decision algorithm is used for detecting the fault and the results are showed in Figure 10. The Figure 11 displays the shading decision algorithm (In Figure 11, A denotes the number of combiner box, and B denotes the number of breach). The branch number of an inverter is 256, and the abnormal data branch number is 52 (v/u=1/5). Then we found that more than 1/2 of the abnormal data points are adjacent, so we can determine the shade in the 2nd inverter square. But the 5th branch of the 3rd combiner box, the 6th branch of the 14th combiner box and 4th branch of the 7th combiner box are in fault.

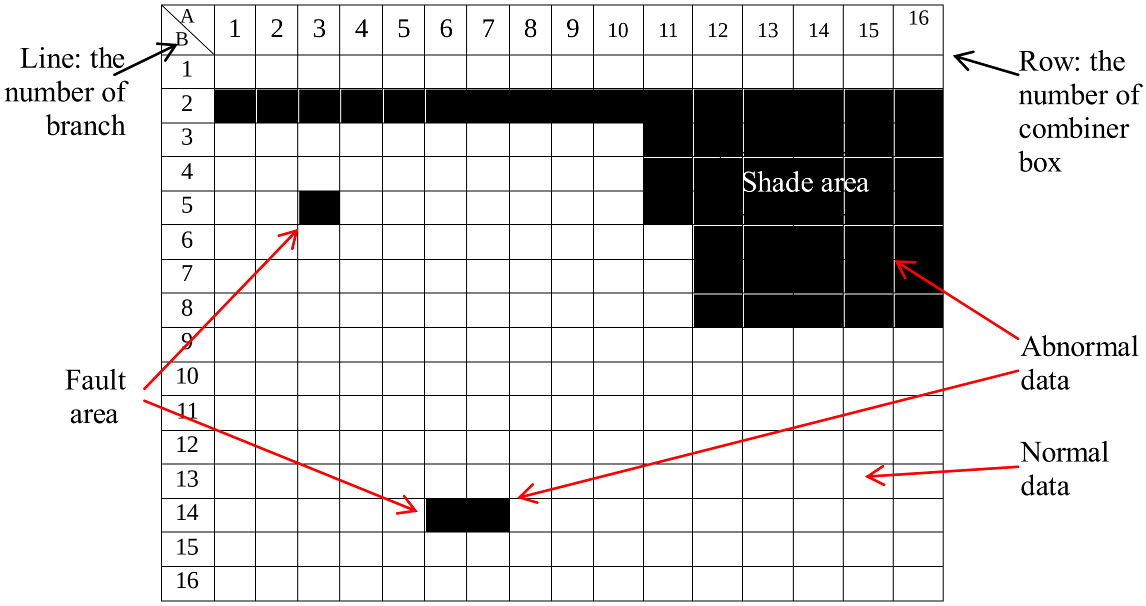


Figure 11. Shade decision diagram of the 2nd square inverter

4.2. Losses measure of PV power station

The power losses occur on the PV power station due to the low fault diagnosis accuracy rate and slower troubleshooting. If the fault is found and maintained in time, the power losses mainly depend on the generated energy of fault branches which are not detected.

20 same (240W) PV panels in series form a branch, and the power of one branch is measured about 4KW (it less than the rated power in normal operation). The total power of this PV power station is 20MW. The different power losses are calculated at different fault detection accuracy rate. The calculated results are shown in Table 2.

Table 2. Power losses under different fault detection accuracy rate

accuracy rate /%	The number of un-detected fault branch	Power losses/KW	Percentage/%
100	0	0	0
94	6	24	0.12
88	12	48	0.24
82	18	72	0.36
76	24	96	0.48

Through the calculating and contrast, the result shows that power losses increases with the loss fault detection accuracy rate. So, fault diagnosis algorithm of high accuracy is helpful to reduce power losses.

5. Conclusions

A new fault diagnosis algorithm, global partitioned gradually approximation method, is proposed for the large-scale PV arrays in high-altitude desert area. The proposed fault diagnosis algorithm takes account for the shade affects, which reduces the misjudgment probability of the algorithm. A system for PV power station monitoring and fault detection is designed and implemented with LabVIEW software based on the new fault diagnosis algorithm. The function of the system includes real time data display, online-browse, data statistics, and fault diagnosis. The proposed fault diagnosis algorithm can predict and locate the fault arrays through the experiments. The developed system can be applied in the large-scale PV plants without changing any original hardware and infrastructure.

Acknowledgments

This research work is sponsored in part by the National Natural Science Foundation of China (No. 51566015), and the Projects from Qinghai Science &Technology Department (No. 2012-G-Y25A-4 and 2014-GX-212).

References

[1] Kelly N A and Gibson T L 2009 Improved PV energy output for cloudy conditions with a solar tracking system *Solar Energy* **83** 2092-2102
[2] Wang Y, Li Z and Wu C 2013 Fault diagnosis technologies for photovoltaic system *Chinese Journal of Power Sources* **37** 1700-05
[3] Cheng Z, Li B F, Liu L and Liu Y L 2010 A fault detection method for new PV module strufure *Journal of Electronic Measurement and Instrument* **24** 131-6
[4] Liu Y and Li B 2010 Research on domestic PV module structure based on fault detection *Intelligent Control and Automation* **20** 171-5
[5] Wang Y Z, Wu C H, Zhou D Q, Fu L, Li Z H. 2013 A survey of fault diagnosis for PV array based on BP neural network *Power System Protection and Control* 41 108-14
[6] Hu Y H, Chen H, Xu R D, Li R. 2011 Photovoltaic(PV) array fault diagnosis strategy based on optimal sensor placement *Proceedings of the Chinese Society of Electrical Engineering* **31** 19-30
[7] Hu Y, Chen H and Xu R 2010 A type of PV array fault diagnosis strategy based on voltage scan *Proceedings of the Chinese Society of Electrical Engineering* **30** 185-91
[8] Hu Y, Deng Y and He X 2013 A summary on PV array fault diagnosis method *Power Electronics* **47** 21-3
[9] Takashima T, Yamaguchi J, Otani K, Oozeki T, Kato K, Ishida M. 2009 Experimental studies of fault location in PV module strings *Solar Energy Materials and Solar Cells* **93** 1079-82
[10] Patel H and Agarwal V 2008 MATLAB-based modeling to study the effects of partial shading on PV array characteristics *IEEE Transactions on Energy Conversion* **23** 302-10
[11] Lan Q and Zhang Y 2011 Application of BP neural network in fault diagnosis of solar photovoltaic system *Telecom Power Technology* **28** 38-40
[12] Zhong D 2011 *Research on fault detection of PV array based on composite data fusion* (Tianjin: Tian Jin University)
[13] Tang J 2012 *The fault detection and simulation research of solar array* (Beijing: Noah China Electric Power University)
[14] Zhang H 2006 *The design and research of distant monitoring system for the PV plant* (Qinhuangdao: Yan Shan University)
[15] Wang K 2013 *Design and implementation of embedded monitoring system for PV power plant* (Bejing: Beijing Jiaotong University)

[16] Zhang X and Zheng J 2010 Design of monitoring system for photovoltaic power station *Electrotechnics Electric* **9** 12-6

[17] Miao J 2012 *The research and implementation of monitoring system for PV power plant* (Beijing: Beijing Jiaotong University)

[18] Zhao Y, Meng C, Chen Y, Sun Y. 2012 Photovoltaic Power Station Monitoring System Based on LabVIEW *Instrument Technique and Sensor* 50-3

[19] Zhang W 2014 *The research and implementation of PV's remote monitoring system* (Yangzhou: Yang Zhou Univesity)

[20] Zhang X, Gao D, Liu H, Ye J, Wang S. 2016 A noval method for multi-sensor fault detection and positioning of PV array *Renewable Energy Resources* **34** 166-72

[21] Chowdhury S R and Saha H 2010 Maximum power point tracking of partially shaded solar photovoltaic arrays *Solar Energy Materials & Solar Cells* **94** 1441-7