

ORIGINAL RESEARCH

Open Access

A power fluctuation evaluation method of PV plants based on RankBoost ranking



Weiye Xia, Zhouyang Ren* , Hui Li and Bo Hu

Abstract

Fluctuation evaluation is an important task in promoting the accommodation of photovoltaic (PV) power generation. This paper proposes an evaluation method to quantify the power fluctuation of PV plants. This consists of an index system and a ranking method based on the RankBoost algorithm. Eleven indices are devised and included in the index system to fully cover diverse fluctuation features. By handling missing and invalid data effectively, the ranking method fuses multiple indices automatically and provides a systematic and comprehensive comparison of power fluctuation. Simulation results based on power data from six PV plants indicate that the evaluation list obtained by the RankBoost ranking method is better represented and more comprehensive than that derived by the equal weight method.

Keywords: Fluctuation evaluation, Photovoltaic power, RankBoost ranking algorithm

1 Introduction

With the aggravation of the energy crisis and environmental pollution, PV power, as one of the most important renewable resources, has developed rapidly and attracted significant attention [1]. Nevertheless, the uncertainty and fluctuation of PV power negatively impact on power quality and power system reliability, and bring huge challenges for energy dispatching and renewable energy accommodation [2]. PV plants differ in their power fluctuation and exerted influence on a power system, and this leads to varying preferences of different PV plants by operators and market participants. Hence, it is necessary to quantify and evaluate fluctuation characteristics, and obtain the comparison and ranking results of different PV plants in their power fluctuation. For power system operators and electricity market participants, a ranking list can provide suggestions on the priority of PV power to be accommodated or traded.

Fluctuation evaluation consists of a quantified index system and a comprehensive fusion method for individual indices. Adequate selection of evaluation indices is of

great significance for reducing computational complexity and improving evaluation credibility. The indices used in PV power evaluation can be sorted into numerical indices, probability indices [3, 4] and time indices. These indices also vary in ranges and can be classified into overall indices or local indices. The mean [1] of PV power gives an overview of the power level while the variance and solar variability focus on the overall fluctuation features. And apart from the numerical description of fluctuation, skewness and kurtosis [1] denote the probability distribution of PV plants. Since the probability distribution of PV power usually shows a stark departure from a normal distribution, it is skewness and kurtosis are inappropriate for measuring the difference of probability distribution. It is more appropriate to depict the probability distribution referring to a typical pattern of PV probability distribution. Paying more attention to the local features, fluctuation ratio or ramp ratio [5, 6] is proposed to quantify the fluctuation at adjacent moments in the time series. However, these two indices contain excessive information on the fluctuation such that they fail to fulfill a single and synthetic evaluation based on a large amount of historical data. To produce an overall evaluation of PV plants, the deviation ratio and sum of gradient [2] and an overall fluctuation tendency index K_{fluc} of

* Correspondence: rzhouyang1108@163.com

The State Key Laboratory of Power Transmission Equipment & System Security and New Technology, School of Electrical Engineering, Chongqing University, Chongqing 400044, PR China

solar radiation is proposed in [7]. However, some extreme and unusual features have been missed. Reference [8] considers the time feature of the solar irradiance by adopting the time of high-level/low-level irradiance to measure the controlled solar irradiance fluctuation. However, it is difficult in practice to set a suitable threshold to PV power between the high-level and low-level. It is rather more practical to discuss zero output and nonzero output.

Any single index is insufficient to characterize the fluctuation of PV power. However, none of the above described literatures has formed a complete, scientific, and reasonable index system for fluctuation evaluation of PV power. In addition, the impact of PV fluctuation on load regulation is of great concern, and thus, quantified indices denoting the ability of load regulation need to be investigated.

The fluctuation analysis and quantization method can be generally divided into the following four groups:

1. Grading and classifying the fluctuation based on a particular index is used in reference [5] where the high and low fluctuations to PV plants based on the ramp rate are allocated. However, although it seems the quantified indices is employed, it returns to qualitative analysis after quantification.
2. Probability models are used to analysis the probability feature of PV power, such as the Gaussian mixture or empirical probability model [8]. It helps to form the overview of all the power data, but it lacks the information of order and fails to analyze the fluctuation of two adjacent sampling points.
3. The Fourier transform and wavelet transform methods are introduced to quantify the fluctuation on different time scales in the wavelet spectrum [9]. Although it differentiates the diffusive and jumpy characteristics of PV power with this method, it is complex and not intuitive enough for the operators and market participants since the raw data has been transformed.
4. A typical step waveform [10] is proposed to simulate the solar irradiance. This method is a theoretical account of fluctuation and it is also not intuitive enough for the understanding of fluctuation.

On the comprehensive combination of multiple indices, the most widely used one is to allocate a weight for every index and summarize all the results to obtain an overall score. However, subjectively selecting weights for different parameters is extremely difficult in practice. Reference [11, 12] puts forward a minimum variance and weighted geometric averaging operator to determine

the weight of each indicator, while reference [5, 12] determines the weight of each index by the deviation and mean value. However, the method of artificially allocating weights lacks theoretical foundation. With the latest development of machine learning, the RankBoost algorithm has been employed to produce the maintenance priority for breakers in a power system by combining multiple indices with different ranges and dimensions [13, 14].

The main contributions of this study on PV output evaluation are as follows:

1. It provides an in-depth study on the power fluctuation features.
2. Considering characteristics of PV output and PV operational issues, a systematic and complete index system is inaugurated to quantify power fluctuation from its three different aspects.
3. An effective approach is developed to quarry out invalid data or errors in PV power data collection.
4. The machine learning method named RankBoost is applied to produce a synthetic evaluation by automatically fusing the indices with different ranges and dimensions.

The completeness and rationality of the proposed evaluation method are discussed through a case study in Section 3.

2 Methods

This section presents the index system and the RankBoost ranking method employed to evaluate the power fluctuation by elucidating the characteristics of PV power.

2.1 The characteristics of PV power

PV power is more deterministic than wind power at the daily scale. Influenced by the day cycle of solar irradiance, the day curve of PV power assumes a unimodal pattern and its shape is similar to the positive semi period of the cosine curve as shown in Fig. 1. PV power usually rises around sunrise, reaches a maximum around midday and ultimately falls to zero near sunset. The earliest time within a day when PV power sets to rise from zero is defined as the initiation time, while the time when PV power falls to zero is the ending time. The period between the initiation time and ending time is called the effective period of PV power. And if there is a day when PV output keeps zero or negative all day long, the the initiation time, ending time and effective period do not exist and the output data of this day would not be included in fluctuation evaluation.

The fluctuation of PV power is reflected in the effective period, power and electricity generation. The

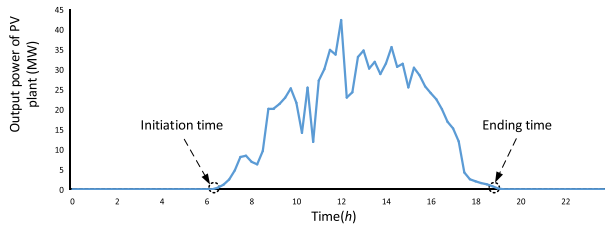


Fig. 1 Initiation time and ending time of PV power

initiation time, ending time and effective period are all variables since there are some horizontal or transverse movements in the power day curve. In addition, vertical movement in the power day curve should also be measured and included in the index system. Finally, integral variation of the power day curve represents the fluctuation of electricity generated by the PV plants. Given these three aspects, eleven indices are proposed to demonstrate and assess the features of PV power fluctuation.

2.2 Index system for PV power fluctuation evaluation

The proposed index system has the following advantages: 1) The fluctuation characteristics are divided into three aspects, providing a systematic and comprehensive perspective on the fluctuation. 2) Every aspect of the fluctuation is clearly described and quantified by certain indices. 3) Every index is intuitively clear, brief, and easy to understand by power system operators. 4) The fluctuation is measured by comparing not only with the adjacent sampling point but also with the load fluctuation, as indices measuring the impact of PV power on load regulation are also designed.

Since fluctuation must be discussed based on reference values, different reference values are designed in different aspects of fluctuation. The fluctuation of effective time is compared with the times of sunrise and sunset, whereas both the power and the electricity generation are compared with their adjacent sampling points and load. The index system involving three aspects of fluctuation is shown in Fig. 2 and will be illustrated in detail in the following.

(1) Fluctuation evaluation of effective time

Since solar radiation exerts a great effect on PV power, the initiation time and ending time of PV power have close relationships with the times of sunrise and sunset. In addition, the sunrise and sunset times in different regions are usually recorded and accessible to researchers. Hence, short time displacement rate (STDR) is proposed to measure the probability of negligible differences between the initiation time and the sunrise time or between the ending time and sunset time as:

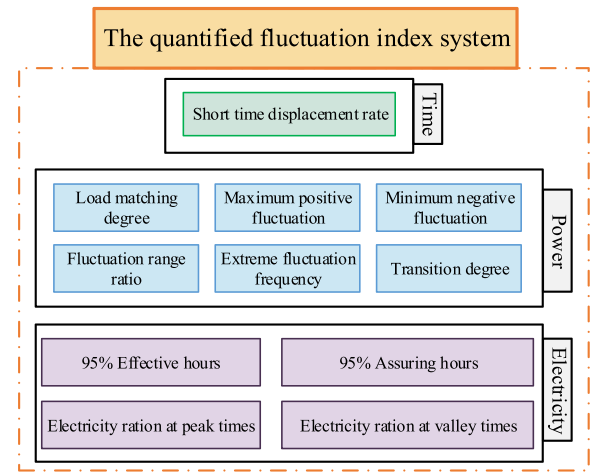


Fig. 2 Structure diagram of the index

$$TD_{starting} = t_{initiation} - t_{up} \tag{1}$$

$$TD_{terminal} = t_{ending} - t_{down} \tag{2}$$

where t_{up} and t_{down} represent the sunrise and sunset times, respectively. $t_{initiation}$ denotes the initiation time of the PV plant while t_{ending} is the ending time of the PV plant. $TD_{starting}$ and $TD_{terminal}$ are the initiation time displacement and the ending time displacement, respectively. If $TD_{starting}$ or $TD_{terminal}$ is positive, it indicates that the initiation time or ending time lags the sunrise or sunset. Otherwise, the initiation time or ending time precedes the sunrise or sunset. To evaluate the probability distributions of the time displacement, 15 min is chosen as the threshold value to calculate STDR:

$$STDR = \frac{d[|TD_{starting}| < 15 \text{ min and } |TD_{terminal}| < 15 \text{ min}]}{2D} \tag{3}$$

where D is the number of days that are calculated and evaluated, and $F_D[con]$ is a function obtaining the number of days in which the condition con is met. This index helps to evaluate the uncertainty and extreme fluctuation of effective period and the rationality of ignoring the gaps between the initiation time and the sunrise or the ending time and sunset. The more likely time displacement is less than 15 min, the more accurately to predict or describe the initiation time and ending time by the sunrise time and sunset time.

(2) Fluctuation evaluation of PV power

A. Fluctuation range ratio

Fluctuation range ratio is given as:

$$f_r = \frac{l}{P_{\max}} = \frac{\sum_{i=1}^N |\Delta P_i|}{P_{\max}} = \frac{\sum_{i=2}^N |P_i - P_{i-1}|}{P_{\max}} \quad (4)$$

where N denotes the number of sampling points in one day. P_i represents the power of the PV plant at the sampling point of i , while P_{\max} is the maximum P_i in one day. The fluctuation range ratio f_r reveals the overall level of the fluctuation. This can provide a quantified overview of the fluctuation range for operators. For example, $f_r = 2$ indicates that the day curve of power assumes a perfect unimodal pattern which means the PV power keeps rising before reaching the peak and then continuously dropping without any oscillation. The larger f_r is, the sharper the fluctuation is. Normalized by the P_{\max} , fluctuation range ratio excludes overall rising and falling tendency from the fluctuation, which is suitable and reasonable for fluctuation evaluation for PV power.

B. Maximum positive fluctuation and minimum negative fluctuation

Maximum positive fluctuation ΔP_{\max} and minimum negative fluctuation ΔP_{\min} are given as:

$$\Delta P_i = P_i - P_{i-1} \quad (5)$$

$$\Delta P_{\max} = \max\{\Delta P_i, i = 2, 3, \dots, N\} \quad (6)$$

$$\Delta P_{\min} = \min\{\Delta P_i, i = 2, 3, \dots, N\} \quad (7)$$

ΔP_{\max} and ΔP_{\min} demonstrate the extreme degree of the fluctuation, which can help operators evaluate whether the range of PV plant fluctuation is acceptable for the power system and examine the obstacles it could bring to power dispatching. Normally, PV plant power is either positive or zero. However, some auxiliary power is consumed so the minimum negative power denotes the maximum auxiliary power that the PV plant draws from the grid.

C. Transition degree

As illustrated before, fluctuation range ratio provides an overview of the fluctuation range. Nevertheless, for real-time power dispatching, the variation of adjacent sample time is also of great interest. Thus, to evaluate the transition frequency of PV power, the coefficient of transition D_T is illustrated as:

$$D_T = \frac{\sum_{i=1}^{N-1} F_P(P_i)}{N-1} \quad (8)$$

$$F_P(P_i) = \begin{cases} 1 & (P_i - P_{i-1})(P_i - P_{i+1}) > 0 \\ 0 & (P_i - P_{i-1})(P_i - P_{i+1}) \leq 0 \end{cases} \quad (9)$$

If $F_P(P_i) > 0$, it means that the PV power changes direction and reaches an extreme at sampling time i . The larger D_T is, the more frequently the PV power changes its direction. This means the PV power is more fluctuating and unpredictable. Thus, the transition degree represents an overall view of how PV power changes in fluctuation direction.

D. Load matching degree

Power system load is of varying and fluctuating nature and exhibits some daily or seasonal regularity. The fluctuation of load and its consistency of the trends with PV power have close relationships with the difficulty of dispatching, as:

$$D_{LM} = \frac{\sum_{i=1}^{N-1} F_L(P_i, L_i)}{N-1} \quad (10)$$

$$F_L(P_i, L_i) = \begin{cases} 1 & (P_{i+1} - P_i)(L_{i+1} - L_i) > 0 \\ 0 & (P_{i+1} - P_i)(L_{i+1} - L_i) \leq 0 \end{cases} \quad (11)$$

where L_i is the total power of the load in the region of PV plants. Load matching helps operators assess the difficulty of peak modulation that the fluctuation of PV plant could bring after being connected to the system. $F_L(\)$ is a defined function which identifies the trend consistency of PV outputs and the load [1].

E. Extreme fluctuation frequency

As mentioned above, the extreme boundary of fluctuation is reflected by the maximum positive fluctuation and minimum negative fluctuation. However, the frequency of extreme margin of fluctuation is also of interest to power system operators. Thus, the extreme fluctuation frequency can be given as:

$$d_+ = d_+ [\Delta P_{i, \max} > (\Delta P_{\max} - \Delta P_{avg+}) \times c + \Delta P_{avg+}] \quad (12)$$

$$d_- = d_- [\Delta P_{i, \min} < (\Delta P_{\min} - \Delta P_{avg-}) \times c + \Delta P_{avg-}] \quad (13)$$

$$f_e = \frac{d_+ + d_-}{D} \quad (14)$$

where ΔP_{avg+} is the annual average of the daily

maximum positive fluctuation and ΔP_{avg-} represents the annual average of the daily minimum negative fluctuation. $\Delta P_{i,max}$ and $\Delta P_{i,min}$ are the maximum positive fluctuation and the minimum fluctuation at day i , respectively. d_+ represents the number of days in which the power of the PV plant meets the condition of $\Delta P_{i,max} > (\Delta P_{max} - \Delta P_{avg+}) \times c + \Delta P_{avg+}$, while d_- represents the number of days in which the power of the PV plant meets the condition of $\Delta P_{i,min} < (\Delta P_{min} - \Delta P_{avg-}) \times c + \Delta P_{avg-}$. c is a constant and is assigned to be 80% in this study but can be adjusted according to the acceptability of PV power fluctuation. f_e demonstrates the frequency of the extreme fluctuation in 1 year, which enables operators to evaluate the difficulty of PV power dispatching and prepare for extreme high or low power.

(3) The fluctuation evaluation of electricity

A. 95% Effective hours and 95% assuring hours

The daily electricity generated by PV plant can be calculated as:

$$E_d = \int_{t_{initiation}}^{t_{ending}} P(t)dt, \tag{15}$$

where d is the sequence number of days in 1 year and $P(t)$ is the power of the PV plant at time t in day d .

E_d ($d = 1, 2, \dots, D$) can be sorted into a descending sequence and a curve can be drawn to describe the decreasing characteristics of the sequence, as shown in Fig. 3. This figure is called the annual sustained electricity curve and the abscissa is the accrued probability of electricity generation. The function of the curve can be represented as $E(p)$ and p is the accrued probability. The 95% effective hours and 95% assuring hours are defined as:

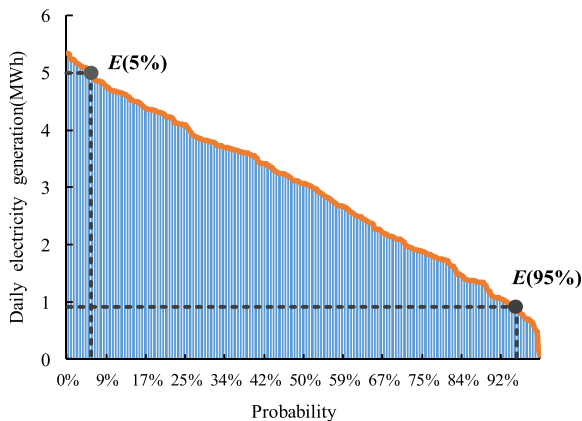


Fig. 3 Annual sustained electricity curve

$$95\%EH = \frac{E(95\%)}{P_N} \tag{16}$$

$$95\%AH = \frac{E(95\%)}{P_N} \tag{17}$$

where P_N is the rated power of the PV plant. $E(5\%)$ means there is 5% probability that daily electricity generated by the PV plant exceeds the $E(5\%)$. In other words, there is 95% probability that daily electricity generated by the PV plant is less than the $E(5\%)$. Similarly, there is 95% probability that daily electricity generated by the PV plant is above the $E(95\%)$. $E(5\%)$ and $E(95\%)$ evaluate the top and bottom boundaries of effective and reliable electricity. The 95% effective hours and 95% assuring hours denote the daily power generation of the PV plant from the perspective of probability. The 95% EH and 95% AH are normalized by the P_N to eliminate the influence of plant size. The 95% EH helps operators evaluate the effective utilization of the PV plant, while the 95% AH helps operators to have a fair idea of assuring power generation which can be considered as a constant power generated by the PV plant despite the fluctuation of daily power.

B. Electricity ratios at peak and valley

To reflect the impact of PV power on load modulation, the load matching degree has been proposed to calculate the consistency of the trends of the PV power and the load. However, the difficulty of load modulation is also associated with the electricity generation, and higher electricity generation by the PV plant at peak time means less difficulty for the power system to modulate the load. Before proposing the electricity ratios at peak and valley, the definitions of the peak period and valley period are described.

Considering L_{avg} is the average power of load in 1 day and $L(t)$ is the total power of the load in the region of the PV plant at time t , for $t \in [t_{1i}, t_{2i}]$, the time period $t_{1i} \sim t_{2i}$ is called the peak period if $L(t) > L_{avg}$, or the valley period if $L(t) < L_{avg}$. The electricity ratios at peak and valley are thus given as:

$$ERP = \begin{cases} \frac{\sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} P(t)dt}{\sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} (L(t) - L_{avg})dt} & \sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} (L(t) - L_{avg})dt \neq 0 \\ 0 & \sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} (L(t) - L_{avg})dt = 0 \end{cases} \tag{18}$$

$$ERV = \begin{cases} \frac{\sum_{j=1}^{N_p} \int_{t_{1j}}^{t_{2j}} P(t) dt}{\sum_{j=1}^{N_p} \int_{t_{1j}}^{t_{2j}} (L(t) - L_{avg}) dt} & \sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} (L(t) - L_{avg}) dt \neq 0 \\ 0 & \sum_{i=1}^{N_p} \int_{t_{1i}}^{t_{2i}} (L(t) - L_{avg}) dt = 0 \end{cases} \quad (19)$$

where *ERP* represents the electricity ratio at peak and the *ERV* denotes the electricity ratio at valley. N_p and N_v are the numbers of the peak period and valley period for 1 day.

It should be noted that: 1) All the 11 indices evaluate the fluctuation at a scale of day except for the extreme fluctuation frequency, which produces an annual result. Therefore, to use the proposed index system, a minimum of one-year power data should be collected. 2) For the daily indices, annual mean should be further calculated to form the annual evaluation index system. 3) Power systems or load exhibit different preferences to different features of fluctuation. For some indices, such as the 95% effective hours, the larger they are, the more favorable the PV plant is, and accordingly, they are called maximum indices. On the contrary, for other indicators such as the maximum positive fluctuation, the larger they are, the less favorable the PV plant is, so they are classified as minimum indicators.

2.3 RankBoost based ranking method for PV power fluctuation evaluation

The index system consists of 11 indices which describe different aspects of fluctuation characteristics and they are in different ranges with different units. To compare the fluctuation of different PV plants and generate an overall list, the RankBoost algorithm [14] is applied to combine individual indices. The method has two main advantages: 1) Different characteristics of fluctuation described by parameters/indices with different ranges and units can be combined by a learning process, which automatically creates fluctuation preference scores to different PV plants and generates an overall ranking list. 2) Missing data, invalid data or errors can be effectively and automatically handled [14].

1. Principle of the RankBoost ranking method

The detail procedure of the RankBoost ranking method has been illustrated in [13, 14], so only its main concept is described here.

The RankBoost ranking method is one type of the pairwise ranking method [13], which determines a comprehensive ranking by analyzing the relative position of every pair of the objections in the ranking sequence, rather than calculating the ranking scores of the individual objects. To find the best comprehensive ranking, the

probability of mis-ordering between the best ranking and individual rankings is defined as ranking loss. The best ranking will be successfully found when the ranking loss reaches the minimum value. It can be proved that minimizing the ranking loss is equivalent to minimizing the following *Z* function [14]:

$$Z = \sum_{i,j=1}^{i,j=m} D(S_k(i), S_k(j)) \times \exp(a(h(i) - h(j))) \quad (20)$$

where *r* is the iteration index. $S_k(i)$ denotes the k^{th} index results of the i^{th} PV plant and *m* denotes the number of the ranking objects. a_r is a factor that is a function of h_r , while h_r is a 0–1 valued function that can be calculated from the individual rankings. D_r is the distribution of every pair of individual indices of the PV plants. This reveals the rationality of the relative position of a pair of objects in the ranking list. By iterative solution as shown in Fig. 4, Z_r can be optimized and the comprehensive ranking can be obtained.

It is worth noting that the absolute *H* values are not important since what the RankBoost algorithm provides is a relative ranking order. The ranking results are non-linear, which means that the change of one input data can change the ranking order.

2. Data preprocessing

The raw data for fluctuation evaluation is the sampling power sequences covering a minimum period of 1 year. However, errors during sampling, empty data, invalid power data etc. can occur and will have a detrimental influence on the accuracy of the fluctuation evaluation.

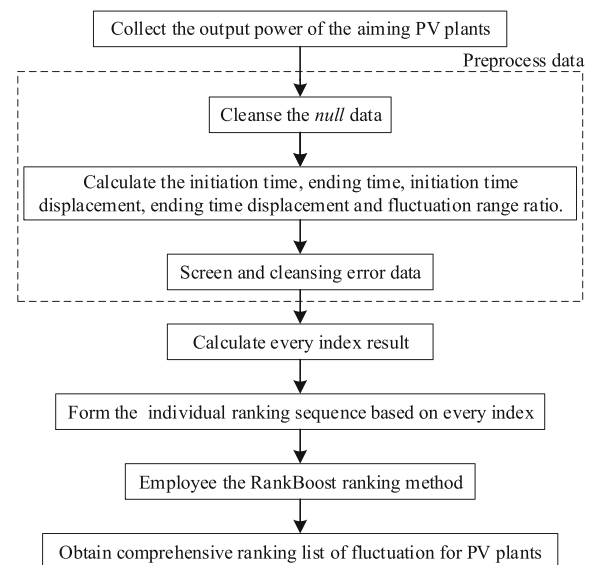


Fig. 4 The procedure of fluctuation evaluation

Table 1 The features of indices for error data

Characteristics of error data	The problems of error data
$\max\{P_i, i=1,2,\dots,4N\} \leq 0$ or $P_i(i=1,2,\dots,4N)$ is constant	If $t_{initiation} \leq 0$ or $t_{ending} \leq 0$, the PV power does not raise or fall, and is keeping either constant or zero during the whole day, which does not conform to the daily features of PV power described in Section 2.1.
$TD_{starting} < -1$ h or $TD_{terminal} > 1$ h	If the initiation time precedes sunrise or the ending time lags the sunset for one hour, it indicates the PV plant starting power generation one hour before the sunrise and continuing one hour after sunset, which is considered unnatural.
$TD_{starting} > 4$ h and $TD_{terminal} < -4$ h	If the initiation time lags sunrise and the ending time precedes the sunset for more than 4 h, the effective time of PV plant is so short and it becomes meaningless to analyze its fluctuation.
$f_r < 2$	If $f_r \geq 2$, the day curve of power generally assumes unimodal pattern. If $f_r < 2$, the power experiences continuous increase during the whole day or keeps constant, which is considered error data.

Therefore, to improve the accuracy and reliability of the ranking results, the existing data reprocessing process considering the features of PV power data is improved using the following procedures.

(1) Data cleansing

When the sampling power is missing, the sampling results may show *null* which is invalid for fluctuation evaluation. Therefore, days containing *Null* data should be eliminated and omitted from the fluctuation evaluation.

(2) Error screening

There are potential errors in the sampling process such as from the measuring equipment or data transmission process. Hence, raw data should be screened for errors which do not conform to the characteristics of PV power described in Section 2.1. By calculating the index results in the index system and combining the features of error data shown in Table 1, the error data can be located and omitted.

3. Procedure of the proposed ranking method for power fluctuation evaluation

The procedure of the proposed RankBoost method to determine the fluctuation priority of power for PV plants includes the following steps:

- (1) Collect the power data of all the target PV plants. In addition, the data should cover a minimum period of 1 year.
- (2) Cleanse the *null* data.
- (3) Calculate $t_{initiation}$, t_{ending} , $TD_{starting}$, $TD_{terminal}$, f_r and screen the error data according to the features shown in Table 1.
- (4) Input the cleansed and screened data to the index system to calculate every index result for all the objects for ranking.
- (5) Input all the index results of all the ranking objects and the individual ranking sequence can be formed.
- (6) Input all the index results of all the ranking objects and the individual ranking sequence, and use the RankBoost ranking method to obtain the comprehensive ranking list of fluctuation.

3 Results and discussion

In this section, the historical data collected from six PV plants are used to demonstrate the correctness and effectiveness of the proposed fluctuation evaluation method.

3.1 Simulation data

The historical power of six PV plants and total load power were collected from a regional power grid from 2017 to 2019, sampled every 15 min. Simulations are performed in MATLAB(R2016a).

Table 2 Data preprocessing results

PV plants	Proportions of being cleansed	Proportions of being screened	Total proportions of invalid data
Plant 1	1.78%	1.95%	3.70%
Plant 2	0.96%	1.11%	2.05%
Plant 3	0.55%	0.69%	1.23%
Plant 4	1.51%	0.70%	2.19%
Plant 5	1.23%	0.28%	1.51%
Plant 6	0.68%	0.28%	0.96%

Table 3 Individual Indices result of six PV plants

Category	Indices	Plant 1	Plant 2	Plant 3	Plant 4	Plant 5	Plant 6
Effective time fluctuation	Short Time Displacement rate	0.405	0.808	0.530	0.3410	0.552	0.3561
Power fluctuation	Fluctuation Range Ratio	4.395	4.694	4.030	4.325	4.682	4.285
	Maximum Positive Fluctuation	0.229	0.299	0.1234	0.231	0.106	0.217
	Minimum Negative Fluctuation	-0.237	-0.307	-0.127	-0.229	-0.109	-0.216
	Transition Degree	0.184	0.160	0.175	0.187	0.197	0.180
	Load Matching Degree	0.254	0.209	0.253	0.259	0.251	0.258
	Extreme Fluctuation Frequency	0.054	0.104	0.038	0.084	0.104	0.087
	Electricity fluctuation	95% Effective Hours(h)	5.013	5.549	3.325	4.808	2.096
95% Assuring Hours(h)		0.775	0.919	0.459	0.592	0.196	0.574
Electricity Ratio at Peak (MW-h)		0.092	0.041	0.078	0.070	0.015	0.145
Electricity Ratio at Valley (MW-h)		0.027	0.020	0.031	0.021	0.004	0.069

3.2 Data preprocessing

As shown in Table 2, the proportions of data cleaning and error screening differ in different PV plants, while Plant 6 exhibits the best reliability and correctness in data measurement and collection. Although the invalid data takes relatively small proportions for all the six plants, the null data and error data of plant 1 have the largest proportion among the six plants. This demonstrates that the data measurement and collection system of plant 1 needs improvement.

After data preprocessing, the indices for all the PV plants are derived, as listed in Table 3. By ranking the PV plants based on the individual indices, the distribution pattern of the PV plants in every individual ranking list is described in Fig. 5.

3.3 Fluctuation evaluation for PV plant power based on the proposed index system

PV plants perform differently in different features of fluctuation described by different indices. Nevertheless, there still exist some consistencies: 1) Maximum positive fluctuation and minimum negative positive fluctuation, 95% effective hours and 95% assuring hours can be considered as two pairs of indices, while all the PV plants

exhibit the same performance on either index of the same pair. 2) By analyzing the individual rankings, it is found that Plant 6 performs badly in most rankings. Hence, the last position in the comprehensive ranking list for fluctuation is assigned to Plant 6, whereas it is hard to arrange other positions directly.

Table 3 reveals that the plant indices are of different ranges and units. In addition, some plants differ from the others in some obvious indicators, while on other indicators plants only show small differences. For example, the load matching degrees of the six PV plants are very close to each other, while the 95% effective hours show significant differences.

3.4 Fluctuation ranking of PV plants based on the proposed ranking method

As shown in Fig. 6, with the RankBoost ranking method, the fluctuation preference priority of the six PV plants is ranked by the ranking score H . A PV plant with higher H value fluctuates less in power and has less negative impact on power prediction and load regulation.

The comprehensive ranking results demonstrate that Plant 3 is the most outstanding in terms of power fluctuation while Plant 5 behaves the worst. This is proved by the distribution pattern of individual rankings shown in Fig. 5.

Equal weight is a traditional approach to deal with multiple parameters for ranking. Figure 7 shows the ranking results using an equal weight for all the 11 indices. The minimum indicators can be transferred into maximum indicators by $x^* = 1/x$, whereas the negative indicators can be transferred into positive ones by $x^* = -x$. Equal weight is then employed to combine the multiple parameters and plants with higher weighted score value perform better in fluctuation.

Comparing the results of these two approaches, it can be seen that Plant 1 and 2 coincidentally have the same ranking. However, the other 4 plants rank differently in

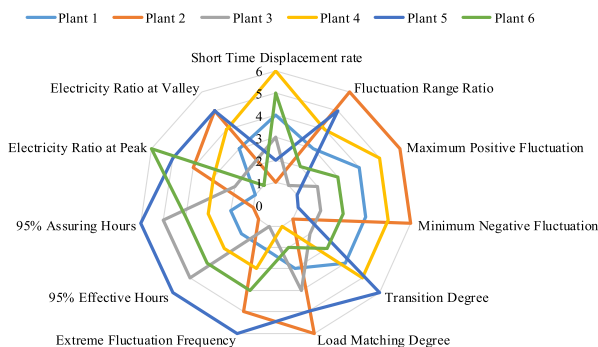


Fig. 5 The distribution pattern of individual rankings

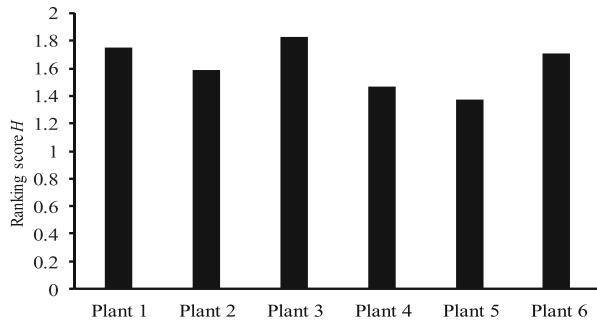


Fig. 6 The ranking results produced by the RankBoost method

the two ranking lists. In particular, Plant 5 is the worst plant in the list of the RankBoost method, whereas in the list of equal weight, it rises to third place. This is because after the conversion from minimum indicators to maximum indicators and from negative indicators to positive indicators, the values of maximum positive fluctuation and minimum negative fluctuation are almost ten times more than the other indicators. Hence, they take the dominant role in the equal weight results, amplifying the good performance of Plant 5 on these two aspects. This proves that the RankBoost ranking method can lead to a more balanced and comprehensive evaluation than the equal weight method.

4 Conclusions

This paper proposes an evaluation method to quantify the power fluctuation of PV plants, which consists of an index system and a ranking method based on the Rank Boost. By handling missing and invalid data effectively and combining the indices of fluctuation, the proposed method can provide a systematic and comprehensive comparison of PV plant power fluctuation. The derived ranking information of PV plants is significant for the power system operator and electricity market participants.

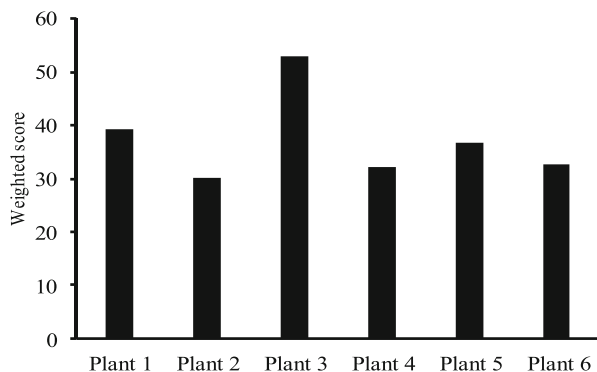


Fig. 7 The ranking results using the equal weight method

The historical power data of six PV plants are used to verify the effectiveness and correctness of the proposed method. The data preprocessing results show that the data collection process has good reliability, while the screening and cleansing method of power data can serve as an evaluation of reliability and correctness for data collection devices. Maximum positive fluctuation and minimum negative positive fluctuation, 95% effective hours and 95% assuring hours can be considered as two pairs of indices, on which all the PV plants exhibit the same performance. It is also shown that the fluctuation of PV power can be comprehensively quantified and precisely described by the RankBoost ranking method. Moreover, a more balanced and comprehensive evaluation can be obtained by the RankBoost ranking method than the equal weight method as the latter may lead to a predominant role for some indices while weakening the others.

Abbreviations

K_{fluc} : Fluctuation tendency index; STDR: Short time displacement rate; PV: Photovoltaic; t_{up} : Sunrise time; t_{down} : Sunset time; $t_{initiation}$: Initiation time of the PV plant; t_{ending} : Ending time of the PV plant; $TD_{starting}$: Initiation time displacement; $TD_{terminal}$: Ending time displacement; N : Number of sampling points in one day; P_i : Power of PV plant at the sampling point i ; P_{max} : Maximum P_i in one day; f_i : Fluctuation range ratio; ΔP_{max} : Maximum positive fluctuation; ΔP_{min} : Minimum negative fluctuation; D_T : Transition degree; D_{LM} : Load matching degree; ΔP_{avg+} : Annual average of daily maximum positive fluctuation; ΔP_{avg-} : Annual average of daily minimum negative fluctuation; $\Delta P_{i,max}$: Maximum positive fluctuation at day i ; $\Delta P_{i,min}$: Minimum fluctuation at day i ; c : A constant assigned to be 80% in this paper; d_+ : Number of days in which the power of PV plant meets $\Delta P_{i,max} > (\Delta P_{max} - \Delta P_{avg+}) \times c + \Delta P_{avg+}$; d_- : Number of days in which the power of PV plant meets $\Delta P_{i,min} < (\Delta P_{min} - \Delta P_{avg-}) \times c + \Delta P_{avg-}$; f_e : Frequency of the extreme fluctuation during one year; E_d : Electricity generated by the plant at day d ; $P(t)$: Power function of the PV plant during day d ; p : Accrued probability of daily electricity; $E(p)$: Function of annual sustained electricity curve; 95%EH: 95% Effective hours; 95%AH95% Assuring hours; P_N : Rated power of the PV plant; L_{avg} : Average power of load in one day; $L(t)$: Total power of the load in the region of PV plant at time t ; ERP : Electricity ratio at peak; ERV : Electricity ratio at valley; N_p : Number of the peak period; N_v : Number of the valley period; Z : Ranking loss function; $S_k(t)$: The k^{th} index results of the i^{th} PV plant; m : Number of the ranking objects; h_i : A 0–1 valued function; a_i : A factor that is a function of h_i ; D_i : Distribution of every pair of individual indices of the PV plants; H : Ranking score of the ranking objections; x^* : Index result after convention; x : Maximum or minimum index result before convention.

Acknowledgements

The authors would like to thank National Key R&D Program of China (Technology and application of wind power / photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) for supporting the project.

Authors' contributions

Xia WY proposed a quantified method to evaluate and compare power fluctuation of PV plants comprehensively and scientifically. The authors read and approved the final manuscript.

Authors' information

Weiyi Xia (1996-), is currently pursuing the M.S. degree with the school of Electrical Engineering, Chongqing University, Chongqing, China. Her research interests include planning and operation of new energy power generation. Zhouyang Ren (1986-), corresponding author, received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2014, where he is currently an Associate Professor with the School of

Electrical Engineering. His research interests include reliability evaluation and planning of power systems, and power system data analysis.

Hui Li (1995-), is currently working toward the Ph. D. degree in the School of Electrical Engineering, Chongqing University, Chongqing, China. His research interests include planning and operation analyses of tidal current generation integrated power system, and the planning of renewable energy integrated distribution networks.

Bo Hu (1983-), received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2010. Currently, he is a professor in the School of Electrical Engineering at Chongqing University, China. His research interests include power system reliability, parallel computing techniques in power systems.

Funding

The research work is supported by National Key R&D Program of China (Technology and application of wind power / photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDK00KJJS1800266).

Availability of data and materials

The power data of PV plants used for case study is from the supporting program, and only allowed to be shared within the program according to the secrecy rules.

Declarations

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 28 January 2020 Accepted: 7 July 2021

Published online: 25 August 2021

References

- Sobri, S., Koohi-Kamali, S., & Rahim, N. A. (2018). Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, *156*, 459–497. <https://doi.org/10.1016/j.enconman.2017.11.019>.
- Chen, J., Pan, G., Ouyang, J., & Zhang, L. (2020). A power quality evaluation method for distribution system with renewable energy. *International Journal of Electronics*, *107*(5), 792–807. <https://doi.org/10.1080/00207217.2019.1680745>.
- Yang, X., Ren, J., Li, X., & Zhang, H. (2018). Typical daily power curve Mining for Energy Storage Systems under smoothing power fluctuation scenarios. *Mathematical Problems in Engineering*, *2018*, 1503092–1503012. <https://doi.org/10.1155/2018/1503092>.
- Tripathy, S. K., Mitra, I., Heinemann, D., Giridhar, G., & Gomathinayagam, S. (2017). Impact assessment of short-term variability of solar radiation in Rajasthan using SRRA data. *Renewable & Sustainable Energy Reviews*, *78*, 798–806.
- Brinkel, N. B. G., Gerritsma, M. K., AlSkaif, T. A., Lampropoulos, I., van Voorden, A. M., Fidder, H. A., & van Sark, W. G. J. H. M. (2020). Impact of rapid PV fluctuations on power quality in the low-voltage grid and mitigation strategies using electric vehicles. *International Journal of Electrical Power & Energy Systems*, *118*, N.PAG-N.PAG. <https://doi.org/10.1016/j.ijepes.2019.105741>.
- Han, S., Zhang, L., Liu, Y., Zhang, H., Yan, J., Li, L., ... Wang, X. (2019). Quantitative evaluation method for the complementarity of wind-solar-hydro power and optimization of wind-solar ratio [J]. *Applied Energy*, *236*, 973–984. <https://doi.org/10.1016/j.apenergy.2018.12.059>.
- Wang, F., Zhen, Z., Mi, Z., Sun, H., Su, S., & Yang, G. (2015). Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. *Energy and Buildings*, *86*, 427–438. <https://doi.org/10.1016/j.enbuild.2014.10.002>.
- Anvari, M., Werther, B., Lohmann, G., Wächter, M., Peinke, J., & Beck, H. p. (2017). Suppressing power output fluctuations of photovoltaic power plants. *Solar Energy*, *157*, 735–743. <https://doi.org/10.1016/j.solener.2017.08.038>.
- Kawasaki, N., Oozeki, T., Otani, K., & Kurokawa, K. (2006). An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis. *Solar Energy Materials & Solar Cells*, *90*(18/19), 3356–3363. <https://doi.org/10.1016/j.solmat.2006.02.034>.
- Cai, Y. H., & Schäfer, A. I. (2020). Renewable energy powered membrane technology: Impact of solar irradiance fluctuation on direct osmotic backwash. *Journal of Membrane Science*, *598*, 117666. <https://doi.org/10.1016/j.memsci.2019.117666>.
- Salarvand, A., Mirzaeian, B., & Moallem, M. (2010). Obtaining a quantitative index for power quality evaluation in competitive electricity market. *IET Generation, Transmission & Distribution*, *4*(7), 810–823. <https://doi.org/10.1049/iet-gtd.2009.0479>.
- Liu, G., Li, M., Zhou, B., Chen, Y., & Liao, S. (2018). General indicator for techno-economic assessment of renewable energy resources. *Energy Conversion & Management*, *156*, 416–426. <https://doi.org/10.1016/j.enconman.2017.11.054>.
- Freund, Y., Iyer, R., Schapire, R. E., Singer, Y., & Dietterich, T. G. (2004). An efficient boosting algorithm for combining preferences. *Journal of Machine Learning Research*, *4*(6), 933–969. <https://doi.org/10.1162/1532443041827916>.
- Zhong, J., Li, W., Wang, C., & Yu, J. (2018). A rank boost-based data-driven method to determine maintenance priority of circuit breakers. *IEEE Transactions on Power Delivery*, *33*(3), 1044–1053. <https://doi.org/10.1109/TPWRD.2017.2748146>.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)