Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 142 (2017) 77–82

Energy Procedia

www.elsevier.com/locate/procedia

9th International Conference on Applied Energy, ICAE2017, 21-24 August 2017, Cardiff, UK

Daily clearness index profiles and weather conditions studies for photovoltaic systems

Chun Sing Lai[a,b], Xuecong Li[a,]*, Loi Lei Lai[a], Malcolm D. McCulloch[b]

[a]*Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou, China, 510006*
[b]*Energy and Power Group, Department of Engineering Science, University of Oxford, 17 Parks Road, Oxford, OX1 3PJ, United Kindom*

Abstract

The increasing number of distributed photovoltaic (PV) systems connected to the power grid has made system planning and performance evaluation a challenging task. This is mainly due to the computational complexity, such as load flow analysis with large irradiance datasets collected from various locations of the installed PV farms. Solar irradiance data are known to possess the characteristic of high uncertainty, due to the random nature of cloud cover and atmospheric conditions. This paper presents the studies on the relationships of clustered clearness index profiles and the weather conditions obtained from the weather forecasting stations. Four years of solar irradiance and weather conditions data from two locations (Johannesburg and Kenya) were obtained and are used for the analysis. The preliminary study shows that the weather condition is related to the daily clearness index profiles. This work will form the basis for estimating the daily clearness index profile with weather conditions.

Keywords: photovoltaic system; clustering; clearness index

1. Introduction

Photovoltaic (PV) systems are being connected to the grid via the distribution systems at an exponential rate [1]. This is the result from reduced cost and increased in economy of scales of PV systems [2]. The transition in the way power is generated is welcomed by the government and environmentalist. Green energy, i.e. solar PV can reduce the

* Corresponding author.
E-mail address: leexuecong@126.com (X. Li), chun.lai@eng.ox.ac.uk (C.S. Lai), l.l.lai@ieee.org (L.L. Lai), malcolm.mcculloch@eng.ox.ac.uk (M.D. McCulloch).

10.1016/j.egypro.2017.12.013

carbon emission and environmental pollutions. However, this also poses a serious problem for system operators. Power and voltage fluctuations on the grid are the main concerns, as a direct result from the non dispatchable PV generation. The problem is worsened with the fact that PV systems in general are of small scale and situated in various locations. Feed-in-tariffs from various countries promote the installation of small scale PV farms, typically of kilowatts. The solar irradiance can be of significant difference at different locations due to the climate and cloud cover. Large scale irradiance datasets are required for detail analysis of the PV farms to the grid in order to prevent brownouts and blackouts.

Currently, irradiance fluctuations are mostly studied with statistical techniques [3-9]. The advantage of these methods is that few parameters can be used to characterize the massive quantity of data. Clearness index (CI) can be useful in studying the fluctuations for solar energy applications [6, 10-12]. The diminish impact of the atmosphere on solar irradiance with respect to the amount of extraterrestrial solar irradiance that reaches the surface of the earth can be presented by CI.

The recent work in [13] presents a study in cluster analysis of daily CI profiles. This work aims to study how well the proposed clustering method correlates with the weather forecasting information, i.e. weather conditions. This will confirm the dependency of the CI and the weather conditions and in addition, the accuracy of the clustering results. The results will also suggest that the clustering procedure can be applied to other locations for PV system evaluations, such as sizing and distribution system load flow analysis. This will be useful in estimating the CI profiles with limited weather conditions data. Also, it provides a validation method for the cluster analysis results and to confirm the generality of the clustering approach.

One of the major issues that needs to be addressed is the question of "How is a 'Clear day' defined?". As an example, can a day with strong cloud cover in the noon for two hours and with no cloud for the rest of the day be classified as a 'Clear day'? This question will be explored and addressed in the paper. Section 2 presents the formal definition of CI. The cluster analysis results of Kenya with Fuzzy C-Means Dynamic time warping (FCM DTW) will be given in Section 3. Section 4 will present the study of weather conditions and clearness index. An algorithm is proposed to determine the percentage of clear days from a set of daily weather conditions. Conclusion and future work are given in Section 5.

2. Clearness Index

The solar irradiance received on the ground will be equal to the value of the solar constant subtract the amount of atmospheric absorption under the ideal atmospheric condition. In general, the global solar irradiance consists of two main components, these are known as diffuse sky irradiance and the direct beam component. Typically, the real-life solar irradiance collected for solar application studies is from a pyranometer device. It measures the solar irradiance on a flat surface and measures the solar radiation flux density in W/m^2. The CI is calculated from the data obtained from the pyranometer data and clear sky model. The CI has a value between 0 and 1. The value 0 signifies that a total cloud cover occurs and no irradiance is to be received on the ground. Conversely, a value of 1 signifies that the maximum theoretical amount will be received on the ground. Exceptional cases need to be made when using CI, such as the definition of CI before sunrise and after sunset where the irradiance will be zero. As these conditions are not useful in this study, they will be neglected during the analysis. The equation for CI calculation is given in Equation (1) below:

$$CI(t) = \frac{I_{pyranometer}(t)}{I_{model}(t)} \tag{1}$$

I_{model} is the clear-sky solar irradiance from the solar model and $I_{pyranomter}$ is the real-life solar irradiance at time t. The clearness index obtained in this work are calculated from the clear-sky solar irradiance model in [13]. Various variables are incorporated in the solar model to compute the clear-sky solar irradiance. These include correction factor to mean solar distance, optical air mass corrected for station height, solar altitude angle and Rayleigh optical depth.

3. Cluster analysis results

The FCM DTW clustering technique is studied with the four years of solar irradiance data (2011-2014) obtained from Solargis company. The corresponding weather conditions are obtained for Nairobi Jomo, Kenya from Weather Underground [14]. The location of the irradiance data is at Gitaru dam, Kenya with coordinates (0.789°S,37.73°E) and elevation at 969 meters. The sampling rate is at 1 sample/15min. The weather conditions for the four years of Johannesburg were also collected from the database in the same website. The clustering results for Johannesburg case are given in [13]. Daily CI profiles are constructed for the four seasons in Kenya and the results are presented in Fig. 1. The mean, plus one and minus one standard deviation for the CI profiles are given in black, blue and red lines respectively. The optimal numbers of clusters obtained from fuzzy decision with cluster reduction are 5, 4, 6 and 6 for Spring, Summer, Autumn and Winter respectively.

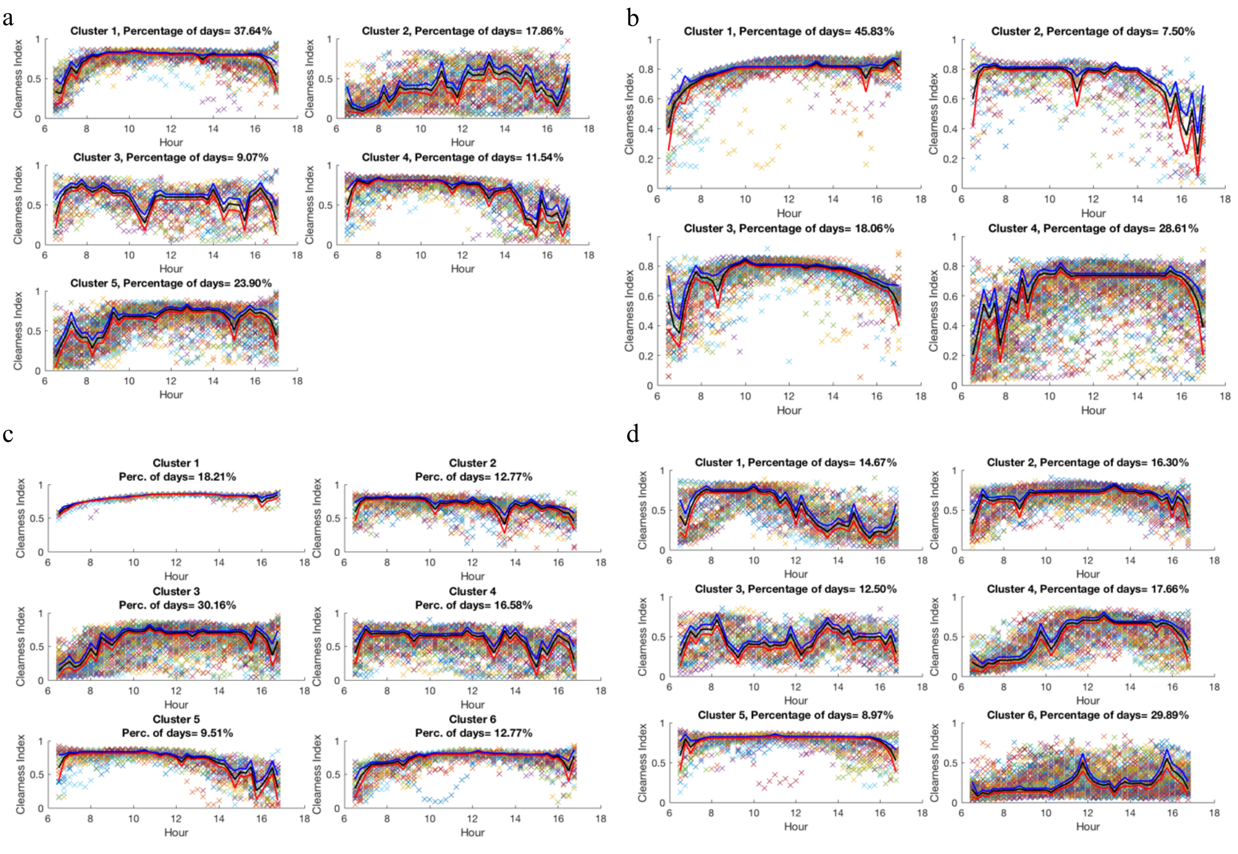

Fig. 1. (a) Daily CI profiles for Spring case in Kenya; (b) Summer case in Kenya; (c) Autumn case in Kenya; (d) Winter case in Kenya.

As shown in Fig. 1a, the 'Clear days' for Spring in Kenya belongs to cluster 1. The cluster has high CIs values with little fluctuations. The CIs with lowest values are situated in cluster 2. The 'Clear days' for Summer in Kenya belongs to cluster 1. Although the other centroids are generally of high mean value, there are some form of perturbations, such as the centroid in cluster 3 reaches a value of 0.3 at 7 am. The 'Clear days' for Autumn in Kenya belongs to clusters 1 and 6. Although clusters 2 and 4 show some high CI values, the CI suffers from cloud covering. The 'Clear days' for Winter in Kenya belongs to cluster 5.

The dataset with weather conditions to be analyzed is shown in Table 1. Columns 1, 2, 3 and 4 represent the time in year, month, day and hour in MATLAB format respectively. The solar irradiance is given in column 5. The weather condition and the CI are given in columns 6 and 7 respectively.

Table 1. Data structure for time, irradiance, weather conditions and CI.

Year	Month	Day	Hour/Minute (MATLAB)	Irradiance (Wm^{-2})	Weather condition	Clearness Index
2011	9	1	0.004166667	0	'Scattered Clouds'	NaN
2011	9	1	0.014583333	0	'Scattered Clouds'	NaN
2011	9	1	0.025	0	'Scattered Clouds'	NaN
2011	9	1	0.035416667	0	'Scattered Clouds'	NaN
2011	9	1	0.045833333	0	'Clear'	NaN
2011	9	1	0.05625	0	'Clear'	NaN
2011	9	1	0.066666667	0	'Clear'	NaN
2011	9	1	0.077083333	0	'Clear'	NaN
2011	9	1	0.0875	0	'Scattered Clouds'	NaN
2011	9	1	0.097916667	0	'Scattered Clouds'	NaN
2011	9	1	0.108333333	0	'Scattered Clouds'	NaN
2011	9	1	0.11875	0	'Scattered Clouds'	NaN
2011	9	1	0.129166667	0	'Partly Cloudy'	NaN
2011	9	1	0.139583333	0	'Partly Cloudy'	0
2011	9	1	0.15	4	'Partly Cloudy'	0.077318929
2011	9	1	0.160416667	20	'Partly Cloudy'	0.166207348
2011	9	1	0.170833333	43	'Mostly Cloudy'	0.219380228
2011	9	1	0.18125	66	'Mostly Cloudy'	0.240601593
2011	9	1	0.191666667	211	'Mostly Cloudy'	0.59744908
2011	9	1	0.202083333	223	'Mostly Cloudy'	0.51692384
2011	9	1	0.2125	313	'Mostly Cloudy'	0.61587608
2011	9	1	0.222916667	390	'Mostly Cloudy'	0.668906293
2011	9	1	0.233333333	472	'Mostly Cloudy'	0.720190264
2011	9	1	0.24375	541	'Mostly Cloudy'	0.746385476
2011	9	1	0.254166667	599	'Scattered Clouds'	0.757265334
2011	9	1	0.264583333	655	'Scattered Clouds'	0.767355712
2011	9	1	0.275	701	'Scattered Clouds'	0.768428757
2011	9	1	0.285416667	731	'Scattered Clouds'	0.7561521
2011	9	1	0.295833333	765	'Mostly Cloudy'	0.752372712
2011	9	1	0.30625	825	'Mostly Cloudy'	0.776718747
2011	9	1	0.316666667	884	'Mostly Cloudy'	0.801696435
2011	9	1	0.327083333	908	'Mostly Cloudy'	0.797818216

4. Weather conditions and clearness index studies

4.1. Background

As shown in Table 1, the weather condition for each time interval can be different. Hence, there will be a degree of ambiguity for defining 'Clear day'. In practice and real-life situations, there are very few instances with '100%' clear day. According to the weather glossary from timeanddate.com [15], the 'Clear' weather condition is defined as a sky condition of less than 1/10 cloud covered. The 'Partly Cloudy' condition is defined as the sky condition when between 7/10ths and 3/10ths of the sky is covered, and is used more frequently at night. The 'Scattered Clouds' condition is defined as a cloud layer that covers between 3/8ths and ½ of the sky.

According to Weather Underground [14], there are 42 other terms such as, 'Rain', 'Thunderstorm', 'Drizzle', 'Snow', 'Haze', etc. that may be used to represent the weather condition for a particular time instance. In this context, the 'Clear day' can be defined with high CI value and in general associate with the weather conditions of 'Clear', 'Partly Cloudy' and 'Scattered Clouds'. Table 2 presents the algorithm to calculate the percentage of 'Clear days' for the season, with the given dataset in Table 1.

Table 2. Algorithm for calculating the total percentage of ‘Clear days’ with weather data.

Input: $d = \{d_1, d_2, \dots, d_k\}$: object with the set of daily time and weather for k time intervals
threshold: Percentage of cloud cover to define if a day is ‘Clear day’
start: element number in the array for sunrise
end: element number in the array for sunset
Output: d'': the percentage of ‘Clear’ condition for the season

```
for i1 = 1:length(d)
  for i2 = start:end
    if (d{1, i1}(i2, 6) == 'Clear' or 'Scattered Clouds' or 'Partly Cloudy') is true
      S(i1, i2) = 1;
    elseif d{1, i1}(i2, 6) == empty is true
      S(i1,i2) = NaN;
    else
      S(i1,i2) = 0;
    end
  end
end
perc_clear = sum(S, 2)/length(S, 2);
for i3 = 1:length(d)
  if (perc_clear(i3) > threshold) is true
    d'(i3) = 1;
  else
    d'(i3) = 0;
  end
end
d'' = sum(d') / length(d');
```

5. Results

As presented in Table 2, the algorithm aims to calculate the percentage of ‘Clear days’ by assigning binary values to the time instances and there is no need to use CI values. It should be worth noting that there are missing weather data in the dataset. These are then converted to ‘NaN’ and are neglected during the summation process in the algorithm.

The threshold can be seen as a sensitivity value and a key parameter in the algorithm. If above the threshold, it is expected that the CI produced for the day will be of high value with little perturbation. It should be recognized that due to different locations, seasons and climates, it may be possible for the threshold to be different. The threshold aims to give the percentage of ‘Clear days’ from weather analysis that represents the closest percentage of the total number of ‘Clear day’ clusters. The comparison of results for cluster analysis and weather analysis are presented in Tables 3 and 4 for Kenya and Johannesburg studies.

Table 3. Percentage of ‘Clear days’ with cluster analysis and weather analysis for Kenya case study.

		Spring	Summer	Autumn	Winter
Cluster analysis	Cluster number	1	1	1 and 6	5
	Total percentage (%)	37.64	45.83	30.98	8.97
Weather analysis		36.47	47.23	31.01	7.72
Threshold (%)		45	50	35	10
Absolute difference (%)		1.17	1.4	0.04	1.25

Table 4. Percentage of ‘Clear days’ with cluster analysis and weather analysis for Johannesburg case study.

		Spring	Summer	Autumn	Winter
Cluster analysis	Cluster number	4	5	3 and 4	3
	Total percentage (%)	39.56	19.10	50.82	77.99
Weather analysis		39.78	17.18	51.38	79.67
Threshold (%)		95	95	85	80
Absolute difference (%)		0.22	1.92	0.56	1.68

From Tables 3 and 4, the total percentage of 'Clear days' shows that there are more 'Clear days' in Johannesburg than Kenya, such as for Spring, Autumn and Winter. However, such comparison is invalid as the threshold value is different for both studies. The amplitude of centroids is generally higher for Kenya than Johannesburg, for all clusters and seasons. It can be concluded that the threshold value can influence the percentage of 'clear days' for a particular season. The threshold for Kenya is lower than Johannesburg, this indicates that it requires less appearances of 'Clear'/ 'Scattered Clouds'/ 'Partly Cloudy' for the day to be classified as 'Clear day'. The threshold is the percentage that represents the 'Clear day' duration. As an example, with a threshold at 25% and an average daily sunshine hours in Kenya at 12 hours, this effectively means that the day will be classified as a 'Clear day' if the day contains three hours of 'Clear'/ 'Scattered Clouds'/ 'Partly Cloudy'. The percentage difference is larger for Summer than other seasons in Johannesburg. This may be due to the increased fluctuations of the CIs for the season, as shown in Fig. 3 in [13] and has proved to be more difficult to provide high quality clusters.

6. Conclusions

This work evaluates the clustering results with the forecast weather conditions. The results show that the weather condition can affect the clearness index, and subsequently the percentage of 'Clear days'. It is possible to estimate the shapes and magnitude of daily clearness index profiles with given limited weather condition data. For future work, a systematic method to define the threshold, possibly through optimization approach will be useful. The meaning and indication of threshold need to be further studied with additional irradiance data from other locations. The threshold will be useful in giving an indication of the number of appearances for 'Clear'/ 'Scattered Clouds'/ 'Partly Cloudy' of the day to produce high value CIs and to be considered as a 'Clear day'.

Acknowledgements

This research is supported in part by the grant from Department of Finance and Education of Guangdong Province 2016[202]: Key Discipline Construction Programme and Education Department of Guangdong Province: New and integrated energy system theory and technology research group [project number 2016KCXTD022]. The authors would like to thank Mr Junyu Li, undergraduate student at Guangdong University of Technology in data cleaning.

References

[1] F. Katiraei and J. R. Aguero, "Solar PV integration challenges," IEEE Power and Energy Magazine, vol. 9, pp. 62-71, 2011.
[2] C. S. Lai and M. D. McCulloch, "Levelized cost of electricity for solar photovoltaic and electrical energy storage," Applied Energy, vol. 190, pp. 191-203, 3/15/ 2017.
[3] A. Maafi and S. Harrouni, "Preliminary results of the fractal classification of daily solar irradiances," Solar Energy, vol. 75, pp. 53-61, 2003.
[4] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of pv power output," IEEE Transactions on Sustainable Energy, vol. 5, pp. 917-926, 2014.
[5] T. Soubdhan, R. Emilion, and R. Calif, "Classification of daily solar radiation distributions using a mixture of Dirichlet distributions," Solar Energy, vol. 83, pp. 1056-1063, 2009.
[6] S. Kheradmanda, O. Nematollahi, and A. R. Ayoobia, "Clearness index predicting using an integrated artificial neural network (ANN) approach," Renewable and Sustainable Energy Reviews, vol. 58, pp. 1357-1365, 2016.
[7] A. Mellit, S. Kalogirou, S. Shaari, H. Salhi, and A. H. Arab, "Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: application for sizing a stand-alone PV system," Renewable Energy, vol. 33, pp. 1570-1590, 2008.
[8] T. Ayodele and A. Ogunjuyigbe, "Prediction of monthly average global solar radiation based on statistical distribution of clearness index," Energy, vol. 90, pp. 1733-1742, 2015.
[9] Z. Ren, W. Yan, X. Zhao, W. Li, and J. Yu, "Chronological probability model of photovoltaic generation," IEEE Transactions on Power Systems, vol. 29, pp. 1077-1088, 2014.
[10] A. Woyte, R. Belmans, and J. Nijs, "Fluctuations in instantaneous clearness index: Analysis and statistics," Solar Energy, vol. 81, pp. 195-206, 2007.
[11] A. Woyte, V. Van Thong, R. Belmans, and J. Nijs, "Voltage fluctuations on distribution level introduced by photovoltaic systems," IEEE Transactions on Energy Conversion, vol. 21, pp. 202-209, 2006.
[12] J. Liu, W. Fang, X. Zhang, and C. Yang, "An improved photovoltaic power forecasting model with the assistance of aerosol index data," IEEE Transactions on Sustainable Energy, vol. 6, pp. 434-442, 2015.
[13] C. S. Lai, Y. Jia, M. D. McCulloch and Z. Xu, "Daily clearness index profiles cluster analysis for photovoltaic system," IEEE Transactions on Industrial Informatics, 2017. DOI: 10.1109/TII.2017.2683519.
[14] "Historical data," Weatherunderground.com, [Online]. Available: https://www.wunderground.com/history/. [Accessed: 06/06/17].
[15] "Weather glossary - terms & definitions," timeanddate.com. [Online]. Available: https://www.timeanddate.com/weather/glossary.html. [Accessed: 06/06/17].