CHARACTERIZATION OF SOLAR SPECTRAL IRRADIANCE USING THE INDEX OF AVERAGE PHOTON ENERGY FOR ITS USE IN THE PERFORMANCE EVALUATION OF PV MODULES

Rafael Moreno-Sáez[1], Mariano Sidrach-de-Cardona[1], Llanos Mora-López[2]
[1]Dept. Applied Physics II. [2]Dept. Computer Sciences and Languages. University of Malaga.
Bulevar Louis Pasteur 35 (29071), Málaga, Spain. Tel: +34 952132772.
e-mail: rmorenosaez@ctima.uma.es; msidrach@ctima.uma.es; llanos@lcc.uma.es

ABSTRACT: A model for characterizing the solar spectral radiation distribution only using APE value as input parameter is proposed. Although the uniqueness of the relationship between these parameters has been previously established there is no expression that allows obtaining the solar spectral curves for each APE value. In this work the APE values for which this uniqueness works have been checked. For this, the cumulative probability distribution function for each spectrum has been estimated. Using a data mining technique, named k-means, all these curves have been clustered in five different groups. In each group a solar spectral cumulative distribution function curve known as centroid has been estimated; this curve is similar to the rest of the curves gathered in the group. The spectra distribution per clusters is quite homogeneous. The centroid of each cluster characterizes all curves in the cluster. For doing the analysis, more than two-hundred-fifty-thousands spectra have been used.

Keywords: Average Photon Energy, k-means, Performance Ratio

1 INTRODUCTION

The spectral influence on photovoltaic modules performance and on the energy production is an issue that has been widely studied. Spectrum seasonal effect has been evaluated by Hirata and Tani [1], Fabero and Chenlo [2] and Gottschalg et al. [3]. In these works they conclude that spectral irradiance curves, its wells and peaks, influences on the module performance. The dependence is different for every kind of technology. In this context a-Si and CdTe cells are the most affected technologies by the spectrum and m-Si and CIGS have a lower influence from the spectral solar irradiation distribution. The reason of this relationship is caused by the spectral response of the module. In the first case module spectral response is only in the ultraviolet and visual range of the solar spectrum, meanwhile the m-Si and CIGS technologies have also spectral response on the near-infrared part of the spectrum.

The solar spectral irradiance influences on the performance as the shape and contribution of irradiance per wavelength changes because there is a high dependence on the hour of the day, the season and the kind of atmosphere. The time and season are influenced by the air mass and the atmosphere influence is due to the Rayleigh scattering and the water vapor that holds on it. According to these issues the spectrum sometimes has a shift to the lowest wavelength, to the blue part of the spectrum and sometimes the shift is to the red part, the highest wavelength.

The aim of this work is to characterize all the solar spectral irradiations recovered for more than one year at the location of Malaga and to gather these spectra in a few groups in order to be able to include this information as another parameter in addition to the typical meteorological parameters. The main scope of the work is to be able to have more accurate predictions of the energy produced by the modules. This can be an interesting tool for large photovoltaic plants which will have a better estimation of the energy production and a more precise integration on the power grid. For this scope we use a parameter that characterizes with a single number the whole solar spectral irradiance distribution, the -APE- Average Photon Energy value. Once the APE is calculated for every spectrum we put them into groups using *k-means* technique that is a clustering method for big amount of data. Finally, we have proved that all APE recovered everyday for more than one year for the location of Malaga can be reduced to a few groups of APE.

2 DATA ACQUISITION SYSTEMS

The device used for spectra data recording is a grating spectroradiometer prepared for continuous outdoor measurement which components have little aging problems, created specifically to capture spectral irradiance distributions. It has a filter glass dome that has been cleaned every day to avoid dust deposition which could affect the measurement. Its silicon sensor sets the wavelength range from 350 to 1050 nm what includes VIS and NIR with a spectral resolution below 8 nm at wavelength interval of 0.75 nm. It makes the measurement time short to between 10 msec and 5 sec automatically controlled so, on days with moving clouds, we get a fine spectrum with no distortion. Manufacturer provides its own software to measure solar spectra.

The spectroradiometer is placed on a fixed 21^{o} slope frame facing the south in the Solar Photovoltaic Energy laboratory flat roof at the University of Malaga (Spain), latitude 36.7^{o} N, longitude 4.5^{o} W, height 50 m. In the same frame, close to the spectroradiometer there are other devices used in photovoltaic such as pyranometers and calibrate cells and a weather station to recover meteorological data. There are also modules of different technologies.

Since November 2010 until May 2012, for more than one year, spectra have been continuously recorded at a rate of one spectrum per minute. In this context the characterization of a complete year can be done so that seasonal effects are included. Based on these premises, around 400.000 spectra were captured. The data were filtered to avoid reflection produced by angle of incidence and other effects. Spectra taken with an elevation angle under 15^{o} are removed as explained in

Nann and Riordan [4] remaining 70% of the spectra captured, around 280.000 spectra.

The data of the electrical parameters of different photovoltaic modules technologies and some meteorological parameters are also recorded for the same period. Specifically, the following parameters have been recorded: module output power, module temperature, ground irradiance in the module surface and temperature. The measurement system is installed nearby the spectra measurement system. These data have been recorded during daytime every five minutes recovering around 60.000 measurements in one year.

A program has been developed for joining all recorded measurement: spectra, modules and meteorological parameters.

3 SPECTRAL CHARACTERIZATION

3.1 Average Photon Energy

To characterize the spectrum we are going to use the APE value, a number that gives information of the spectrum. APE stands for Average Photon Energy and is calculated from the solar spectral distribution using the following expression, Eq. 1:

$$APE = \frac{\int_a^b E(\lambda)\cdot d\lambda}{q\cdot\int_a^b \phi(\lambda)\cdot d\lambda} \quad [\mathrm{eV}] \qquad (1)$$

where:
E.- is the irradiance distribution per wavelength,
ϕ.- is the photon flux density per wavelength,
q.- is the electron charge.

We want to prove that spectra with the same APE values are similar. To perform this comparison we gather spectra indexed by APE with an interval of 0.01 eV. The result of this can be seen in Fig. 1.

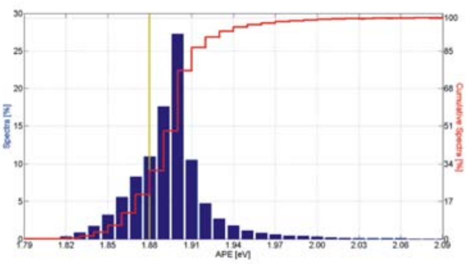

Figure 1: Spectra distributed in terms of APE

Over 60% of the spectra have an APE value over 1.88 eV –the APE value for the standard spectrum AM 1.5- in the spectral range from 350 to 1050 nm. This means that spectrum in Malaga can be considered as shifted to the 'blue'. This effect is caused by the water vapor that holds on the atmosphere, a seaside town.

3.2 Spectral distribution

To carry out the comparison between two spectra we need to normalize them. We propose to use the cumulative probability distribution function as a normalization procedure. The expression is as follows, Eq.2:

$$c.p.d.f.(\lambda) = P(\Lambda \le \lambda) = \int_{\lambda_o}^{\lambda_f} f(\lambda)\cdot d\lambda \qquad (2)$$

The result can be seen in Fig. 2 where one spectrum and its c.p.d.f. are drawn.

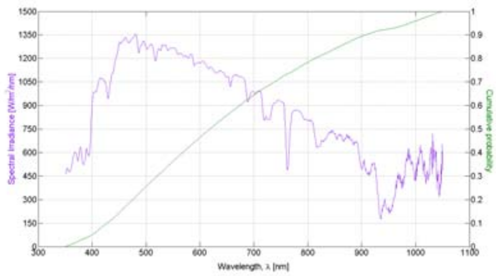

Figure 2: Spectrum and c.p.d.f.

Once calculated the c.p.d.f. for every spectrum we compare them by calculating the distance between curves with the same APE value as can be seen in Fig. 3.

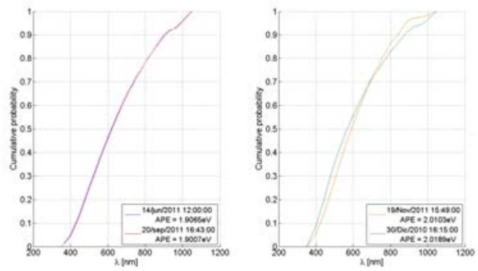

Figure 3: c.p.d.f. superposition with same APE values.

Curves on the right figure are considered as similar, and curves on the left are considered as different.

3.3 Spectra data clustering

In order to reduce the number of different c.p.d.f. measured, we use a clustering method to gather all the spectra in a few groups: the *k-means* clustering technique [5]. It is a partitional algorithm that distributes every sample in one of the *k* clusters. It determines a centroid per cluster by calculating the samples mean in every centroid. The number of clusters and the initial placement of the centroids are a key issue. Although there are initialization methods to calculate the number of clusters [6, 7] we are going to chose the number of cluster taking into account the results obtained by the PR.

Once we have decided the number of groups we assign every spectrum to a cluster by calculating the minimum Euclidean distance between the spectrum c.p.d.f. and the cluster centroid.

What we are going to do is to gather all the spectra in a few groups by comparing their c.p.d.f. and see if sample distributes in an appropriate manner.

4 PERFORMANCE RATIO ANALYSIS

The performance ratio –PR- has been estimated to analyze how spectrum influences on the module performance. The PR has been indexed to APE and T_{MOD}

according to the procedure developed by Minemoto et al. [8].

The result for one of the modules installed in the laboratory facilities is depicted in Fig. 4. It is a m-Si module with peak power provided by the manufacturer of 94 Wp. The spectral range for this module is the same than the spectroradiometer wavelength sensor, up to 1050 nm.

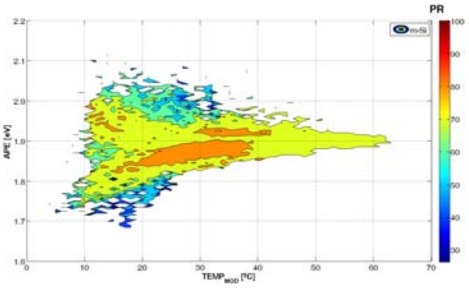

Figure 4: Contour graph of the module PR

4 RESULTS

The results obtained when samples are clustered using the *k-means* method are shown in Table 1 where the percentage of spectra assigned to every cluster for different values of *k* (number of clusters) used is shown.

Table 1: Spectral distribution per cluster

	Cluster (%)					
No. clusters	**1**	**2**	**3**	**4**	**5**	**6**
4	11.6	26.4	52.7	9.3		
5	9.8	23.5	48,4	13.9	4.4	
6	5.9	15.4	23,2	40.7	11.1	3.7

It seems that samples distribution is quite homogeneous and elements are well distributed. This means that we have chosen a proper number of clusters.

However we calculate the number of spectra per cluster for the different *k* by means of the APE value. Results are shown in Table 2.

Table 2: Spectra by means of APE per cluster

	No. Of Clusters (Cluster) & % of Spectrum per Cluster														
APE	**4(1)**	**4(2)**	**4(3)**	**4(4)**	**5(1)**	**5(2)**	**5(3)**	**5(4)**	**5(5)**	**6(1)**	**6(2)**	**6(3)**	**6(4)**	**6(5)**	**6(6)**
1.79	0.01				0.01					0.01					
1.80	0.02				0.02					0.02					
1.81	0.07				0.07					0.07					
1.82	0.34				0.34					0.34					
1.83	0.86				0.86					0.86					
1.84	1.74				1.74					1.74					
1.85	3.22				3.22					2.87	0.35				
1.86	5.25	0.39			3.48	2.16				0.03	5.61				
1.87	0.1	8.38				8.49					8.46	0.03			
1.88		11.12				11.1	0.02				1	10.12			
1.89		6.47	11.08			1.78	15.76					13.04	4.5		
1.90			27.54			0.01	27.54					0.01	27.54		
1.91			10.69				5.07	5.62				0.01	8.65	2.04	
1.92			3.39	1.44				4.83					0.01	4.82	
1.93				2.73				2.74						2.74	
1.94				1.81				0.74	1.08					1.4	0.41
1.95				1.14				0.01	1.14					0.02	1.11
1.96				0.79					0.79						0.79
1.97				0.57					0.57						0.57
1.98				0.45					0.45						0.45
1.99				0.01					0.01						0.01

Looking at Table 2 we can see that a logical distribution of spectra per cluster is done. This can be explained because spectrum with similar APE value gets into the same cluster. This makes sense as APE indicates the contribution of irradiance per wavelength. Elements with very low difference on the APE value are considered as similar.

A very interesting result that can be seen from Table 2 is the fact that the percentage of spectra with the same APE value that are in two different clusters is very low.

Finally, if we pay attention to the contour graph in Fig. 4 and we look at the APE axis we can see that they are grouped in five different colors. If we choose a cluster per color then we will make *k=5* as the number of cluster that explains the behavior of analyzed photovoltaic module. That is, the 5 types of c.p.d.f., and then types of spectra.

5 CONCLUSIONS

The spectra measured for Malaga can be represented using only 5 different types of spectrum. The 5 spectra explain the different distributions of spectral irradiances observed in Malaga. The parameters used for selecting these spectra are the cumulative probability distribution of spectra and the APE value of each spectrum.

We prove that all the spectra recovered at the laboratory of photovoltaic systems in Malaga, more than 250.000 spectral irradiances can be grouped into a reduced number of clusters and that the cluster centroid characterizes every spectrum on it. For clustering all measured spectra we have used the k-means data mining technique.

We have found that once calculated the APE value of one spectrum it will be assigned to one of the five clusters according to Table 2 characterizing spectrum.

Although this study has been done at a specific location it will probably be valid the same assumption for different locations.

6 REFERENCES

[1] Y. Hirata and T. Tani, Solar Energy, Vol.55, no.6 (1995) 463-468.

[2] F. Fabero and F. Chenlo, Proceedings of the 22nd IEEE Photovoltaic Specialists Conference. IEEE Press (1991) 812-817.

[3] R. Gottschalg, T. R. Betts, D.G. Infield and M. J. Kearney, Meas. Sci. Technol., Vol. 15 (2004) 460–466

[4] S. Nann and C. Riordan, J. Appl. Meteor., Vol. 30 (1991) 447-462.

[5] S. Z. Selim and M.A. Ismail, Pattern Analysis and Machine Intelligence, IEEE Transactions, Vol. PAMI-6, no. 1 (1984) 81-87.

[6] G. H. Ball and D. J. Hall, Behavioral Science, Vol 12, no.2 (1967) 153–155.

[7] T. Gonzalez, Theoretical Computer Science, Vol. 38, no. 2–3 (1985) 293–306.

[8] T. Minemoto, S. Nagae and H. Takakura, Solar Energy Materials & Solar Cells, Vol. 91 (2007) 919-923.