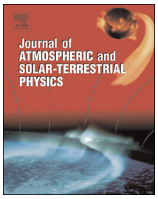
Journal of Atmospheric and Solar-Terrestrial Physics 130-131 (2015) 142–150

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar-Terrestrial Physics

journal homepage: www.elsevier.com/locate/jastp



Research Paper

Statistical functions and relevant correlation coefficients of clearness index

Diego Pavanello [a,*], Willem Zaaiman [a], Alessandra Colli [b], John Heiser [b], Scott Smith [b]

[a] European Commission DG Joint Research Centre, Institute of Energy and Transport, Renewable Energies Unit, 21027 Ispra, Italy
[b] Brookhaven National Laboratories, 11973 Upton, NY, USA

ARTICLE INFO

Article history:
Received 19 February 2015
Received in revised form
20 May 2015
Accepted 21 May 2015
Available online 9 June 2015

Keywords:
Solar radiation
PV system
Statistical analysis
Fuzzy logic
Correlation coefficients

ABSTRACT

This article presents a statistical analysis of the sky conditions, during years from 2010 to 2012, for three different locations: the Joint Research Centre site in Ispra (Italy, European Solar Test Installation – ESTI laboratories), the site of National Renewable Energy Laboratory in Golden (Colorado, USA) and the site of Brookhaven National Laboratories in Upton (New York, USA). The key parameter is the clearness index k_T, a dimensionless expression of the global irradiance impinging upon a horizontal surface at a given instant of time. In the first part, the sky conditions are characterized using daily averages, giving a general overview of the three sites. In the second part the analysis is performed using data sets with a short-term resolution of 1 sample per minute, demonstrating remarkable properties of the statistical distributions of the clearness index, reinforced by a proof using fuzzy logic methods. Successively some time-dependent correlations between different meteorological variables are presented in terms of Pearson and Spearman correlation coefficients, and introducing a new one.

1. Introduction

The penetration of PV systems into the overall energy budget creates a series of consequences to the grid behavior during the daylight hours, when the PV contribution is most important. In particular, the PV production depends on the variability of the sky conditions; fluctuations of the incident irradiance are directly correlated to the energy fluctuations of the grid system. This becomes more evident for short-term fluctuations, which can cause voltage flickers and imbalances during the day. PV installers and operators have to deal not only with long-term climatic characteristics of the site when designing a PV system, but also with short-term characteristics, which are dominated by statistics instead of periodicity and determinism.

The sky clearness index k_T is a dimensionless form of the irradiance: it's the ratio between the measured global horizontal irradiance and the extraterrestrial irradiance incident on a horizontal plane at the same instant of time. The interest in the properties of the statistical cumulative distributions and probability density functions of the clearness index started in 1960, after the work of (Liu and Jordan, 1960), in which a set of generalized curves of monthly distributions based on daily k_T values were proposed. The cumulative function represents the fractional time distribution of insolation, while the density function describes the two main levels of irradiance for the site in cloudless and cloudy conditions. Other successive works investigated the statistical properties of solar radiation using different time resolutions, from daily averages (Bendt and Collares-Pereira, 1981; Saunier et al., 1987), five minutes intervals (Jurado et al., 1995) and one minute intervals (Colli et al., 2014; Tovar et al., 1998; Suehrcke and McCormick, 1988). The analysis of high time resolution data sets permits to include in the analysis the optical air mass, which has an important impact on the shape of the statistical functions (Tovar et al., 1998), the correlations with other meteorological variables for specific sites (Colli et al., 2014), and the analysis of fluctuations of the irradiance which have an important role in PV energy production (Woyte et al., 2007; Peled and Appelbaum, 2013). The increasing volume of available data from different sites and climates is symptomatic of the general interest of a deeper knowledge of solar radiation, and offers the background for modeling (Badescu, 2008; Hollands and Suehrcke, 2013) and databases for applications (Gueymard and Thevenard, 2009).

1.1. Data acquisition and instrumentation

In order to perform the analysis, data sets collected with instantaneous values at 1-min rate through all the years 2010, 2011 and 2012, both during daylight hours and nights, have been collected at three different locations: (A) the Joint Research Centre site in Ispra (Italy, European Solar Test Installation – ESTI

---

* Corresponding author.
E-mail address: diego.pavanello@ec.europa.eu (D. Pavanello).

http://dx.doi.org/10.1016/j.jastp.2015.05.012
1364-6826/

Table 1
Instrumentation for irradiance measurements.

JRC	
Global Horiz.	Pyranometer Kipp&Zonen CM11
Diffuse	Pyranometer Kipp&Zonen CM10
DNI	Pyrheliometer CHP1
BNL	
Global Horiz.	Pyranometer Kipp&Zonen CMP22
Diffuse	Pyranometer Kipp&Zonen CMP22
DNI	Pyrheliometer CHP1
NREL	
Global Horiz.	Pyranometer Kipp&Zonen CM22
Diffuse	Pyranometer Kipp&Zonen CM22
DNI	Pyrheliometer CH1

laboratories, coordinates 45°48′N, 8°37′E, altitude 220 m); (B) the site of National Renewable Energy Laboratory in Golden (Colorado, USA, coordinates 39°74′N, 105°18′W, altitude 1829 m); (C) the site of Brookhaven National Laboratories in Upton (New York, USA, coordinates 40°52′N, 72°53′W, altitude 27 m, only for 2012). Not only global horizontal irradiance has been acquired, but other relevant meteorological variables, such as ambient temperature at ground, relative humidity, ambient pressure, DNI have been recorded. Referring to irradiance only, the following Table 1 gives an overview of the instrumentation used by the three laboratories.

The specification and classification of pyranometers and pyrheliometers is provided by the ISO 9060 International Standard. Pyranometers have been used to measure the hemispherical normal solar radiation, and all the models used for the present study are classified as secondary standard according to ISO 9060:1990 par. 4.3. The pyrheliometers have been used to measure the DNI, and are classified as first class instruments according to ISO 9060:1990 par. 5.3. For such instruments, the combined expanded uncertainty is about $\pm 2\%$ of the measured value.

In the present analysis the data sets have been filtered, according to the following criteria:

a. points with negative sun elevation;
b. points with $k_T < 0$ or $k_T > 1$;
c. points with global horizontal irradiance $G_H > 1200\ \mathrm{W/m^2}$ or $G_H < 10\ \mathrm{W/m^2}$.

The aim of point "a." is to remove points acquired during nights, while points "b." and "c." are aimed to reduce the data sets within a range of values suitable for the actual standard procedures of calibration and testing of PV devices. Even if values with $k_T > 1$ are possible in conditions of atmospheric concentration of sunlight, they have been removed because not suitable for PV calibration measurements, and typically they appear in not stable conditions.

Hereafter, if not otherwise specified, the analysis refers to the filtered data sets.

Moreover, the three laboratories synchronized the clocks of their data acquisition systems to the NIST (Boulder, USA) time server, in order to minimize time errors which may decouple the measured irradiance and the computed extraterrestrial one (Suehrcke, 1994).

1.2. Solar constant

The extraterrestrial irradiance has been computed following the algorithm proposed by NREL (Reda and Andreas, 2005); the most important variables are the geographical position on the Earth's surface, the period of interest, and the solar constant to be used. The solar constant is the rate at which solar energy is impinging upon a unit surface, normal to the rays, in free space, at the Earth's average distance from the Sun. Consequently, even if the solar constant varies during the year, a unique consistent value is used through this study, equal to $1367\ \mathrm{W/m^2}$.

2. Analysis with daily values

The clearness index is sensitive to both long-term effects and short-term effects. As already stated, the former ones depend upon the Earth's movement and are described by astronomy, the latter ones depends upon atmospheric effects and are described by statistics. A suitable parameter for the long-term analysis is the daily $k_{T,day}$, given by the equation

$$k_{T,day} = \frac{H}{H_0} \tag{1}$$

where H is the measured global horizontal solar irradiation at ground during the day and H_0 the extraterrestrial solar irradiation during the same day for the same location.

Following the criteria proposed by (Colli et al., 2014), it is possible to make a classification of the daily average sky conditions of a specific location (see Table 2).

An analysis of this kind is suitable to make considerations about the general features of a location, but the short-time fluctuations are not described. For the three sites used in this analysis, Fig. 1 shows the fraction of days falling into each day type during the three years. Each fraction $f_{i,\%}$ has been calculated by dividing the number of days in the specific category n_i for the total number of days N:

$$f_{i,\%} = 100\frac{n_i}{N} \tag{2}$$

This approach evidences that the two sites near sea level have similar daily conditions, while the site at 1829 m over sea level is less influenced by clouds, having 75% of clear or sunny days, compared with the 55% and 59% of the other sites. The corresponding cloudy days reduce from 25% down to 7% only (Fig. 2).

The monthly exploded diagrams show more in detail this difference, highlighting the different behavior during months; having a number of daily values comprised between 28 and 31 for each month, this subdivision is the most possible detailed one using daily values, not having a sufficiently large population of samples to compute density functions. Nevertheless, this information is interesting in the preliminary phase of the choice of a site, suitable for the installation of a PV plant.

3. Analysis with 1-min values

3.1. Density functions

The picture changes if the timescale of the analysis changes. Having a large population of samples, it is possible to compute density and cumulative functions. The statistical effects of clouds become more and more evident while reducing the timeframe for

Table 2
Day classifications used in this analysis.

$k_{T,Day}$	Day class
< 0.3	Cloudy
$0.3 \le k_{T,Day} \le 0.5$	Partly cloudy
> 0.5	Sunny or clear