Solar Energy 158 (2017) 49–62

Contents lists available at ScienceDirect

Solar Energy

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

Quality control of global solar radiation data with satellite-based products

Ruben Urraca[a], Ana M. Gracia-Amillo[b,*], Thomas Huld[b], Francisco Javier Martinez-de-Pison[a], Jörg Trentmann[e], Anders V. Lindfors[f], Aku Riihelä[f], Andres Sanz-Garcia[c,d]

[a] EDMANS Group, Department of Mechanical Engineering, University of La Rioja, Logroño, Spain
[b] European Commission, Joint Research Centre, Via E. Fermi 2749, 21027 Ispra, Italy
[c] University of Helsinki, Viikinkaari, 5 E, P.O. Box 56, 00014 Helsinki, Finland
[d] Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan
[e] Deutscher Wetterdienst, 63067 Offenbach, Germany
[f] Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland

ARTICLE INFO

Keywords:
Satellite-based models
Global horizontal irradiance
Quality control (QC)
Ground radiation records
Solar radiation

ABSTRACT

Several quality control (QC) procedures are available to detect errors in ground records of solar radiation, mainly range tests, model comparison and graphical analysis, but most of them are ineffective in detecting common problems that generate errors within the physical and statistical acceptance ranges. Herein, we present a novel QC method to detect small deviations from the real irradiance profile. The proposed method compares ground records with estimates from three independent radiation products, mainly satellite-based datasets, and flags periods of consecutive days where the daily deviation of the three products differs from the historical values for that time of the year and region. The confidence intervals of historical values are obtained using robust statistics and errors are subsequently detected with a window function that goes along the whole time series. The method is supplemented with a graphical analysis tool to ease the detection of false alarms.

The proposed QC was validated in a dataset of 313 ground stations. Faulty records were detected in 31 stations, even though the dataset had passed the Baseline Surface Radiation Network (BSRN) range tests. The graphical analysis tool facilitated the identification of the most likely causes of these errors, which were classified into operational errors (snow over the sensor, soiling, shading, time shifts, large errors) and equipment errors (miscalibration and sensor replacements), and it also eased the detection of false alarms (16 stations). These results prove that our QC method can overcome the limitations of existing QC tests by detecting common errors that create small deviations in the records and by providing a graphical analysis tool that facilitates and accelerates the inspection of flagged values.

1. Introduction

Solar radiation has been historically recorded at ground level by different meteorological agencies in order to provide reliable data for the assessment of the solar resource. These records are not only the most accurate source of solar radiation data, but are also crucial for validating satellite-based models, which are swiftly becoming the most widely used option to obtain spatial estimates of solar radiation (McArthur, 2005; Polo et al., 2016; Sengupta et al., 2015). Different parameters can be used to measure the amount and type of solar radiation reaching the Earth. The most common one is the global horizontal irradiance (G), i.e. the total shortwave incoming radiation received by a horizontal surface. More specialized monitoring stations can also measure the radiation components, i.e. the direct normal irradiance (B_N) and diffuse horizontal irradiance (D), providing more information about the type of radiation being received. Other parameters also recorded are the longwave radiation (upwelling and downwelling) and the sunshine duration, a parameter historically used to indirectly estimate G. However, in most cases the parameter normally recorded is only the G and the B_N and D are derived using decomposition models (Gueymard and Ruiz-Arias, 2015).

Measuring G is more prone to errors than other meteorological variables (Moradi, 2009). Younes et al. (2005) proposed the classification of these errors into two broad groups: equipment and operational errors. Equipment errors are inherent to the type of pyranometer used and the calibration applied, and include the zenithal error (cosine error), azimuthal error, stability, non-linearity, temperature dependence and spectral response. Highest-quality records are obtained with thermopile pyranometers, which are based on the thermoelectric effect. Within thermopiles, three levels of quality are established by the ISO

* Corresponding author.
E-mail address: ana.gracia-amillo@ec.europa.eu (A.M. Gracia-Amillo).

http://dx.doi.org/10.1016/j.solener.2017.09.032
Received 5 May 2017; Received in revised form 17 August 2017; Accepted 15 September 2017

Nomenclature

CM-SAF	Satellite Application Facility on Climate Monitoring
B	beam/direct surface irradiance received on a horizontal plane
B_N	beam/direct surface irradiance received on a plane always normal to sun rays
BSRN	Baseline Surface Radiation Network
CAMS	Copernicus Atmosphere Monitoring System
CI	Confidence Interval
D	diffuse surface irradiance received on a horizontal plane
E_{0N}	solar constant adjusted to Earth - Sun distance
E_0	extraterrestrial irradiance received on a horizontal plane
ECMWF	European Center for Medium-range Weather Forecast
G	global surface irradiance received on a horizontal plane
K	diffuse ratio
KN	beam transmittance
KT	clearness index
MAD	Median Absolute Deviation
MAE	Mean Absolute Error
MFG	Meteosat First Generation
MSG	Meteosat Second Generation
n	parameter to adjust the level of restriction (width) of the CIs
PV	photovoltaic
QC	quality control
w	window width, i.e. number of consecutive days analyzed at once
WMO	World Meteorological Organization

Greek letters

δ	deviation (estimated - observed)
θ_S	solar zenith angle

Subscripts

d	day
g	group of stations - spatial group
h	hour
m	total months of the time series
m'	twelve months of the year (Jan. to Dec.)
s	station

Superscript

p	product

9060:1990 (ISO, 1990) and World Meteorological Organization (WMO) (WMO, 2008): (i) Secondary Standard or High quality, (ii) First Class or Medium quality and (iii) Second Class or Low Quality. A low-cost option to record solar radiation is the use of radiometers based on the photovoltaic effect, such as silicon-based photodiodes and solar-reference cells. However, they are considered not compliant with the quality rules of the ISO 9060:1990 due to the limited spectral response of the silicon (400–1000 nm). On the other hand, operational errors are independent of the type of sensor and involve different factors such as shading by nearby objects, dew, frost, snow or dust (soiling) covering the dome of the pyranometer, incorrect leveling, station shut-downs, electric fields in the vicinity of cables or a malfunction in the datalogger, among others. An adequate selection of the place to install the pyranometers, as well as a regular maintenance, can prevent most of these operational errors. Another classification proposed by Zahumenský (2004) distinguishes between random errors, which are symmetrically distributed around zero, systematic errors, asymmetrically distributed, large errors mainly caused by malfunctions of the devices and errors in data processing, and micrometeorological errors, which are incoherences of the ground records compared to the surrounding regions. Overall, all types of error introduce a certain degree of uncertainty in the radiation measurements and applying a quality control (QC) procedure becomes an essential step before using ground datasets.

Many QC methods have been proposed by meteorological agencies and independent researchers. Some well-known examples are the QC tests from the Baseline Surface Radiation Network (BSRN) (Long and Dutton, 2002), the MESOR recommendations (Hoyer-Klick et al., 2008), the NREL SERI QC procedure (NREL, 1993), the QCRad methodology (Long and Shi, 2008) or the web-based services from MINES ParisTech (Geiger et al., 2002) and AQC test (Molineaux and Ineichen, 2003). These QC procedures flag those samples identified out of the normal ranges of data and usually leave the decision of removing flagged cases to the user. They can be classified in four broad categories (Ohmura et al., 1998): range tests (physically and extremely rare limits), across-quantities relationships, model comparisons and graphical analysis. The least restrictive level of most QC procedures is a range test based on the physically possible limits, with the upper limit being equal to the extraterrestrial irradiation (E) and a lower limit lying within −4 and 0 W/m^2 (Long and Dutton, 2002; Hoyer-Klick et al., 2008; Long and Shi, 2008). In a second step, the physical ranges are narrowed imposing more strict conditions. The upper limit is usually reduced with estimations from a clear-sky model (Geiger et al., 2002; Journée and Bertrand, 2011; Hoyer-Klick et al., 2008; Younes et al., 2005) such as the ESRA clear-sky model (Rigollier et al., 2000), the Page model (Page and Lebens, 1986) or the Bird clear-sky model (Bird and Hulstrom, 1980, 1981). Simulation with these clear-sky models are usually carried out under clean atmospheric conditions (aerosols and water vapor set to 0) instead of using estimated or climatological values. The lower limit is increased up to the level of extremely overcast conditions. This is typically imposed with the dimensionless clearness index ($KT = G/E$), with values around 0.03 (Geiger et al., 2002), or with the modified clearness index (Perez et al., 1990). Other approaches to reduce the acceptance ranges are to use climatological values (Long and Shi, 2008), to interpolate records from nearby locations or to use estimations from meteorological variables (Tang et al., 2010), mainly sunshine duration (Journée and Bertrand, 2011; Moradi, 2009; Muneer and Fairooz, 2002). Other QC tests also check the stability of the time series generated, analyzing the step between consecutive samples (Journée and Bertrand, 2011). Besides, some authors have proposed QC methods tailored to detect time shifts by analyzing the symmetry between morning and afternoon records (Ineichen, 2013) and by using graphical analysis (Moreno-Tejada et al., 2015). When the diffuse or direct components are also available, coherence or consistency tests are commonly imposed (Long and Dutton, 2002). Some authors have proposed the use of envelope tests in a dimensionless k-space consisting of the clearness index, the direct-beam clearness index ($KN = B/E$) and the global-to-diffuse ratio ($K = G/D$). These envelope tests are based on setting empirical or statistical limits either on the $KT-K$ space or the $KT-KN$ one (Younes et al., 2005; Journée and Bertrand, 2011; NREL, 1993; Pashiardis and Kalogirou, 2016) and the subsequent graphical analysis of the envelopes obtained.

The majority of these tests are designed to detect only large deviations in ground records, while some of the most common errors just introduce small deviations from the real irradiance profile. These faulty records pass most tests because they are acceptable from a statistical or a physical perspective, but they still have a negative influence on the assessment of solar radiation. Our aim is to develop a new QC