Sources of uncertainty in annual global horizontal irradiance data

Ruben Urraca[a,], Thomas Huld[b], Francisco Javier Martinez-de-Pison[a], Andres Sanz-Garcia[c]

[a] *EDMANS Group, Department of Mechanical Engineering, University of La Rioja, 26004, Logroño, Spain*
[b] *European Commission, Joint Research Centre, Via Fermi 2749, I-21027 Ispra, Italy*
[c] *University of Helsinki, Viikinkaari, 5 E, P.O. Box 56, 00014, Helsinki, Finland*

Abstract

The major sources of uncertainty in short-term assessment of global horizontal radiation (G) are the pyranometer type and their operation conditions for measurements, whereas the modeling approach and the geographic location are critical for estimations. The influence of all these factors in the uncertainty of the data has rarely been compared. Conversely, solar radiation data users are increasingly demanding more accurate uncertainty estimations. Here we compare the annual bias and uncertainty of all the mentioned factors using 732 weather stations located in Spain, two satellite-based products and three reanalyses.

The largest uncertainties were associated to operational errors such as shading (bias = -8.0%) or soiling (bias = -9.4%), which occurred frequently in low-quality monitoring networks but are rarely detected because they pass conventional QC tests. Uncertainty in estimations greatly changed from reanalysis to satellite-based products, ranging from the gross accuracy of ERA-Interim ($+6.1^{+18.8}_{-6.7}$%) to the high quality and spatial homogeneity of SARAH-1 ($+1.4^{+5.6}_{-5.3}$%). Finally, photodiodes from the Spanish agricultural network SIAR showed an uncertainty of $^{+6.9}_{-5.4}$%, which is far greater than that of secondary standards ($\pm$1.5%) and similar to SARAH-1. This is probably caused by the presence of undetectable operational errors and the use of uncorrected photodiodes. Photodiode measurements from low-quality monitoring networks such as SIAR should be used with caution, because the chances of adding extra uncertainties due to poor maintenance or inadequate calibration considerably increase.

Keywords: horizontal irradiance; uncertainty; satellite-based model; pyranometer; quality control

Nomenclature

Δ	difference between test and reference value
B	beam (direct) surface irradiance received on a horizontal plane
D	diffuse surface irradiance received on a horizontal plane
G	global surface irradiance received on a horizontal plane

Email address: ruben.urraca@unirioja.es (Ruben Urraca)

u uncertainty

Subscripts

d daily

h hourly

min minutely

t tilted

y annual

Superscript

est estimated

meas measured

ph photodiode pyranometer

ref reference value

ss secondary standard pyranometer

test test value

1. Introduction

Solar resource assessment is essential for many disciplines such as environmental sciences, climatology or energy production. They constantly demand more accurate solar radiation data with high spatial and temporal coverage, but there is also a growing interest on the uncertainty of the data. This information allows performing uncertainty propagation studies of models that use solar radiation as input (Thevenard and Pelland, 2013). A good example of such is yield estimations for new PV systems (Müller et al., 2017), where large uncertainties in solar data lead to high financial costs. A better understanding of these uncertainties would also contribute to mitigate their impact, as well as to select the best source of data for each application.

Uncertainty of solar radiation data depends on the source of data used and the type of radiation analyzed. Data is typically available as global horizontal irradiance (G), that is the surface-downwelling shortwave radiation received on a horizontal plane. For short-term assessments, the uncertainty of G primarily depends on whether the data is measured or estimated. The type of pyranometer and the maintenance procedures are the dominant factors in measured data (McArthur, 2005), while the quality of estimations strongly varies with the modeling approach (Urraca et al., 2017c). For long-term assessments, the inter-annual variability of solar radiation and the decadal trends, known as global dimming and brightening (Wild, 2009; Müller et al., 2014), must be also accounted. Additional uncertainties appear if other variables are used, such as the diffuse (D) and beam (B) components or the irradiance at tilted surfaces (G_t). This is because these variables are rarely measured, especially G_t, and are usually derived from G using decomposition (Gueymard and Ruiz-Arias, 2016; Moretón et al., 2017) and transposition (Ineichen, 2011; Gracia and Huld, 2013) models. Herein we will only address the sources of uncertainty in short-term assessment of G. We refer to the works listed above for evaluations of the uncertainties related to long-term effects, decomposition and transposition models.

Pyranometers are the most accurate source of G data when they are well-calibrated and properly maintained. The main types of outdoor sensors are thermopiles and silicon-based photodiodes (Vignola et al., 2012). Thermopiles are based on the thermoelectric effect and typically achieve the lowest uncertainty. They are the only ones compliant with the requirements of WMO (WMO, 2008) and ISO 9060:1990 (ISO, 1990), which classifies them from highest to lowest accuracy in (i) secondary standard, (ii) first class and (iii) second class. Photodiode sensors are based on the photovoltaic effect and are an attractive alternative to thermopiles for remote areas and agricultural monitoring stations because of their significantly lower cost and less maintenance. Besides, their fast time response makes them the detectors used by rotating shadowband irradiometers (RSI) with continuous rotation, which provide simultaneous records of G, B and D by shading and unshading the detector periodically (Sengupta et al., 2017). However, photodiodes generally have a lower accuracy than thermopiles, mainly due to the narrow spectral response of silicon. Overall, uncertainty in ground measurements largely varies with the type and cost of the instrument. Some of the factors limiting their accuracy are the cosine error, linearity, spectral effects and temperature dependence (Diresse et al., 2016). The implementation of corrections for these defects becomes essential to achieve acceptable uncertainties (Al-Rasheedi et al., 2018). All these errors are inherent to the sensor and its calibration, and are referred to as *equipment errors* (Younes et al., 2005; Journée and Bertrand, 2011). Excluding large deviations in the calibration constants, equipment errors cannot be detected with quality control (QC) methods and are commonly present in measured data.

The operation conditions of measurement stations introduce additional uncertainties in ground records. This is the case of shading by surrounding objects, accumulation of dust or snow, incorrect leveling of the sensor and electronic problems (Younes et al., 2005; Journée and Bertrand, 2011). All of them are referred to as *operational errors* and are independent of the type of sensor employed. Their magnitude highly varies with the severity of the defect but is generally larger than that of equipment errors. The probability of detecting operational errors should be therefore higher, but common defects, such as shading and soiling, produce acceptable records from a physical perspective. Hence, finding most operational errors in practice is also unlikely (Urraca et al., 2017a) and they are frequent in ground datasets, especially on those from low-quality networks and stations under extreme weather conditions.

Estimations are used in the absence of ground records, which is the most common case due to the sparsity and limited temporal coverage of ground stations. Satellite-based and reanalysis models are the most extended approaches (Bojanowski et al., 2014; Urraca et al., 2017c), as they provide long-time series with spatially continuous estimations. Satellite-based models use images from geostationary and polar-orbiting satellites to estimate cloud properties, and are the most popular method due to their superior quality (Sengupta et al., 2015; Polo et al., 2016). Reanalyses are based on the combination of numeric weather prediction (NWP) models with ground and satellite observations, but they generally have less accuracy than satellite-based models, mainly due to their coarse spatial resolutions (30-80 km). On their plus side, they provide hourly estimations of surface irradiance with global coverage, without gaps and include many other climatic variables. The uncertainty of these products greatly varies spatially because it depends not only on the characteristics of the database but also on the particular conditions of the place being assessed (Urraca et al., 2017b).

The uncertainty in some of the sources of solar radiation data listed above has been analyzed individually. Radiation databases are commonly validated against measurements from high quality ground stations (Suri and Cebecauer, 2014; Ineichen, 2014; Bojanowski et al., 2014; Urraca et al., 2017b, 2018). The uncertainty of thermopiles (Habte et al., 2015; Vuilleumier et al., 2014; Reda, 2012) and photodiodes (Al-Rasheedi et al., 2018; Wilbert et al., 2015; Geuder et al., 2014) has been evaluated with side-by-side comparisons against reference sensors, limiting the

number of radiometers used in these studies. The magnitude of cosine errors, linearity effect, temperature dependence and spectral mismatch has also been estimated (Sengupta et al., 2012; Driesse et al., 2015; Diresse et al., 2016). Nonetheless, uncertainties in estimated and measured data are rarely evaluated and compared within a common framework, and there is a lack of information about the impact of operational errors because they are rarely detected by QC tests.

Our main goal in this study is to evaluate the uncertainty in annual G associated to (i) estimations, (ii) operational errors and (iii) equipment errors. For that, we use measurements from 732 Spanish stations and estimations from two satellite-based products and three reanalyses. The study is conducted in annual terms from 2005 to 2013 because this is the common temporal resolution used for the prospection of new PV systems (Müller et al., 2017); however daily uncertainties are also reported for the comparison with previous studies. In case of equipment errors, the lack of collocated reference radiometers hinders a strict uncertainty estimation. We assume that the field uncertainty of the 53 secondary standard thermopiles from the national meteorological network should be close to their calibration uncertainty. Based on this, we make a rough estimate of the uncertainty of photodiodes that have a secondary standard closer than 20 km. Note that part of this uncertainty may be due to the reference instrument and to the validation procedure. Finally, using a novel QC method (Urraca et al., 2017a) we detect small operational errors by comparing the measurements against estimations from different independent radiation databases. The cause of each operational error is identified by visual inspection of the plots generated with the QC method, allowing the estimation of the uncertainty associated to each type of operational defect.

2. Data

2.1. *Measurements: Weather stations*

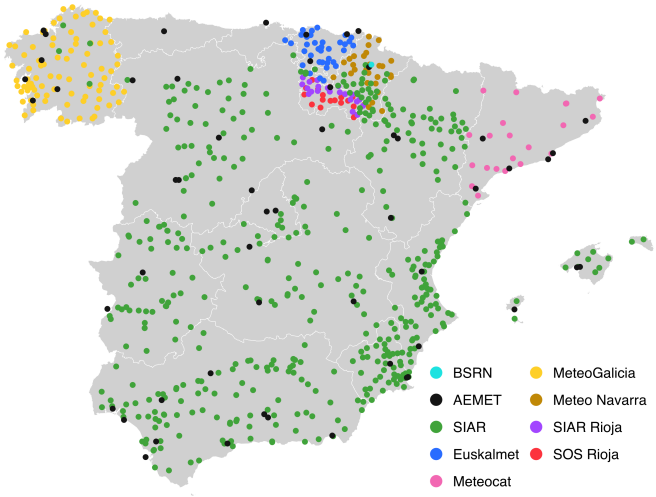

Fig. 1: Locations of the monitoring stations used in the study.

Ground records of G were retrieved from all Spanish weather stations that provided them at no cost (Fig. 1). This results in a ground dataset comprised by 732 stations distributed in 9 networks, including global networks such as the Baseline Radiation Surface Network (BSRN)

(BSRN, 2017), national networks such as the Agencia Estatal de Meterología (AEMET) (AEMET, 2017) and the Servicio Integral de Asesoramiento al Regante (SIAR) (SIAR, 2015), and several regional networks such as Meteo Navarra (Meteo Navarra, 2017), Meteocat (Meteocat, 2017), Euskalmet (Euskalmet, 2017), MeteoGalicia (MeteoGalicia, 2017), the SIAR branch in La Rioja (SIAR Rioja) (SIAR La Rioja, 2017) and SOS Rioja (SOS Rioja, 2017). G was downloaded at the highest temporal resolution freely provided for the period 2005-2013 (Table 1). Only years with at least 7500 hours of valid data were included in this study. The number of years available varies between the stations. For each station, the information about the pyranometer model was collected and sensors were accordingly classified as thermopile (281 stations) and photodiode pyranometers (380 stations). Note that some SIAR stations belong to the Spanish Ministry of Agriculture and some others to the regional governments, so some differences may exist in the maintenance and calibration routines.

Table 1: Description of the monitoring networks included in this study.

				Type of pyranometer [b]					
Network	Type [a]	Extent	Temporal resolution	SS	FC	SC	Ph.	NR	**Stations**
BSRN	met.	World	1 min	1	-	-	-	-	**1**
AEMET	met.	Spain	1 d	53	-	-	-	-	**53**
SIAR	agr.	Spain	30 min	-	19	35	348	66	**468**
Meteo Navarra	met.	Navarra	1 d	26	-	-	-	-	**26**
Meteocat	met.	Cataluña	1 d	-	15	-	5	-	**20**
Euskalmet	met.	Euskadi	10 min	43	-	-	-	-	**43**
MeteoGalicia	met.	Galicia	10 min	1	21	34	27	1	**84**
SIAR Rioja	agr.	La Rioja	1 h	-	-	21	-	-	**21**
SOS Rioja	eme.	La Rioja	1 h	-	12	-	-	4	**16**
Total				**124**	**67**	**90**	**380**	**71**	**732**

[a] rad. = radiometric, met. = meteorological, agr. = agricultural, eme. = emergencies.
[b] SS = secondary standard, FC = first class, SC = second class, Ph. = photodiode, NR = Not Reported

2.2. *Estimations: Radiation products*

Two satellite-based products and three reanalyses were used to study the uncertainty of radiation products (Table 2). Both satellite-based products are developed by the Satellite Application Facility on Climate Monitoring (CM SAF): SARAH-1 (Müller et al., 2015b,c), using images from the Meteosat geostationary satellites, and CLARA-A1 (Karlsson et al., 2012), using observations from polar-orbiting satellites. Both databases have been produced following a semi-empirical modeling approach where the cloud coverage is calculated from the satellite images and then it is combined with estimations from a clear-sky model. Products based on geostationary satellites are the most widely used because these satellites have temporal resolutions up to 15-min and spatial resolutions around 3-6 km. However, they only cover latitudes within $\pm 65^{\circ}$. Conversely, polar-orbiting satellites cover the whole world but products using these satellites can only provide daily values. This is because each satellite passes over the equatorial region only twice per day (one diurnal observation and a complementary observation 12 hours later at night). Since 2000, CLARA-A1 uses records from 3-5 polar satellites (NOAA and Metop seires) that are distributed in different observation nodes, passing over a region at different diurnal times. Therefore, it uses 3-5 diurnal satellite images to calculate the cloud coverage, limiting its output resolution to daily averages. More detailed information and validations of these products can be found in the literature (Urraca et al., 2017b; Riihelä et al., 2015).

The reanalysis datasets selected are two global reanalyses and one high-resolution regional reanalysis. The two global reanalyses are developed by the European Centre of Medium-range Weather Forecast (ECMWF): ERA-Interim (Dee et al., 2011), which is currently the operational dataset of ECMWF, and ERA5 (ECMWF, 2017), which will substitute ERA-Interim by the end of 2019. Both products use a similar modeling approach but ERA5 presents higher spatial (31 vs. 81 km) and temporal (1h vs. 3h) resolutions. A preliminary validation of surface irradiance data from ERA5 (Urraca et al., 2018) revealed a substantial improvement in the estimation of surface irradiance compared to former global products such as ERA-Interim. The regional reanalysis used is COSMO-REA6 (Bollmeyer et al., 2015), developed by the Hans-Ertel-Centre for Weather Research of Deutscher Wetterdienst (HErZ/DWD) for Europe. COSMO-REA6, based on the regional NWP model from the DWD, uses a high-resolution grid of 6.2 km based on the CORDEX EUR-11 domain.

Table 2: Main features of the radiation products included in this study.

Product	Spatial Coverage	Period	Spatial resolution	Temporal resolution	Variables
ERA-Interim	Global	1979 - present	0.75° ×0.75°(81 km)	3 h	*G*
ERA5	Global	1950 - present[a]	0.28° × 0.28°(31 km)	1 h	*G*, *B*
COSMO-REA6	Europe	1995 - 2014	0.055° × 0.055°(6.2 km)	1 h	*B*, *D*
SARAH-1	Eurasia, Africa (65°S - 65°N)	1983 - 2013	0.05° × 0.05°(5 km)	1 h	*G*, *B*
CLARA-A1	Global	1998 - 2015	0.25° × 0.25°(25 km)	1 d	*G*, *B*

[a] Preliminary release: 2010 - 2016

3. Methods

3.1. Data aggregation and quality control

The daily means of *G* were calculated for all stations to homogenize the temporal resolution of the dataset. In the case of 1-min data, the 15-min averages were calculated if at least 5 minutes were available, and then the hourly means were obtained if all four 15-min values were valid. For time resolutions between 5-min to 30-min, the hourly means were directly calculated if all values were available. Finally, daily means were obtained by averaging the hourly means if at least 20 hourly values were present.

Ground data with sub-daily time resolution were quality controlled using the BSRN range tests for physically possible and extremely rare limits (Long and Dutton, 2002). The daily means of all stations were additionally verified using the procedure described in Urraca et al. (2017a), which compares records against estimations from three independent radiation products: SARAH-1, CLARA-A1 and ERA-Interim. First, the confidence intervals (CIs) within which the daily differences between products and records (ΔG_d) lie for a region and a time of the year were calculated. Then a window function was run at each station flagging periods of consecutive days where the deviations of the three products were out of these CIs. The CIs were defined as n times the mean absolute deviation (MAD) around the median:

$$CI = median(\Delta G_d) \pm n \times MAD(\Delta G_d) \quad (1)$$

where n is a tuning parameter to control the restriction level. The CIs were calculated in a monthly basis (temporal aggregation) and the same CIs were used for all Spanish stations (spatial aggregation). Only the 53 AEMET stations were used to obtain the CIS because they have ventilated secondary standard pyranometers and are strictly maintained. These CIs were used to filter all Spanish stations, including AEMET stations. If the QC method finds defects in AEMET stations, the CIs are recalculated excluding those records. The window function was

run two times, first looking for short-lived high deviations (n = 2.4, width = 20 days), and second looking for long-term small deviations (n = 0.4, width = 90 days). Two plots were generated for visual inspection of the errors detected: (i) the daily differences between all products and measurements, and (ii) the hourly irradiance values of the sensor and SARAH-1 (the only product with hourly resolution) overlapped. Samples flagged by the BSRN tests and the daily QC method were not automatically removed. They were instead analyzed using the two previous plots, investigating the presence of false alarms and the cause of the errors.

Satellite-based and reanalysis data were also retrieved at the highest temporal resolution available (Table 2) and the aggregation procedure described for ground data was used to calculate the daily values. In case of SARAH-1, daily values were directly retrieved from the CM SAF website because hourly values present some gaps at sunrise and sunset and the use of an aggregation procedure without gap filling would introduce a bias in the daily values. For the products with high spatial resolutions (SARAH-1, CLARA-A1 and COSMO-REA6), values extracted are those of the corresponding pixel in the raster files. CLARA-A1 is included in this group because the available raster files (25 km) are a spatial average of an original resolution of 4 km. Inverse distance weighting (IDW) interpolation from the nearest four data points was used in products with coarse spatial grids (ERA-Interim and ERA5).

3.2. Analysis of the influencing factors

The influencing factors were evaluated in terms of the annual global annual irradiance (G_y) using the years with at least 312 available daily values (G_d). For each factor, the annual relative difference ($\Delta G_y[\%]$) between the irradiance provided by the source/factor being analyzed, the test irradiance (G_y^{test}), and the most likely true irradiance, the reference irradiance (G_y^{ref}), was calculated. The term "difference" (ΔG) was used in lieu of "error" to stress that reference values have their own uncertainty (Gueymard, 2014). The conventional method to obtain the annual differences is as follows:

$$\Delta G_y[\%] = \frac{G_y^{test} - G_y^{ref}}{G_y^{ref}} 100 \tag{2}$$

However, the calculation of G_y^{test} and G_y^{ref} separately introduces a systematic deviation in annual statistics if missing values are present either in the *test* or *ref* sets of data. Although this is mitigated using a minimum number of 312 days to calculate the annual differences, the effect of missing values can be removed estimating $\Delta G_y[\%]$ based on the annual relative bias ($bias_y$) of daily irradiances (Eq.3).

$$\Delta G_y[\%] \approx bias_y[\%] = \frac{\frac{1}{N_d}\sum_{d=1}^{N_d} G_d^{test} - G_d^{ref}}{\frac{1}{N_d}\sum_{d=1}^{N_d} G_d^{ref}} 100 \tag{3}$$

Based on Eq 3, the annual comparisons between *test* and *ref* sets are made at daily level guaranteeing that the same daily samples (N_d) are used from both sets. Besides, the use of at least 312 days assures that the annual statistics cover most of the annual variability. In the following sections, the specific details for the analysis of each individual factor are described, focusing on defining which are the *test* and *ref* irradiances in each case.

3.2.1. Estimations from radiation products

Estimations from radiation products (G_y^{est}) were analyzed using ground records from 53 AEMET secondary standard pyranometers ($G_y^{meas,ss}$) as reference values (Eq. 4).

$$\Delta G_y^{product} = G_y^{est} - G_y^{meas,ss} \quad (4)$$

Ground records were quality controlled by removing the periods flagged by the daily QC method. The analysis was limited to the period in which data from the five products were simultaneously available (2010-2013).

3.2.2. Operational errors

The quality flags resulting from the BSRN range test and the daily QC method were classified into the following categories: shading, soiling, time shifts, stability issues in the sensor (leveling), diurnal samples with zero radiation (diurnal G = 0) and large errors. The flags were visually inspected using two plots generated by the QC method. Examples of how each operational error is identified in these two plots can be found in Urraca et al. (2017a). Rain data was retrieved from stations including a pluviometer to ease the identification of soiling cases. In case of doubt with a potential failure, it was not considered for the analysis of operational errors. If the presence of a defect was clear but not its cause, samples were classified as "Unknown cause". Time shifts were not considered for the present analysis as they do not affect the annual irradiance. Years with more than one type of defect were also excluded. For each type of operational error, the annual difference between the ground measurements (G_y^{meas}) and SARAH-1 was calculated (Eq. 5).

$$\Delta G_y^{oper} = G_y^{meas} - G_y^{est,sarah} \quad (5)$$

The differences were calculated only at the stations and years in which an operational error was found, using all ground data available (732 stations, 2005-2013).

3.2.3. Equipment errors

Uncertainties associated with equipment errors depend on the pyranometer model and calibration procedure. In this study, the evaluation of equipment errors was simplified to the comparison of the 53 AEMET secondary standard pyranometers against the 348 SIAR photodiodes. Sensors from the same network were used in order to mitigate uncertainties related to the calibration procedures followed by the networks. Sensors from the regional networks, first class sensors and second class sensors were excluded because they are sparsely distributed in Spain so their inclusion could lead to additional uncertainties related to the spatial location.

AEMET sensors are secondary standard pyranometers from Kipp&Zonen with an achievable uncertainty at 95% CI of $\pm$3% and $\pm$2% for hourly and daily values respectively (Sengupta et al., 2017). The uncertainty of annual values should be smaller due to the compensation of seasonal deviations, so we assume an annual uncertainty of $\pm$1.5%. SIAR photodiodes are the SP1110 model (Skye Instruments), and their datasheet reports an absolute accuracy always better than 5% and most times under 3%. However, their specifications do not detail either the temporal resolution or the confidence level of this value. Besides, the field uncertainty of photodiodes varies according to the empirical corrections applied. Our goal is therefore to make a rough estimate of the real uncertainty of SIAR photodiodes. Two analyses were design to overcome the lack of collocated measurements from photodiodes and thermopiles.

In the first analysis, the differences between both types of sensor ($\Delta G_y^{ph|ss}$) and SARAH-1 were calculated (Eq. 6). SARAH-1 was defined as the reference because the goal is to evaluate uncertainties on ground records, but measurements from both types of sensors should be a priori closer to the true irradiance.

$$\Delta G_y^{ph|ss} = G_y^{meas,ph|ss} - G_y^{est,sarah} \quad (6)$$

The second analysis was based on the direct comparison of photodiodes ($G_y^{meas,ph}$) against secondary standard ($G_y^{meas,ss}$) (Eq. 7), using a subset of 34 stations equipped with photodiodes separated less than 20 km from a secondary standard sensor.

$$\Delta G_y^{ph} = G_y^{meas,ph} - G_y^{meas,ss} \tag{7}$$

Quality controlled records were used in both analyses removing the impact of operational errors (1 AEMET station, 74 SIAR stations). The analysis was restricted to the period 2007-2013, because the time series of most AEMET stations start in 2007. Stations selected were those with data for all the years from 2007 to 2013 avoiding the inclusion of any artificial trends.

3.3. Estimation of uncertainties from annual differences

Annual differences calculated in previous sections included the uncertainties of both test and reference values. The uncertainty of the test values alone was estimated following the GUM rules (JCGM 100:2008, 2008; JCGM 104:2009, 2009) in two steps: (i) calculating the random and systematic components of the annual differences, and (ii) removing the uncertainty of the reference values by using uncertainty propagation rules.

GUM analyzes random and systematic errors separately (Farrance et al., 2016), and it calculates the uncertainty (u) only from random errors because it considers that systematic errors can be detected and corrected. For instance, site-adaptation techniques are applied to eliminate systematic errors in radiation databases (Polo et al., 2016), whereas systematic deviations in radiometers are mitigated using empirical correction factors. However, from the perspective of users of solar radiation data, systematic errors will be generally present in either measured or estimated databases. Therefore, we kept the systematic error, hereafter referred as bias, for the analysis of the data but we reported it separately from the uncertainty ($bias \pm u$).

Parametric statistics were discarded to calculate the bias and uncertainty because the annual differences are not normally distributed. The bias was calculated using the median, a more robust central measurement than the mean for non-normal distributions. Uncertainties were estimated as the distance between the 95% CIs and the bias, where the 95% CIs were obtained using percentiles 2.5 ($p_{2.5}$) and 97.5 ($p_{97.5}$). The distance between the CIs and the median was not symmetric for non-normal distributions. Thus, both negative (u^-) and positive (u^+) uncertainties were reported (Eq. 8).

$$\pm u = \begin{matrix} u^+ \\ u^- \end{matrix} = \begin{matrix} +(p_{97.5} - bias) \\ -(bias - p_{2.5}) \end{matrix} \tag{8}$$

Biases in reference values were directly removed to obtain the bias of the test values (Eq. 9). Contrary, all reference values had their own uncertainty because the true irradiance is never known. The uncertainty of test values (u_y^{test}) was obtained following the functional relationships for uncorrelated uncertainties (Farrance and Frenkel, 2012). In the case of a subtraction ($\Delta = test - ref$), the uncertainties of the test and reference values (Eq. 10) add in quadrature. For non-normal distributions, the functional relationships were applied separately to $u+$ and u^-.

$$bias_y^{\Delta} = bias_y^{test} - bias_y^{ref} \rightarrow bias_y^{test} = bias_y^{\Delta} + bias_y^{ref} \tag{9}$$

$$(u_y^{\Delta})^2 = (u_y^{test})^2 + (u_y^{ref})^2 \rightarrow u_y^{test} = \sqrt{(u_y^{\Delta})^2 - (u_y^{ref})^2} \tag{10}$$

Daily uncertainties were obtained similarly to annual uncertainties, but using daily differences ($\Delta G_d = G_d^{test}$ - G_d^{ref}) instead of annual ones.

4. Results and Discussion

4.1. Estimations from radiation products

Satellite-based products exhibited the smallest bias overall (Fig 2), with mean biases of +1.4 and -1.6% for SARAH-1 and CLARA-A1, respectively. Our previous study made with 313 European weather stations (Urraca et al., 2017b) also showed a positive bias of +0.29% for SARAH-1, whereas Gracia Amillo et al. (2014) and Müller et al. (2015a) obtained mean biases of +0.73% and +1.3 W/m^2 (< 1%), respectively. In the same study (Urraca et al., 2017b), we reported a negative bias of -0.49% for CLARA-A1 over Europe. Similarly, Trentmann and Müller (2012) obtained a bias of -3.3 W/m^2 for CLARA-A1 using several BSRN stations worldwide. The bias was also low for ERA5 (+1.3%), but ERA5 showed greater variability than satellite-based products. ERA5 presented strong intra-annual variations, overestimating G on cloudy days and underestimating G on clear-sky days. We also found positive biases for ERA5 in a validation made at 277 European weather stations (+3.47%) and 40 BSRN stations (+2.92%) (Urraca et al., 2018). We observed the same bias variability, which is most likely caused by different failures in the modeling of clouds. The impact of using data with a low annual bias but high intra-annual variability should be further investigated.

ERA-Interim and COSMO-REA6 presented the largest bias among all the products evaluated. ERA-Interim showed a positive bias of +6.1% produced by a strong overestimation of G under cloudy conditions that aggravates the bias variability discussed for ERA5. According to Träger-Chatterjee et al. (2010) and Boilley and Wald (2015), this effect is related to an underestimation of the cloud fraction. The bias of ERA-Interim increased even more at coastal locations due to its coarse spatial resolution. The strong overestimation of ERA-Interim was also reported in the studies conducted over Europe by Bojanowski et al. (2014) (+9.08%) and Urraca et al. (2018) (+9.28%), as well as in the study published by Träger-Chatterjee et al. (2010) at Germany (+5.2 W/m^2). COSMO-REA6 showed the same bias variability than ERA5 and ERA-Interim, but it particularly underestimates G under clear-sky days. This led to a mean negative bias of -8.2%, in line with the biases of -9.2 W/m^2 ($\approx$ -5%) and -3.22% for Europe reported by Frank et al. (2018) and by Urraca et al. (2018), respectively. Frank et al. (2018) outlined that this underestimation is caused by the use of an aerosol climatology that overestimates the aerosol content, and they proposed a bias-corrected version to mitigate the bias variability.

The spatial variability of the results substantially varied between products. ERA-Interim exhibited the largest variability, with annual differences exceeding +20% along the Atlantic coast while being below +10% for most inland regions. This is due to the coarse spatial grid of ERA-Interim (81 km) that includes portions of water and land in a single grid point whereas solar irradiance patterns rapidly change in coastal areas (Hazuba et al., 2013). This variability was somewhat mitigated by the new ERA5 due to the finer grid used (31 km). COSMO-REA6 is the only reanalysis that partly corrects this issue due to its high-resolution grid (6.2 km). Both satellite-based products had the most homogeneous distributions for Spain, although certain spatial patterns were still observed. Nonetheless, these variations are negligible if compared to that of reanalysis products, making these types of products more reliable.

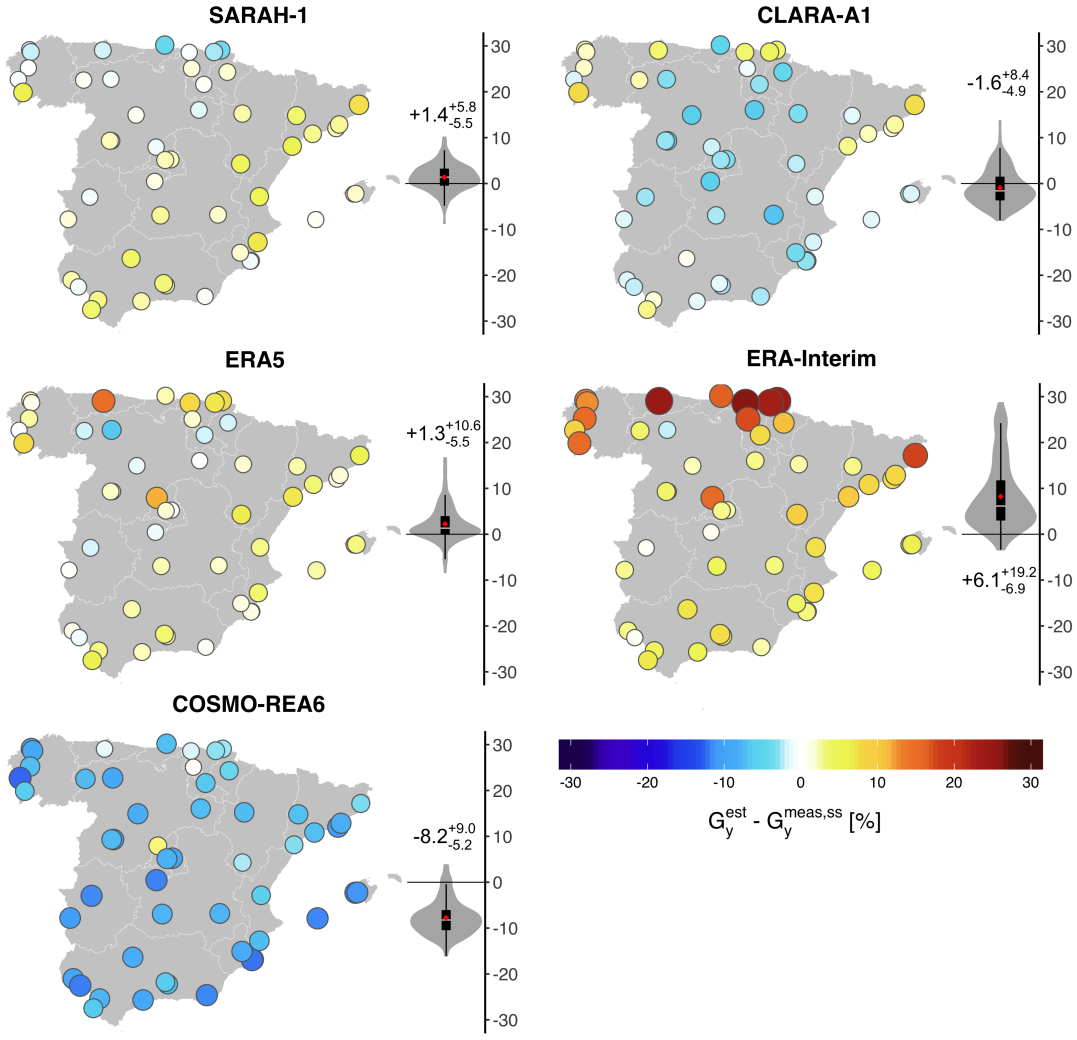

Fig. 2: Annual difference [%] between radiation products (G_y^{est}) and AEMET secondary standard pyranometers ($G_y^{meas,ss}$) in the period 2010-2013. Violin plots outline the kernel probability density. Boxplots visualize the lower quartile, median and upper quartile. The red diamond represents the mean.

Uncertainties in estimated values are caused by the particular characteristics of the radiation products and the conditions of the location to be assessed. Some of the strongest influencing factors are the spatial and temporal resolution, and the type of model implemented. Differences were observed not only between satellite-based or reanalysis models, but also between products within each type of technique. A key factor is the model representation of the atmospheric components that attenuate solar radiation, mainly clouds (e.g. ERA-Interim and ERA5) but also under clear-skies aerosols (e.g. COSMO-REA6) and water vapor. The influence of all these factors in the quality of the estimations also changes spatially, which is a major difference between ground records and radiation products. When estimating solar irradiance, all the regions covered by the product do not show the same uncertainty. Well-known examples of regions where satellite-based products have large uncertainties are the mountains, regions with bright surfaces (deserts and snow) and the edges of the satellite images (Suri and Cebecauer,

2014). On the contrary, shortcomings of reanalyses become more evident in cloudy regions and coastline locations (Urraca et al., 2017a).

4.2. Operational errors

Most operational failures introduced annual differences around 10% (Fig. 3), while some severe cases exceeded ±30%. The differences were generally negative, with some exceptions in the leveling and large errors groups. Note that these differences include the bias and uncertainty of SARAH-1 ($+1.4^{+5.6}_{-5.3}$%) so they are negatively biased. The true magnitude of the defects is shown in Subsection 4.4 after removing the bias and uncertainty of SARAH-1. Anyway, the magnitude of operational errors was much higher than that of SARAH-1, justifying the use of SARAH-1 as the reference irradiance. All the defects showed a high variability as a consequence of the different severity and duration of the defects. While the whole year was typically affected in the case of shading, same cases of electronic errors lasted just few days or few hours, having different impacts in the annual G. Excluding extreme cases of large errors and leveling, most operational errors passed the BSRN range tests and were only detected by the daily QC method, proving its capacity to detect low-magnitude defects.

Shading caused either by natural or artificial objects was the most common operational error (56 stations). The bias was constant from year to year in most stations, indicating a bad selection of the location, but the magnitude of the bias steadily increased in a few stations, suggesting a poor maintenance. In all cases it resulted in negative differences (bias = -9.4%) that exceeded -30% in extreme cases. Soiling was another common defect detected in 21 stations. It rarely lasted more than 2-3 months and disappeared with maintenance or rain. While shading typically occurs near sunset or sunrise, where the sun elevation is low, the attenuation due to soiling keeps constant during the day. This hindered the identification of low-intensity soiling errors because the resulting irradiance profile has a similar shape to the true irradiance. This may be the most likely explanation for the 6 stations showing unidentified errors and classified as "Unknown cause".

Large errors were also a common defect (23 stations) that involves malfunctions either in the instrumentation or in the data processing (Zahumenský, 2004). The effect can equally be positive or negative, lasting typically no more than one month and occurring even at night (MeteoGalicia and Euskalmet). Large errors shown in Fig. 3 only include daily values in the range -200 < G < 500 W/m^2, but more extreme values were even obtained in some stations with high temporal granularity such as Euskalmet. An incorrect leveling of the sensor was also detected in 10 stations, and this is the only type of defect along with large errors that can lead to positive differences when the pyranometer is facing towards the sun. A particular case of large errors are diurnal periods with G=0. These were found in 7 stations, without taking into account two SOS Rioja stations with G=0 all year round. This defect is probably caused by the data logger returning '0' instead of 'Not Available'. It could be straightforwardly identified by visual inspection, but calculations involving large datasets would make its detection difficult because conventional QC methods, such as the BSRN tests, cannot filter it.

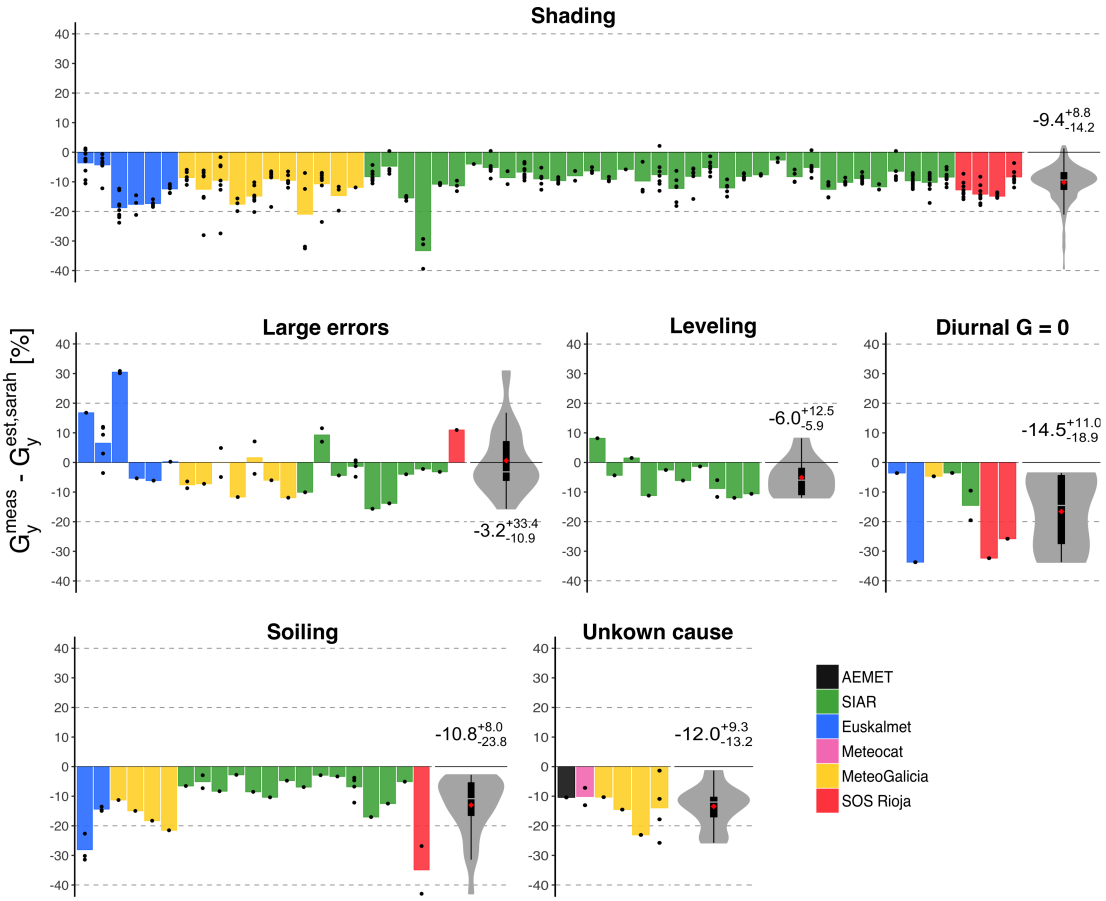

Fig. 3: Annual difference [%] between stations (G_y^{meas}) and SARAH-1 ($G_y^{est,sarah}$) for the stations and years in the period 2005-2013 in which defects were detected. Bars depict the bias at each station. Violin plots outline the kernel probability density. Boxplots visualize the lower quartile, median and upper quartile. The red diamond represents the mean. Operational errors were not found in BSRN, SIAR Rioja and Meteonavarra.

The number and the intensity of the defects strongly varied between the different monitoring networks (Fig. 3). One single case was found in AEMET and BSRN that implement the strictest maintenance procedures and keep their pyranometers ventilated reducing the accumulation of dust and humidity over the sensors. Similarly, SIAR Rioja, Meteonavarra and Meteocat showed only one defect, which suggests that the maintenance is acceptable as well. The majority of the errors occurred in SIAR, SOS Rioja, MeteoGalica and Euskalmet, which are either agricultural or regional networks. Several deficiencies exist in the maintenance as well as on the internal quality checks of these networks. In case of agricultural monitoring stations, the selection of the location might be also inadequate because it is made based on agricultural criteria instead of looking for the best location for solar radiation purposes. The number of defects detected also increased with the temporal resolution of the data available. For instance, most of large errors were detected in MeteoGalicia and Euskalmet, both providing 10-min data. As some of these short-lived defects get masked when aggregating to hourly or daily values, high-resolution data should be preferred to perform a good QC.

4.3. Equipment errors

4.3.1. Comparison between secondary standards and photodiodes using SARAH-1 as reference

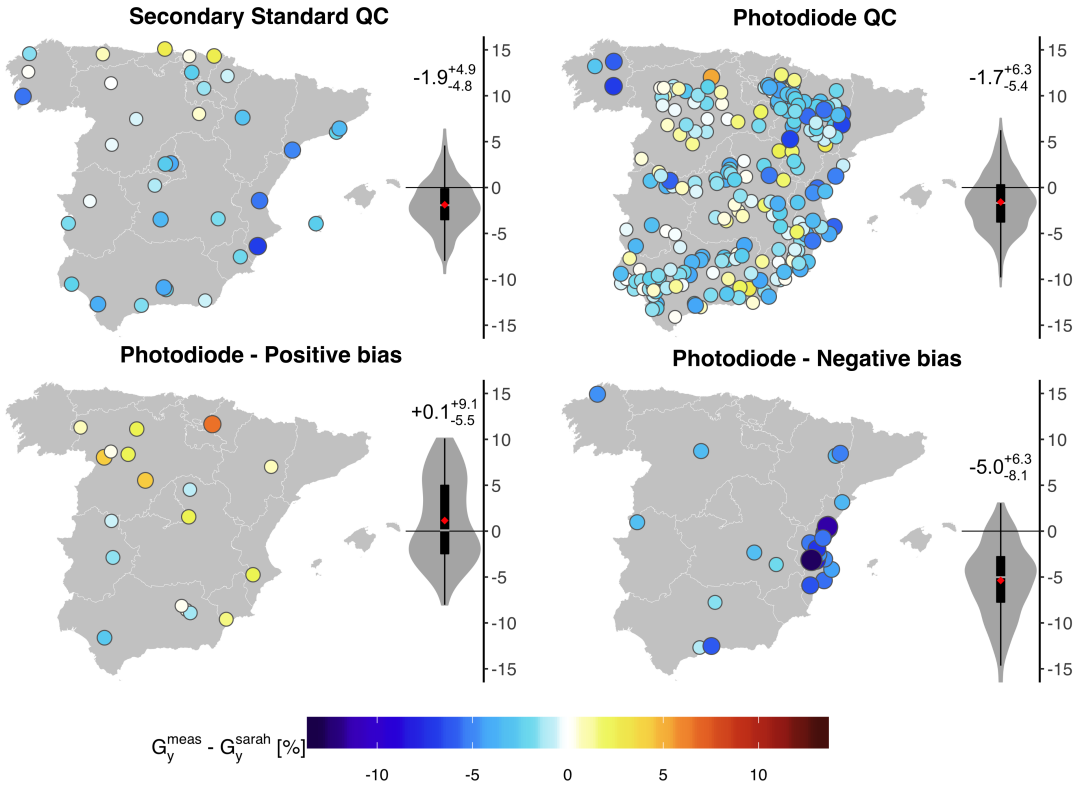

Fig. 4: Annual difference [%] between measurements (G_y^{meas}) and SARAH-1 ($G_y^{est,sarah}$) for quality controlled photodiodes (220 sensors) and secondary standards (35 sensors). Photodiodes without an evident operational error but flagged by the QC method are depicted separately (27 stations with periods of negative bias, 18 stations with periods of positive bias). Stations shown have valid annual values for all the years in the period 2007-2013. Violin plots outline the kernel probability density. Boxplots visualize the lower quartile, median and upper quartile. The red diamond represents the mean.

Annual differences between measurements (Fig. 4) and SARAH-1 presented a similar bias for thermoelectric and photoelectric detectors (-1.9 vs. -1.7%). The spatial distributions were also comparable, whereas photodiodes showed a slightly greater variability. The uncertainty of secondary standards is known to be substantially lower than that of photodiode pyranometers. Hence, the similarity between the distributions could be due to the fact that the uncertainty of photodiodes is close to that of SARAH-1, and random errors in sensors and products might be compensated. The results also indicate that the photodiodes are a priori unbiased, but the interannual analysis in Fig. 5 reveals that differences obtained with secondary standards steadily increased from 2007 to 2013, while that obtained with photodiodes randomly oscillated during the same period. The steady increase of ($G_y^{meas,ss} - G_y^{est,sarah}$), which means that the bias of SARAH-1 is decreasing, is in line with the trend of -1.7 W/m^2 ($\approx$ 1%) documented in the validations of SARAH-1 (Pfeifroth et al., 2016). This trend is caused by artefacts in the satellite product and by the use of an aerosol climatology, and it hinders the precise modeling of climate trends (Sanchez-Lorenzo et al., 2017). The oscillation of ($G_y^{meas,ph} - G_y^{est,sarah}$) is therefore caused

by artificial variations in the photodiodes. This is confirmed when the differences obtained with both types of sensor are subtracted (Fig. 5). Assuming that the bias obtained with secondary standard pyranometers is the true value, photodiodes overestimated around +2% in the period 2007-2009 and underestimated around -1% in the period 2011-2013. The most likely cause of this sharp change is a major calibration campaign around 2009-2011, but the only information provided by SIAR is the last calibration date (typically made in the last 2 years) and the installation date (around 2005-2006). Nonetheless, we found clear evidence that significant biases exist either in recently installed sensors (2007-2009) or in potentially re-calibrated photodiodes (2011-2013), while the spread of the differences is larger in photodiodes than in secondary standards.

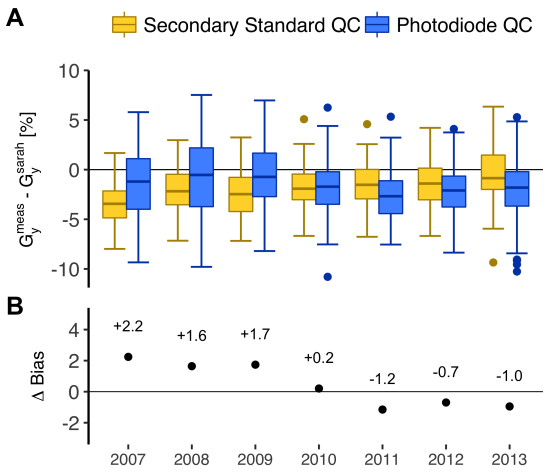

Fig. 5: (**A**) Inter-annual variation of the annual difference [%] between stations (G_y^{meas}) and SARAH-1 ($G_y^{est,sarah}$) for each type of pyranometer. (**B**) Difference between the bias of photodiodes and the bias of secondary standards against SARAH-1 ($\Delta Bias = Bias_y^{meas,ph} - Bias_y^{meas,ss}$). Stations shown have valid annual values for all the years in the period 2007-2013.

The previous comparison between photodiodes and secondary standards was made excluding not only operational errors but also 60 SIAR photodiodes (45 without gaps in 2007-2013) flagged by the QC method, in which it was not possible to identify operational errors. In all these stations, samples flagged were long-term periods of at least 90 consecutive days with a small but persistent bias around ± 50 W/m^2 between measurements and the three products. The hypothesis of a simultaneous failure of all radiation products, i.e. a false alarm of the QC, was discarded because this type of error was only found in SIAR photodiodes. Further, although some of the negative biases (36 stations) could be caused by a undetected operational errors, this type of error rarely produces long-term positive bias (24 stations). Thus, the most likely cause is an equipment error related to the photodiode pyranometers. This is why these stations were studied neither as an operational error of unknown cause nor in the Photodiode QC group, because low-severity operational errors might exist in the case of negative biases.

Annual differences at these stations were higher, showing values of -5.0% and +0.1% for stations with periods of negative and positive biases, respectively. The spread of these differences was similar to that of the main group of photodiode pyranometers. Most of these biases followed a random pattern, spanning from 3 to 6 months, at any season of the year, and happening typically just one year of the time series. The causes of these biases could be related to changes in temperature, aerosol or water vapor leading to temperature and spectral errors of the instrument (Driesse et al., 2015). The negative biases may be caused by a small accumula-

tion of dust over the pyranometer not detected during the QC visual inspection. On the other hand, some of the positive biases appeared for a consecutive number of years at the beginning of the time series (2005-2009). This agrees with the overestimation observed in Fig.5 of around +2% for that period, suggesting that the QC method flagged those stations with large deviations in the calibration constants. We assume that most of these errors were caused by limitations of the photodiodes (spectral response, cosine error or temperature dependence) and miscalibrations of the sensor, which increases the uncertainty reported for the main group of photodiodes (Photodiode QC).

4.3.2. Direct comparison between photodiodes and secondary standards

The direct comparison of photodiodes against secondary standard pyranometers corroborated that significant differences exist between the annual readings of the two types of sensors (Fig 6). The annual uncertainty was $^{+7.1}_{-5.6}$%, with 56% of annual differences out of the $\pm$1.5% uncertainty interval of secondary standards, while 15.3% of annual differences exceeded $\pm$5%. The number of years with positive differences was higher than the negative (60 vs. 40%), but the bias, and hence the sign of the differences, also showed a high inter-annual variation (Fig. 7). Similarly to the results obtained using SARAH-1 as reference, a sharp change was observed in many photodiodes from 2009 to 2011. Photodiodes overestimated annual G by around +2.5% in the period 2007-2009, while they underestimated it by around -1% in the period 2011-2013. Some of the inter-annual variations might be also caused by the different performance of photodiodes with the amount and type of irradiance received each year, but they are in a second order of magnitude compared to changes in the calibration constant.

Differences depicted in Figs. 6 and 7 include the uncertainty of secondary standard measurements, and additional uncertainties because the sensors compared are not exactly in the same place. However, no trend, neither in bias nor in the spread of the data, was observed with increasing distance, while some of the largest differences were in fact obtained between pairs of sensors closer than 8 km (Fig. 6). In addition, agricultural stations are typically installed in low-lying areas with low spatial variability of irradiance, which means that the spatial representativeness of these stations is typically high. This suggests that the uncertainty of the comparison is low so the differences observed between photodiode and secondary standard records actually exist.

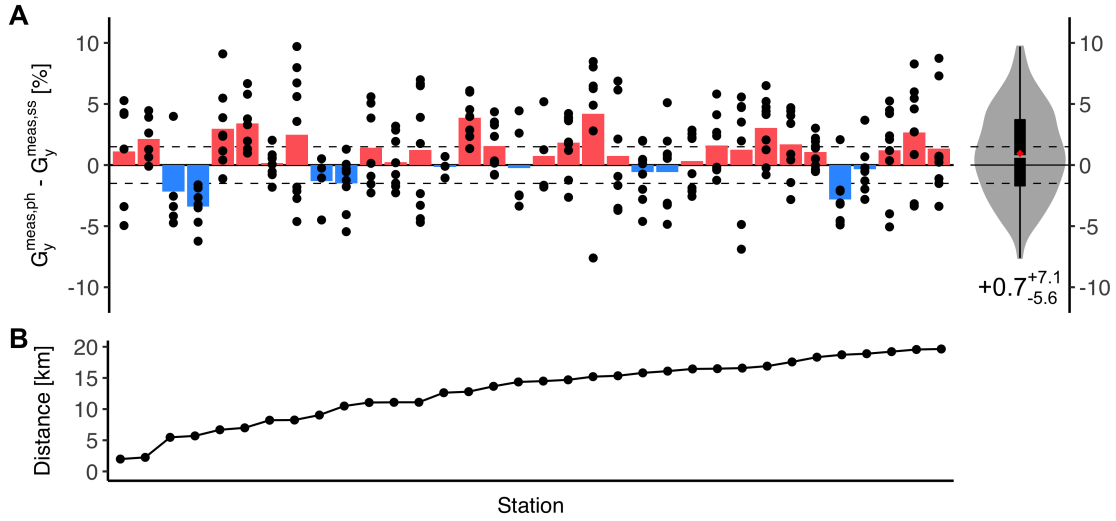

Fig. 6: (**A**) Annual difference [%] between photodiodes ($G_y^{meas,ph}$) and secondary standards ($G_y^{meas,ss}$), for the photodiodes that have a secondary standard in a radius of 20 km. (**B**) Distance between the photodiode and the secondary standard sensor being compared. The dashed line represents the uncertainty of secondary standard pyranometers. Violin plots outline the kernel probability density. Boxplots visualize the lower quartile, median and upper quartile. The red diamond represents the mean.

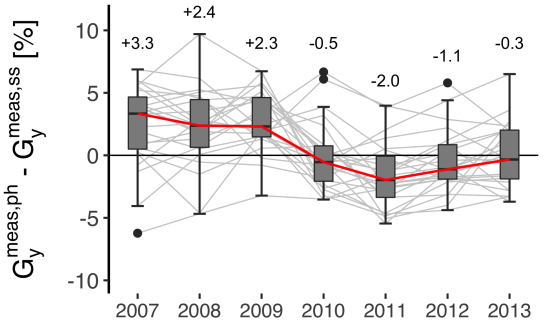

Fig. 7: Inter-annual variation of the annual difference [%] between photodiodes ($G_y^{meas,ph}$) and secondary standards ($G_y^{meas,ss}$), for the photodiodes that have a secondary standard in a radius of 20 km. The 24 pairs of stations visualized have valid annual values for all the years in the period 2007-2013. The grey lines show the inter-annual variation for each pair of stations. The red line visualizes the median. The black dots are extreme cases lying beyond 1.5 times the interquartile range (length of the whiskers).

Different factors could be behind the differences observed in photodiodes. The main sources of uncertainty in photoelectric detectors are the spectral response, cosine error, temperature dependence and linearity. Reda (2012) estimated that the contributed uncertainty of each of these defects was 5%, 2%, 1% and 1%, respectively. Driesse et al. (2015) analyzed the individual impact of linearity, spectral response, temperature and the angle of incidence on photodiode and thermopile pyranometers. They found positive errors up to +2% for high temperatures, up to +3% for medium and low angles of incidence and up to +2% for high irradiances, while negative errors up to -3% were only observed for low irradiance. Similarly, Sengupta et al. (2012) found an overprediction of photodiode pyranometers when compared to secondary standard

sensors in the morning and the afternoon. Both studies emphasized the spectral limitations of these pyranometers. The spectral response of silicon is limited to the range 350-1100 nm that includes about 70-75% of the total energy (Meyers, 2011), whereas 96% of broadband irradiance is received in the range from 300 to 3000 nm. Their calibration factors are consequently calculated to estimate the broadband irradiance from the narrowband measurement, but problems arise when the part of the spectrum not seen by the sensors varies non-linearly. This occurs with changes on the aerosols and water vapor concentrations, as well as with intra-daily and intra-annual variations of the sun elevation angle that modify the dominant scattering process in the atmosphere.

The previous defects introduce systematic deviations in the measurements that can be mitigated using empirical correction functions. Ideally, these corrections should be applied individually per defect, type of sensor and location after side-by-side comparisons with a reference radiometer. The calibration should last at least 1-2 months (Geuder et al., 2014; Al-Rasheedi et al., 2018) to include different atmospheric conditions, and photodiodes should be recalibrated every 2 years to mitigate the effects of sensitivity drifts (Sengupta et al., 2017). Previous studies have reported that the uncertainty of uncorrected photodiodes oscillates within 5-10% for 1-min values, doubling that of corrected ones that are generally below below 5% (Al-Rasheedi et al., 2018). Specifically, Al-Rasheedi et al. (2018) obtained 91% of G records within $\pm$5%, Vuilleumier et al. (2017) reported an uncertainty of $\pm$10% for presumably uncorrected photodiodes installed at Payerne, and Wilbert et al. (2015) found that the uncertainty was reduced from 5.2% to 2.2% using empirical corrections (Wilbert et al., 2015). All these values refer to 1-min records of G, which hinders the comparison with our study where only daily and annual uncertainties are reported. Besides, we did not carry a side-by-side validation against a reference instrument, which increases the uncertainty in our results. Nonetheless, the results obtained suggest that annual and daily uncertainties of SIAR photodiodes are larger than those reported in previous studies. This might be explained by the presence of undetected operational errors in the quality controlled data, which is likely due to the questionable maintenance of SIAR stations. Besides, SIAR photodidoes were probably uncorrected, and even though corrected, it is likely to believe that general empirical corrections were applied to most of the sensors. This would explain the abrupt inter-annual changes observed in the measurements from SIAR photodiodes. Overall, our results suggest that the poor maintenance and inadequate calibration of SIAR photodiodes reduces substantially the accuracy of the measurements obtained.

4.4. Comparison of the different sources of uncertainty

The uncertainty associated to radiation databases, operational errors and equipment errors (Fig. 8) was calculated from the annual differences shown in previous subsections by removing the uncertainty of the reference irradiance, i.e, the uncertainty of secondary standards in the analysis of radiation products and photodiodes, and the uncertainty of SARAH-1 in the analysis of operational failures. The expected uncertainty of secondary standards was used ($\pm$1.5%) (Sengupta et al., 2017), while the annual uncertainty SARAH-1 was estimated based on the validation against AEMET pyranometers ($+1.4^{+5.6}_{-5.3}$%). The uncertainty of daily values was also reported, which is the highest temporal resolution in which the different radiation sources can be compared. The daily bias is equal to the annual bias because this statistic is independent of the temporal resolution of the data. The daily uncertainty of SARAH-1 increased up to $^{+14.7}_{-14.9}$%, whereas the expected daily uncertainty for secondary standards is $\pm$2%. Note that the field uncertainty of secondary standards could be somewhat larger due to the presence of operational failures or an inadequate calibration. However, all secondary standards used as reference are ventilated, have passed the two QC tests, and are strictly maintained by AEMET. Besides, previous studies conducted by Vuilleumier et al. (2014) and Reda (2012) reported even

lower field uncertainties of $\pm$1.8% and $\pm$2.6%, respectively, for 1-min records obtained with properly calibrated and maintained secondary standard sensors. Therefore, the uncertainty values assumed for secondary standard seems reasonable. Finally, the variations between the differences discussed in previous sections and the uncertainties calculated here were almost negligible, corroborating the adequacy of the reference value selected for each analysis.

The largest deviations were found in operational errors, with biases around 10% for most of the defects and uncertainties within 10-20%. Despite the high magnitude of these values, these defects can rarely be detected with conventional QC methods and they are commonly present in ground datasets, especially in those from secondary networks. From user's perspective, the impact of these defects should be smaller because we have only used the years in which an operational error occurred. The second largest source of uncertainty was the use of estimations from radiation products. Here great differences were observed between reanalyses such as ERA-Interim and COSMO-REA6 (bias = +6.1 and -8.2%), and satellite-based products such as SARAH-1 and CLARA-A1 (bias = +1.4 and -1.6%). The gap between reanalysis and satellite-based products increased even more for daily estimations, because the bias variability observed in reanalysis datasets was compensated when aggregated to annual values. These systematic errors in reanalysis are caused by failures in the modeling of clouds (ERA-Interim, ERA5) and aerosols (COSMO-REA6). The daily uncertainty also increased more in CLARA-A1 ($^{+18.4}_{-20.5}$%) than in SARAH-1 ($^{+14.7}_{-14.9}$%), probably caused by the lower number of daily images available from polar satellites to estimate the cloud coverage. We confirm that satellite-based products using images from geostationary satellites, such as SARAH-1, provide the estimations with the lowest bias and uncertainty, and with the most homogeneous spatial performance (Polo et al., 2016). Reanalysis data should be only used as gross estimates when satellite-based data is missing.

Significant differences were also found between different types of pyranometer. The annual uncertainty of photodiodes ($^{+6.9}_{-5.4}$%) was substantially larger than that of secondary standards ($\pm$1.5%). The uncertainty in photodiodes would be even higher from the user's perspective, because operational defects and 60 doubtful stations were removed from this analysis whereas photodiodes are typically installed in secondary networks where the amount of operational errors is typically high. The daily ($^{+15.1}_{-16.1}$%) and annual ($^{+6.9}_{-5.4}$%) uncertainties obtained were larger than the values reported in previous studies made with photodiodes. These results should be interpreted cautiously because part of this uncertainty could be attributed to the validation procedure used. However, the presence of undetected operational errors and the inadequate calibration of these sensors are the most likely causes of the high uncertainties obtained. Besides, the results suggest that poorly maintained and incorrectly calibrated photodiodes could obtain similar uncertainties than the best satellite-based radiation databases such as SARAH-1. Therefore, SIAR photodiodes should avoided in applications in which high-quality data is essential, such as the validation of satellite-based products or the analysis of climate trends.

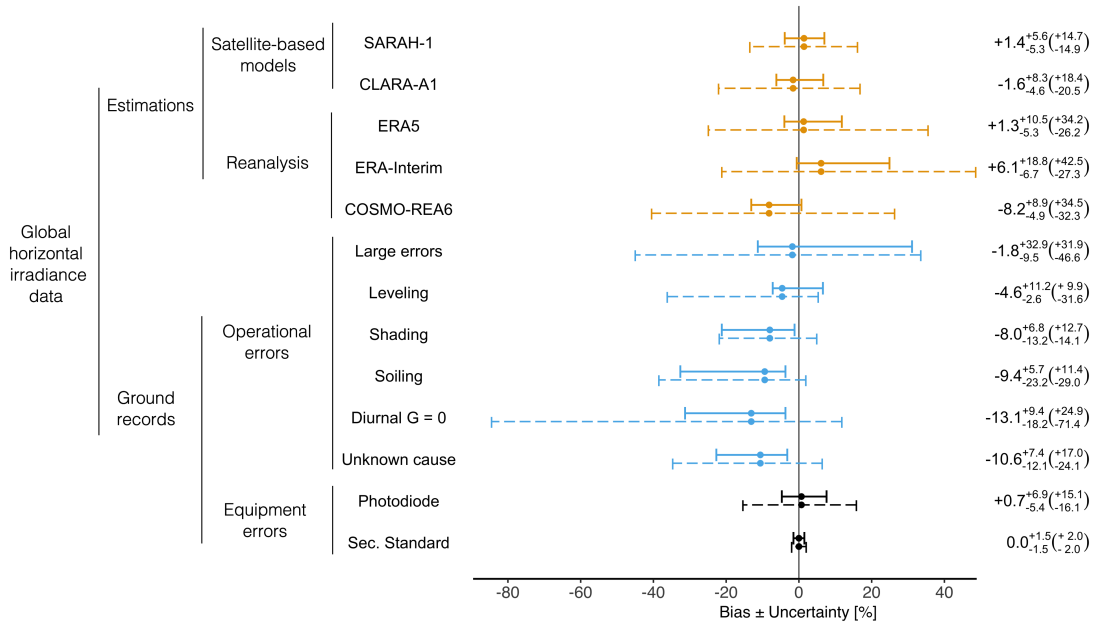

Fig. 8: Bias and uncertainty (95% CI) of the different sources of global horizontal radiation data evaluated. Annual and daily values are depicted with solid and dashed lines, respectively. Uncertainties for daily values are shown between parentheses.

5. Conclusions

In this study we evaluated the principal sources of uncertainty in the assessment of annual global horizontal irradiance. For estimated data, large differences were observed between reanalysis and satellite-based products, and also between different geographic conditions. Satellite-based products showed the lowest uncertainty, especially those based on images from geostationary satellites such as SARAH-1 ($+1.4^{+5.6}_{-5.3}$%). Therefore, they should be generally preferred to reanalyses if available. When using measurements from monitoring stations, operational errors produced the largest uncertainties among all the factors analyzed with annual biases of -8.0% and -9.4% for common errors such as shading and soiling, respectively. They primarily occurred in secondary networks with low maintenance of the stations. Importantly, they generally go unnoticed due to passing most of the conventional QC methods. The uncertainty was also significantly different between thermoelectric and photoelectric pyranometers, with the annual uncertainty of SIAR photodiodes ($^{+6.9}_{-5.4}$%) being far above than that of secondary standard pyranometers (±1.5%) and similar to that of SARAH-1. The high uncertainty of photodiodes might be explained by the presence of undetected operational errors and the lack of corrections for cosine errors, spectral response and temperature dependence. Part of this uncertainty might be also attributed to the validation procedure used. We recommend looking for records from secondary standards when the accuracy of the results is a critical factor. Photodiodes should be especially avoided if they belong to low-quality monitoring networks because the probability of having operational errors and uncorrected sensors substantially increases.

Acknowledgements

R. Urraca is funded by the fellowship FPI-UR-2014 granted by the University of La Rioja. A.S.G. is funded by Academy of Finland under the FINSKIN Project (num. 273689). A.S.G. also wants to express his gratitude to the Instituto de Estudios Riojanos (IER). This work used the Beronia cluster (Universidad de La Rioja), which is supported by FEDER-MINECO grant number UNLR-094E-2C-225. We would like to thank the different networks for providing their data free of charge and we specially appreciate the collaboration of Alberto Berastegui Arbizu from Tragsa, Natalia Murugarren from INTIA and Pedro Gavilán from IFAPA. We would also like to acknowledge the CM SAF project, the ECMWF and the HErZ/DWD for producing the radiation products used in this study. We would like to thank Ana Gracia-Amillo, Wim Zaaiman and Harald Müllejans for their valuable comments and suggestions.

References

AEMET, 2017. Agencia Estatal de Meteorología. URL: http://eportal.magrama.gob.es/websiar/Inicio.aspx. [accessed 05.02.17].

Al-Rasheedi, M., Gueymard, C., Ismail, A., Hussain, T., 2018. Comparison of two sensor technologies for solar irradiance measurement in a desert environment. Solar Energy 161, 194–2016. URL: https://doi.org/10.1016/j.solener.2017.12.058.

Boilley, A., Wald, L., 2015. Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface. Renewable Energy 75, 135–43. URL: http://dx.doi.org/10.1016/j.renene.2014.09.042.

Bojanowski, J.S., Vrieling, A., Skidmore, A.K., 2014. A comparison of data sources for creating a long-term time series of daily gridded solar radiation for Europe. Solar Energy 99, 152–71. URL: http://dx.doi.org/10.1016/j.solener.2013.11.007.

Bollmeyer, C., Keller, J., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S., Steinke, S., 2015. Towards a high-resolution regional reanalysis for the European CORDEX system. Quarterly Journal of the Royal Meteorological Society 141, 1–15. URL: https://dx.doi.org/10.1002/qj.2486.

BSRN, 2017. Baseline Surface Radiation Network. URL: http://bsrn.awi.de. [accessed 15.08.17].

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137, 553–97. URL: http://dx.doi.org/10.1002/qj.828.

Diresse, A., Zaaiman, W., Riley, D., Taylor, N., Stein, J.S., 2016. Investigation of pyranometer and photodiode calibrations under different conditions, in: Photovoltaic Specialist Conference (PVSC), 5-10 June 2016, Portland, USA. URL: http://dx.doi.org/10.1109/PVSC.2016.7749562, doi:10.1109/PVSC.2016.7749562.

Driesse, A., Zaaiman, W., Riley, D., Taylor, N., Stein, J.S., 2015. Indoor and outdoor evaluation of global irradiance sensors, in: 31st European Photovoltaic Solar Energy Conference, 14-18 September 2015, Hamburg, Germany, pp. 1704–1709. URL: http://dx.doi.org/10.4229/EUPVSEC20152015-5CO.5.3, doi:10.4229/EUPVSEC20152015-5CO.5.3.

ECMWF, 2017. ERA5 data documentation. European Centre for Medium-range Weather Forecast (ECMWF). URL: https://software.ecmwf.int/wiki/display/CKB/ERA5+data+documentation.

Euskalmet, 2017. Agencia vasca de meteorología. URL: http://www.euskalmet.euskadi.eus. [accessed 07.01.17].

Farrance, I., Badrick, T., Sikaris, K.A., 2016. Uncertainty in measurement and total error - are they so incompatible. Clin Chem Lab Med 54, 1309–1311. URL: http://dx.doi.org/10.1515/cclm-2016-0314.

Farrance, I., Frenkel, R., 2012. Uncertainty of measurement: A review of the rules for calculating uncertainty components through functional relationships. Clinical Biochemist Reviews 33, 49–75. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3387884/.

Frank, C., Wahl, S., Keller, J., Pospichal, B., Hense, A., Crewell, S., 2018. Bias correction of a novel european reanalysis data set for solar energy applications. Solar Energy 164, 12–24. URL: https://doi.org/10.1016/j.solener.2018.02.012.

Geuder, N., Affolter, R., Krass, B., Wilbert, S., 2014. Long-term behavior, accuracy and drift of LI-200 pyranometers as radiation sensors in Rotating Shadowband Irradiometers (RSI). Energy Procedia 49, 2330–2339. URL: http://dx.doi.org/10.1016/j.egypro.2014.03.247.

Gracia, A.M., Huld, T., 2013. Performance comparison of different models for the estimation of global irradiance on inclined surfaces. Validation of the model implemented in PVGIS. JRC Technical Report. URL: https://publications.europa.eu/en/publication-detail/-/publication/4ef8c4e1-4397-4e27-8487-448786327f27.

Gracia Amillo, A., Huld, T., Müller, R., 2014. A new database of global and direct solar radiation using the Eastern Meteosat satellite, models and validation. Remote Sensing 6, 8165–8189. URL: http;//dx.doi.org/10.3390/rs6098165.

Gueymard, C., 2014. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. Renewable and Sustainable Energy Reviews 39, 1024–1034. URL: http://dx.doi.org/10.1016/j.rser.2014.07.117.

Gueymard, C.A., Ruiz-Arias, J.A., 2016. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Solar Energy 128, 1–30. URL: https://doi.org/10.1016/j.solener.2015.10.010.

Habte, A., Wilcox, S., Stoffel, T., 2015. Evaluation of Radiometers Deployed at the National Renewable Energy Laboratory's Solar Radiation Research Laboratory. NREL technical report, NREL/TP-5D00-60896. URL: https://www.nrel.gov/docs/fy14osti/60896.pdf.

Hazuba, M., Folini, D., Sanchez-Lorenzo, A., M., W., 2013. Spatial representativeness of ground-based solar radiation measurements. Journal of Geophysical Research 88, 8585–8597. URL: http://dx.doi.org/10.1002/jgrd.50673,2013.

Ineichen, P., 2011. Global irradiance on tilted and oriented planes: model validations. University of Geneva. URL: https://archive-ouverte.unige.ch/unige:23519.

Ineichen, P., 2014. Long term satellite global, beam and diffuse irradiance validation. Energy Procedia 48, 1586–1596. URL: http://dx.doi.org/10.1016/j.egypro.2014.02.179.

ISO, 1990. ISO 9060:1990: Specification and classification of instruments for measuring hemispherical solar and direct solar radiation. Geneva, Switzerland.

JCGM 100:2008, 2008. Evaluation of measurement data - guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology (JCGM).

JCGM 104:2009, 2009. Evaluation of measurement data - An introduction to the "Guide to the expression of uncertainty in measurement" and related documents. Joint Committee for Guides in Metrology (JCGM).

Journée, M., Bertrand, C., 2011. Quality control of solar radiation data within the RMIB solar measurements network. Solar Energy 85, 72–86. URL: http://dx.doi.org/10.1016/j.solener.2010.10.021.

Karlsson, K., Riihelä, A., Müller, R., Meirink, J., Sedlar, J., Stengel, M., Lockhoff, M., Trentmann, J., Kaspar, F., Hollmann, R., Wolters, E., 2012. Clara-a1: Cm saf clouds, albedo and radiation dataset from avhrr data - edition 1 - monthly means / daily means / pentad means / monthly histograms URL: http://dx.doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V001.

Long, C.N., Dutton, E.G., 2002. BSRN Global Network recommended QC tests, V2.0. BSRN Technical Report URL: http://ezksun3.ethz.ch/bsrn/admin/dokus/qualitycheck.pdf.

McArthur, L.J.B., 2005. World Climate Research Programme - Baseline Surface Radiation Network (BSRN) - Operations Manual Version 2.1. Downsview, Ontario, CANADA, Experimental Studies Division, Atmospheric Environment Service.

Meteo Navarra, 2017. Web de meteorología y climatología de Navarra. URL: http://meteo.navarra.es. [accessed 07.01.17].

Meteocat, 2017. Servicio Meteorológico de Cataluña. URL: http://www.meteo.cat. [accessed 07.01.17].

MeteoGalicia, 2017. Unidad de Observación y Predición Meteorológica de Galicia. URL: http://www.meteogalicia.gal. [accessed 07.01.17].

Meyers, D.R., 2011. Quantitative Analysis of Spectral Impacts on Silicon Photodiode Radiometers. NREL technical report, NREL/CP-5500-50936. URL: https://www.nrel.gov/docs/fy11osti/50936.pdf.

Moretón, R., Lorenzo, E., Pinto, A., Muñoz, J., Narvarte, L., 2017. Renewable and Sustainable Energy Reviews 78, 886–903. URL: https://dx.doi.org/10.1016/j.rser.2017.05.020.

Müller, B., Bostock, P., Farnung, B., Reise, C., 2017. A framework to calculate uncertainties for lifetime energy yield predictions of PV systems, in: IEEE PV Specialist Conference (PVSC), 25-30 June 2017, Washington, USA, pp. 1–5.

Müller, B., Wild, M., Driesse, A., Behrens, K., 2014. Rethinking solar resource assessments in the context of global dimming and brightening. Solar Energy 99, 272–282. URL: http://dx.doi.org/10.1016/j.solener.2013.11.013.

Müller, R., Pfeifroth, U., Träger-Chatterjee, C., 2015a. Towards optimal aerosol information for the retrieval of solar surface radiation using Heliosat. Atmosphere 6, 863–878. URL: http://www.mdpi.com/2073-4433/6/7/863.

Müller, R., Pfeifroth, U., Träger-Chatterjee, C., Cremer, R., Trentmann, J., Hollmann, R., 2015b. Surface solar radiation data set - Heliosat (SARAH) - Edition 1. Satellite Application Facility on Climate Monitoring (CM SAF). URL: http://dx.doi.org/10.5676/EUM_SAF_CM/SARAH/V001, doi:10.5676/EUM_SAF_CM/SARAH/V001.

Müller, R., Pfeifroth, U., Träger-Chatterjee, C., Trentmann, J., Cremer, R., 2015c. Digging the METEOSAT treasure - 3 decades of solar surface radiation. Remote Sensing 7, 8067–101. URL: http://dx.doi.org/10.3390/rs70608067.

Pfeifroth, U., Kothe, S., Trentmann, J., 2016. Meteosat solar surface radiation and effective cloud Albedo climate data record, SARAH-2. Validation report. CM SAF, EUMETSAT Satellite Application Facility on Climate Monitoring.

Polo, J., Wilbert, S., Ruiz-Arias, J.A., Meyer, R., Gueymard, C., Súri, M., Martín, L., Mieslinger, T., Blanc, P., Grant, I., Boland, J., Ineichen, P., Remund, J., Escobar, R., Troccoli, A., Sengupta, M., Nielsen, K.P., Renne, D., Geuder, N., Cebecauer, T., 2016. Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Solar Energy 132, 25–37. URL: http://dx.doi.org/10.1016/j.solener.2016.03.001.

Reda, I., 2012. Method to Calculate Uncertainties in Measuring Shortwave Solar Irradiance Using Thermopile and Semiconductor Solar Radiometers. NREL technical report, NREL/TP-3B10-52194. URL: https://www.nrel.gov/docs/fy11osti/52194.pdf.

Riihelä, A., Carlund, T., Trentmann, J., Müller, R., Lindfors, A.V., 2015. Validation of CM SAF surface solar radiation datasets over Finland and Sweden. Remote Sensing 7, 6663–82. URL: http://dx.doi.org/10.3390/rs70606663.

Sanchez-Lorenzo, A., Enriquez-Alonso, A., Wild, M., Trentmann, J., Vicente-Serrano, S., Sanchez-Romero, A., Posselt, R., Hakuba, M., 2017. Trends in downward surface solar radiation from satellite and ground observations over europe 1983-2010. Remote Sensing of Environment 189, 108–117. URL: https://doi.org/10.1016/j.rse.2016.11.018".

Sengupta, M., Gotseff, P., Stoffel, T., 2012. Evaluation of Photodiode and Thermopile Pyranometers for Photovoltaic Applications. NREL technical report, NREL/CP-5500-56540. URL: http://www.nrel.gov/docs/fy15osti/63112.pdf.

Sengupta, M., Habte, A., Gueymard, C., Wilbert, S., Renné, D., 2017. Best practices handbook for the collection and use of solar resource data for solar energy applications: Second edition. NREL technical report, NREL/TP-5D00-68886. URL: https://www.nrel.gov/docs/fy18osti/68886.pdf.

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., Stoffel, T., Renné, D., Gueymard, C., Myers, D., Wilcox, S., Blanc, P., Perez, R., 2015. Best practices handbook for the collection and use of solar resource data for solar energy applications. NREL technical report. URL: http://www.nrel.gov/docs/fy15osti/63112.pdf.

SIAR, 2015. Servicio de Información Agroclimática para el Regadío. URL: http://eportal.magrama.gob.es/websiar/Inicio.aspx. [accessed 10.10.16].

SIAR La Rioja, 2017. Servicio de Información Agroclimático de La Rioja. URL: http://www.larioja.org/agricultura/es/informacion-agroclimatica/datos-estaciones. [accessed 07.01.17].

SOS Rioja, 2017. Estaciones meteorológicas del Gobierno de La Rioja. URL: http://www.larioja.org/emergencias-112/es/meteorologia. [accessed 07.01.17].

Suri, M., Cebecauer, T., 2014. Satellite-based solar resource data: model validation statistics versus user's uncertainty, in: ASES SOLAR 2014 Conference, San Francisco, 7-9 July 2014, pp. 7–9.

Thevenard, D., Pelland, S., 2013. Estimating the uncertainty in long-term photovoltaic yield predictions. Solar Energy 91, 432–445. URL: http://dx.doi.org/10.1016/j.solener.2011.05.006.

Träger-Chatterjee, C., Müller, R.W., Trentmann, J., Bendix, J., 2010. Evaluation of ERA-40 and ERA-interim re-analysis incoming surface shortwave radiation datasets with mesoscale remote sensing data. Meteorologische Zeitschrift 19, 631–640. URL: https://dx.doi.org/10.1127/0941-2948/2010/0466.

Trentmann, J., Müller, R., 2012. CM SAF Cloud, Albedo, Radiation dataset, AVHRR-based, Edition 1 (CLARA-A1), Surface Radiation Products, Validation Report. CM SAF. URL: http://dx.doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V001, doi:10.5676/EUM_SAF_CM/CLARA_AVHRR/V001.

Urraca, R., Gracia Amillo, A., Huld, T., Martinez-de Pison, F.J., Trentmann, J., Lindfors, A., Riihelä, A., Sanz Garcia, A., 2017a. Quality control of solar radiation data with satellite-based products. Solar Energy 158, 49–62. URL: http://dx.doi.org/10.1016/j.solener.2017.09.032.

Urraca, R., Gracia-Amillo, A., Koubli, E., Huld, T., Trentmann, J., Riihelä, A., Lindfors, A., Palmer, D., Gottschalg, R., Antonanzas-Torres, F., 2017b. Extensive validation of CM SAF surface radiation products over Europe. Remote Sensing of Environment 199, 171–186. URL: https://dx.doi.org/10.1016/j.rse.2017.07.013.

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de Pison, F.J., Kaspar, F., A., S.G., 2018. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. Solar Energy 164, 339–354. URL: https://dx.doir.org/10.1016/j.solener.2018.02.059.

Urraca, R., Martinez-de Pison, E., Sanz-Garcia, A., Antonanzas, J., Antonanzas-Torres, F., 2017c. Estimation methods for global solar radiation: Case study evaluation of different approaches in central Spain. Renewable & Sustainable Energy Reviews 77, 1098–1113. URL: http://dx.doi.org/10.1016/j.rser.2016.11.222.

Vignola, F., Michalsky, J., Stoffel, T., 2012. Solar and infrared radiation measurements. Taylor & Francis.

Vuilleumier, L., Félix, C., Vignola, F., Blanc, P., Badosa, J., Kazantzidis, A., Calpini, B., 2017. Performance evaluation of radiation sensors for the solar energy sector. Meteorologische Zeitschrift 26, 485–505. URL: https://dx.doi.org/10.1127/metz/2017/0836.

Vuilleumier, L., Hauser, M., Félix, C., Vignola, F., Blanc, P., Kazantzidis, A., Calpini, B., 2014. Accuracy of ground surface broadband shortwave radiation monitoring. Journal of Geophysical Research: Atmosphere , 838–860URL: http://dx.doi.org/10.1002/2014JD022335.

Wilbert, S., Kleindiek, S., Nouri, B., Geuder, N., Habte, A., Schwandt, M., Vignola, F., 2015. Uncertainty of rotating shadowband irradiometers and Si-pyranometers including the spectral irradiance error. AIP Conference Proceedings 1734, 150009. URL: http://dx.doi.org/10.1063/1.4949241.

Wild, M., 2009. Global dimming and brightening: A review. Journal of Geophysical Research 11, 1–31. URL: httpx://doi.org/10.1029/2008JD011470.

WMO, 2008. WMO guide to meteorological instruments and methods of observation: WMO No. 8. 7th edition ed., Geneva, Switzerland. URL: http://www.wmo.int/pages/prog/gcos/documents/gruanmanuals/CIMO/CIMO_Guide-7th_Edition-2008.pdf.

Younes, S., Claywell, R., Muneer, T., 2005. Quality control of solar radiation data: Present status and proposed new approaches. Energy 30, 1533–1549. URL: http://dx.doi.org/10.1016/j.energy.2004.04.031.

Zahumenský, I., 2004. Guidelines on Quality Control Procedures for Data from Automatic Weather Stations. World Meteorological Organization. Geneva, Switzerland. URL: https://www.wmo.int/pages/prog/www/OSY/Meetings/ET-AWS3/Doc4(1).pdf.