Quality Assurance Procedure for Solar Radiation at Minute Resolution

Diego Miranda[1], Leonardo Petribú[2], Janis Galdino[2], João Victor Furtado[2], Lucas Barboza[2], Olga Vilela[2], Alexandre Costa[2], Emerson Gomes[2], Alex Pereira[3], Eduardo Jatobá[3], Alcides Codeceira Neto[3] and José Bione Filho[3]

[1] Former affiliation: Center for Renewable Energy of the Federal University of Pernambuco (CER-UFPE), Recife (Brazil), now at German Aerospace Center (DLR), Institute of Networked Energy Systems, Oldenburg (Germany)

[2] Center for Renewable Energy of the Federal University of Pernambuco (CER-UFPE), Recife (Brazil)

[3] Companhia Hidro Elétrica do São Francisco (CHESF), Recife (Brazil)

Abstract

This article proposes a quality assurance (QA) procedure for evaluating measurements of global horizontal, diffuse horizontal and direct normal irradiances. The QA is divided into physical, comparative and refining tests, including Baseline Surface Radiation Network (BSRN) recommended quality tests. The comparative tests (tracker off and coherence tests) evaluate the connection among the three main radiation components. The coherence test as proposed by BSRN is complemented with a comparison between the measured radiation series with a validated clear sky model to define which of the three radiation components is presenting problems. Three stations in Pernambuco, Brazil and one in Gobabeb, Namib have been qualified using the proposed QA, with two of these stations belonging to BSRN (Petrolina and Gobabeb). The results show that the direct normal irradiance series can present considerable outliers due to soiling in pyrheliometers. For solar stations without periodic maintenance, the series of direct normal radiation measured by pyrheliometers can have more soiling problems than other radiation components measured by pyranometers.

Keywords: quality assurance, soiling in pyrheliometer

1. Introduction

Measuring solar radiation with accuracy is vital to analyse radiation models (for example, separation models, transposition models and power photovoltaic output models), apply site adaptation techniques, evaluate the accuracy of solar maps, and many other applications in solar energy. Thus, quality assurance (QA) of solar radiation data is an important tool to increase the reliability of the measured data. Many other works present QA solar radiation data procedures using different tests and techniques (NREL, 1993; Long and Dutton, 2002; Urraca et al., 2017; Perez-Astudillo et al., 2018). This work presents an approach for solar radiation quality assurance based on the Baseline Surface Radiation Network (BSRN) recommended quality tests. The quality assurance procedure focuses on the coherence test with a clear sky model validation and the tracker off test to increase the comparative tests analysis of the three radiation components.

2. Methodology

The QA for solar radiation evaluates time series of global horizontal (G), diffuse horizontal (G_d), and direct normal (G_b) irradiances in W/m², divided into three sets of tests: physical tests, comparative tests, and refining tests, as shown in Fig. 1 and Tab. 1 (Petribú et al., 2017). The global checks identify inconsistencies in the record of timestamps by the datalogger. First, the chronological record of date and hour is checked, ordering eventual timestamps and their respective measurements. Identical registered timestamps are checked as well. The gaps between consecutive records are filled with 'Not a Number' (NaN). This work uses six quality indicators, called flags, to classify the data. If one data is classified as 'good' (flag 1), it means the data was approved in all quality

doi:10.18086/eurosun.2022.15.04 Available at http://proceedings.ises.org

tests of the procedure. One data is classified as 'suspicious' (flag 2) when the test is not conclusive if the data was a rare occurrence or anomalous data. If one data is classified as 'anomalous' (flag 3) in one test, it will not be checked on the subsequent tests, being classified as 'previously anomalous' (flag 4). Finally, flag 5 is for non-tested data, and flag 6 is for unavailable data (NaN values). The tests are applied in the numerical sequence as shown in Tab. 1. The physical tests use equations based on solar geometry to evaluate the quality of irradiance data. The first test proposes limits to the measured irradiances, while the third test proposes limits to the clearness index ($G/G_{eff,h}$ for k_t), diffuse ratio (G_d/G for k_d), and normal transmittance (G_b/G_{eff} for k_n), called here k's limits, with the threshold values adapted from Raichijk (2012). G_{eff} is the extraterrestrial radiation with a mean value of 1361.1 W/m², according to Gueymard et al. (2018), varying throughout the year due to the sun-earth distance. $G_{eff,h}$ is the extraterrestrial radiation in a horizontal plane which is obtained multiplying G_{eff} by the cosine of solar zenith angle (Θ_z). The second test indicates that the procedure does not test the data above 7º of solar elevation (α), adapted from Maxwell (1973). The fourth test (Long and Dutton, 2002) defines lower and upper limits for irradiance measurements, while the fifth test is based on Rigollier et al. (2000) clear sky model using a link turbidity factor equal to 2.5, according to Ruiz-Arias et al. (2010). The comparative tests include the 'tracker off' to evaluate the solar tracker's operation, and the coherence test (BSRN), to judge the coherence among the three radiation components. It is important to note that the coherence test is analysed on clear days with a clear sky model validated by the QA procedure – here, the McClear model (Lefèvre et al., 2019) was chosen. Then, if the clear sky model reproduces the behavior of the global horizontal irradiance and direct normal irradiance on clear days, this model can be used to decide which of the three components is with measurement problems. Finally, the refining tests aim to evaluate abrupt solar radiation variations between consecutive data on minute resolution (step test) and the repetition of the same value of radiation consecutively in large windows (persistence test), as proposed by Sönmez (2013). In Tab. 1, data is classified as anomalous (flag 3) when the conditions are not satisfied for tests 1, 3, 4i and 8; suspicious (flag 2) when the conditions are not satisfied for tests 4ii and 5; and anomalous (flag 3) when the conditions are satisfied for the tracker off and persistence tests.

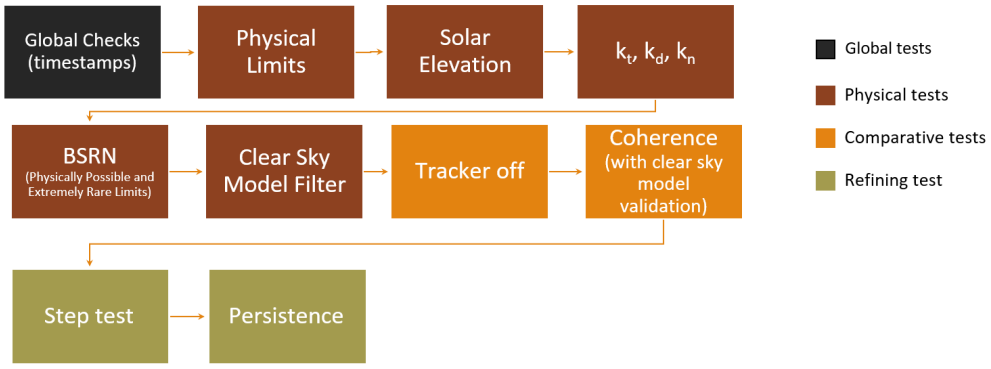


Fig. 1: Methodology of QA procedure.

Tab. 1: Tests proposed by the Quality Assurance procedure.

Physical Tests	
1 – Limits Test	$-5 < G, G_b$ and $G_d < 2000$ for irradiances in W/m²
2 – Solar Elevation	$\alpha < 7^o$ (non-tested data – flag 5); where α is the solar elevation (in degrees)
3 – k's Limits	$0 < k_t \leq 1.3; 0 < k_d \leq 1.1; 0 \leq k_n \leq 1.1$
4 – Extremely rare limits (i) and physically possible limits (ii) (BSRN)	i) Anomalous data: $-4 \leq G \leq G_{eff} 1.5 \cdot cos(\Theta_z)^{1,2} + 100; -4 \leq G_b \leq G_{eff}; -4 \leq G_d \leq G_{eff}\ 0.95 \cdot cos(\Theta_z)^{1,2} + 50$ for irradiances in W/m²
	ii) Suspicious data: $-2 \leq G \leq G_{eff} 1.2 \cdot cos(\Theta_z)^{1,2} + 50; -2 \leq G_b \leq G_{eff} 0.95 \cdot cos(\Theta_z)^{0,2} + 10; -2 \leq G_d \leq G_{eff} 0.75 \cdot cos(\Theta_z)^{1,2} + 30$ for irradiances in W/m²
5 – Clear Sky Filter Test	$G < 2{,}1 \cdot G_{csRig}; G_b < 2{,}1 \cdot G_{b,csRig}; G_d > 0{,}8 \cdot G_{d,csRig}$, where $G_{csRig}, G_{b,csRig}$ and $G_{d,csRig}$ are the Rigollier clear sky model for G, G_b and G_d

Comparative Tests	
6 – Tracker Off Test (anomalous data if the conditions are satisfied)	i) Daily Basis: $K_d > 0.75$ & $K_c > 0.7$ & $G_{b,daily} < 25\ W/m^2; K_c = \frac{G_{daily}}{G_{daily,McClear}}$ where K_d and K_c are the daily diffuse fraction and daily clear sky index, respectively; $G_{b,daily}$ and G_{daily} are the daily values for G_b and G; and $G_{dailyMcClear}$ is the daily G clear sky irradiance for McClear model.
	ii) Minutely Basis: $k_d > 0.75$ & $k_c > 0.85$ & $G_b < 5\ W/m^2$ & $\theta_z < 75°;\ k_c = \frac{G}{G_{cs,McClear}}$ where $G_{cs,McClear}$ is the McClear clear sky model for G (minute resolution).
7 – Coherence Test (BSRN). The results of this test should be compared with a validated clear sky model.	i) $\left\|\frac{G_d+G_b cos(\Theta_Z)-G}{G}\right\| \leq 0{,}08\ if\ \Theta_Z < 75°$ and $G_d + G_b\ cos\ (\Theta_Z) > 50\ W/m^2$ ii) $\left\|\frac{G_d+G_b cos(\Theta_Z)-G}{G}\right\| \leq 0{,}15\ if\ 75° < \Theta_Z < 90°$ and $G_d + G_b\ cos\ (\Theta_Z) > 50\ W/m^2$ iii) $k_d < 1{,}05\ for\ \Theta_Z < 75°\ and\ \mathrm{G} > 50\ W/m^2$ iv) $k_d < 1{,}1\ for\ 75° < \Theta_Z < 90°\ and\ G > 50\ W/m^2$
Refining Tests	
8 – Step test	$(\mathrm{G}, G_b, G_d)_t - (\mathrm{G}, G_b, G_d)_{t-1} \leq 1100\ W/m^2$ where t is the current timestamp and $t-1$ the previous (minute resolution)
9 – Persistence test	i) G data with successive repetitions ≥ 20 min are considered anomalous ii) G_b non-zero data with successive repetitions ≥ 30 min are considered anomalous

The BSRN coherence test does not identify which of the three radiation components presents problems when the first two equations (Tests 7i and 7ii) detect a considerable difference between the sum $G_d + G_b cos(\Theta_Z)$ and G. To determine which component is not measuring well, we proposed a three-step procedure for running the coherence test, as presented in Fig. 2. First, the procedure identifies clear sky days on the radiation time series using the G measurements and the McClear model for global irradiance (G_{cs}). The metrics clear sky ratio (K_c) and variability index (VI) are calculated on a daily resolution from the data in minute resolution, as proposed by Stein et al. (2012) with adaptations. These metrics aim to identify clear sky days in the conditions specified on the first diagram block ($K_c > 0.8$ and VI lower than the fifth percentile of VI), because clear sky days have high K_c and low VI. The VI metric quantifies how much solar radiation varied in a specific time window. It is the ratio between the sum of lengths (or line segments) of G and the sum of lengths of G_{cs}. In Eq. 1, it was used n = 1440 to obtain the VI values on a daily resolution from the data in minute resolution. Clear sky and overcast days have low values of VI, while days with variable radiation have high values of VI. In the first step of the proposed method, it is necessary to choose a well-adjusted global horizontal irradiance clear sky model.

$$VI = \frac{\sum_{i=2}^{n}\sqrt{(G_i - G_{i-1})^2 + 1}}{\sum_{i=2}^{n}\sqrt{(G_{cs,i} - G_{cs,i-1})^2 + 1}} \quad \text{(eq.1)}$$

After selecting the clear sky days, the clear sky model must be validated for the coherence test. The validation is made by statistically comparing the measurements of the previously selected clear sky days and the clear sky model. First, in step 2, the days containing data detected by the first equation (Test 7i) of the coherence test are dropped. Then, only the data without incoherence among the radiation components remain in the dataset. G and G_b are validated according to the statistics presented in the two last diagrams of step 2. The clear sky model can be used in the coherence test only if both G_{cs} and $G_{b,cs}$ (McClear model for G_b) are validated. The $G_{d,cs}$ (McClear model for G_d) was not validated because the diffuse measurement in clear sky conditions varies depending on the location, making it challenging to define an excellent statistical range for its validation.

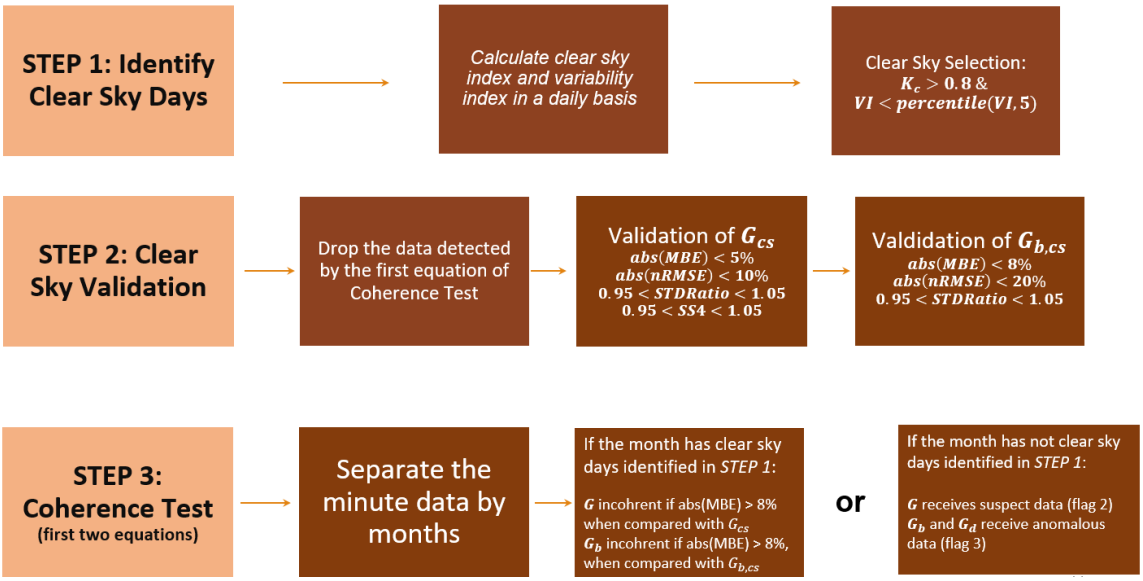


Fig. 2: Methodology of coherence test with clear sky validation.

The statistics used are the normalized mean bias error (nMBE), the normalized root mean square error (nRMSE), the Pearson correlation (*Corr*), the ratio between the standard deviation of the model and the standard deviation of the observation (STDRatio) and the skill score of Taylor (SS4). Eqs. 2 to 4 show the formulas for these statistics, where x_{mod} is the modeled time series, x_{obs} is the observed time series and $\bar{x}_{obs}$ is the mean of the observed time series.

$$nMBE = \frac{1}{N\,\bar{x}_{obs}} \sum_{i=1}^{N} (x_{mod}^{i} - x_{obs}^{i}) \quad \text{(eq. 2)}$$

$$nRMSE = \frac{1}{\bar{x}_{obs}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{mod}^{i} - x_{obs}^{i})^2} \quad \text{(eq. 3)}$$

$$SS4 = \frac{(1+Corr)^4}{4\left(STDRatio + \frac{1}{STDRatio}\right)^2} \quad \text{(eq. 4)}$$

Finally, step 3 applies the first two equations of the coherence test (Tests 7i and 7ii). It consists of applying the first two equations presented in Table 1 for the coherence test monthly. Then, the minute measurements of G, G_b and G_d are analysed month by month. For each month, the script searches if clear sky days from the previously selected clear sky days (step 1) exists in the month in question. If there were one or more clear sky days on this particular month, a comparison is made between the clear sky measurements and the clear sky model on these specific days for G and G_b. The idea here is to analyse if the coherence issue lies in components G or G_b, instead of flagging all components detected by Tests 7i and 7ii. The absolute value of nMBE statistic was chosen to compare the clear sky model and measurements on clear sky days because the bias shows if a model is overestimating or underestimating the observations on the mean. The threshold values adopted were a mean bias error of 8% for the absolute difference $G - G_{cs}$ and $G_b - G_{b,cs}$. The value of 8% was chosen after a sensitive analysis of the ground measurements data employed. For example, if the absolute value of nMBE is higher than 8% for the difference $G_b - G_{b,cs}$, all G_b detected by the first two equations of the coherence test (Tab. 1 – Tests 7i and 7ii) for that month are considered incoherent. The evaluation of each month follows the cases shown in Tab. 2. In the cases where no clear sky days from the previous selection are present in a particular month, the detected data by Tests 7i and 7ii might attribute suspicious data for G (flag 2) and anomalous data for G_b and G_d (flag 3). For example, in the third row of Tab. 2, if the G_b data detected on clear sky days were considered incoherent (which was decided based on the difference $G_b - G_{b,cs}$ higher than 8% of nMBE on the selected clear

sky days) and the G data were coherent for the particular month, then the G_b component is causing the differences detected by Tests 7i and 7ii on that month and receives flag 3, while G receives good data and G_d suspicious data. Finally, the last two equations of coherence tests (Tab. 1 – Tests 7iii and 7iv) always attribute anomalous data for G and G_d, independent of the clear sky validation.

Tab. 2: Analysis of G and G_b on clear sky days of a particular month.

If on the clear sky days of a particular month...	Then, the detected data by the first two equations of the coherence test (Tab. 1, Tests 7i and 7ii) for that month will be flagged as		
	G	G_b	G_d
G_b is **coherent** and G is **coherent**	Suspicious data (flag 2)	Anomalous data (flag 3)	Anomalous data (flag 3)
G_b is **coherent** and G is incoherent	Anomalous data (flag 3)	Good data (flag 1)	Suspicious data (flag 2)
G_b is incoherent and G is **coherent**	Good data (flag 1)	Anomalous data (flag 3)	Suspicious data (flag 2)
G_b is incoherent and G is incoherent	Anomalous data (flag 3)	Anomalous data (flag 3)	Anomalous data (flag 3)

The tracker off test aims to find issues in the operation of the solar tracker. In the QA procedure, this test is applied in two steps: a daily analysis to find days whose tracker did not work all day and an analysis on minute resolution to find the moments when the tracker is not working for part of the day. To perform the daily analysis, the daily mean of G, G_b, G_d and G_{cs} are extracted, accepting an amount until 25% of missing data (NaN values) to calculate the mean of the variables considering only sunshine time. If there are more than 25% of NaN values on a particular day during sunshine, this particular day receives NaN instead of the daily mean. After having the four variables daily, the tracker off condition on a daily basis is applied (Tab. 1 – Test 6i). With this simple test, we can detect days when the tracker did not work all day. Otherwise, in the second step, the conditions presented in Tab. 1 (Test 6ii) are used to detect moments when the tracker did not work for part of the day, as proposed by Long and Shi (2006), with some adaptations.

Besides the objective tests presented in this QA procedure, it is worth to be noted that the quality assurance of measuring data should always be complemented with a rigorous analysis of the results and time series plots. In some cases, the tests cannot detect all kinds of anomalies, and an analysis by an expert could be fundamental to complement the objective procedure. Therefore, the QA procedure's results are manually analysed to complement the flagged process of possible anomalous data that were not detected by the objective QA procedure.

3. Datasets and results

Three meteorological stations located in the state of Pernambuco (CRESP Petrolina, SONDA Petrolina and Araripina), Brazil, and one located in the Gobabeb Namib Research Institute, Namib, were analysed in this study. Their stations' coordinates and solar sensors are presented in Tab. 3. All ground measurements data are in minutely resolution. Tab. 4 shows the final results of the QA procedure for all stations. The tracker off and the coherence tests were the tests that mostly identified anomalous data. Then, the results of these two tests are presented in major details in this section. The results of the three steps of the coherence test will be presented for CRESP Petrolina station to illustrate the methodology described for the coherence test. First, in step 1, 45 clear sky days were selected from CRESP Petrolina station. In step 2, 11 days were dropped from the selected clear days because they had measurement problems detected by the first equation of coherence test (Tab. 1 – Test 7i). Fig. 3 shows the 34 remained clear sky days used to validate the clear sky model for G_{cs} and $G_{b,cs}$. The statistical comparison results are shown in Tab. 5 for all stations. All the statistics are inside the ranges proposed on the diagram methodology for stations CRESP Petrolina, Araripina and SONDA Petrolina, so the clear sky models G_{cs} and $G_{b,cs}$ could be used to analyse the results of the first two equations of the coherence test. For Gobabeb station, $G_{b,cs}$ was not validated on the site because the STDRatio exceeds the proposed range. Then, the first two equations of the coherence test will attribute suspicious data for G and anomalous data for G_b and G_d for the detected data of Gobabeb station.

Tab. 3: Meteorological stations used for applying the QA procedure.

Station	Location and country	Coordinates	Elevation (m)	Solar variables and sensors
CRESP	Petrolina, Brazil	-9.1069, -40.4414	376	G (EKO MS-80A), G_b (EKO MS-57) and G_d (EKO MS-80A)
SONDA Petrolina (BSRN - PTR)	Petrolina, Brazil	-9.0689, -40.3197	387	G (Kipp&Zonen CMP22), G_b (Kipp&Zonen CHP1) and G_d (Kipp&Zonen CMP22)
Araripina	Araripina, Brazil	-7.5742, -40.5144	633	G (Kipp&Zonen CMP3), G_b (Kipp&Zonen CHP1) and G_d (Kipp&Zonen CMP3)
Namib (BSRN - GOB)	Gobabeb, Namib	-23.5614, 15.04198	407	G (Kipp&Zonen CMP22), G_b (Kipp&Zonen CHP1) and G_d (Kipp&Zonen CMP22)

Tab. 4: Results of QA procedure for all stations analysed.

Quality Flag	1 - Good data	2 - Suspicious data	**3 - Anomalous data**	5 - Non-tested data	6 - Data not available
G Araripina	42.54%	2.10%	**2.13%**	48.33%	4.90%
G_d Araripina	35.42%	5.85%	**5.66%**	48.10%	4.97%
G_b Araripina	33.20%	0.00%	**12.21%**	49.65%	4.94%
G SONDA Petrolina	34.42%	1.29%	**0.54%**	42.42%	21.33%
G_d SONDA Petrolina	30.76%	1.08%	**4.72%**	42.09%	21.35%
G_b SONDA Petrolina	30.43%	0,00%	**6.39%**	41.82%	21.36%
$G1$ CRESP Petrolina	38.81%	0.14%	**0.01%**	46.24%	14.80%
$G2$ CRESP Petrolina	38.88%	0.08%	**0.00%**	46.24%	14.80%
G_d CRESP Petrolina	37.81%	0.97%	**0.18%**	46.24%	14.80%
G_b CRESP Petrolina	37.72%	0.00%	**1.30%**	46.19%	14.79%
G Namíbia	44.60%	0.03%	**0.01%**	53.70%	1.66%
G_d Namíbia	44.50%	0.08%	**0.05%**	53.70%	1.67%
G_b Namíbia	44.52%	0.00%	**0.05%**	53.69%	1.74%

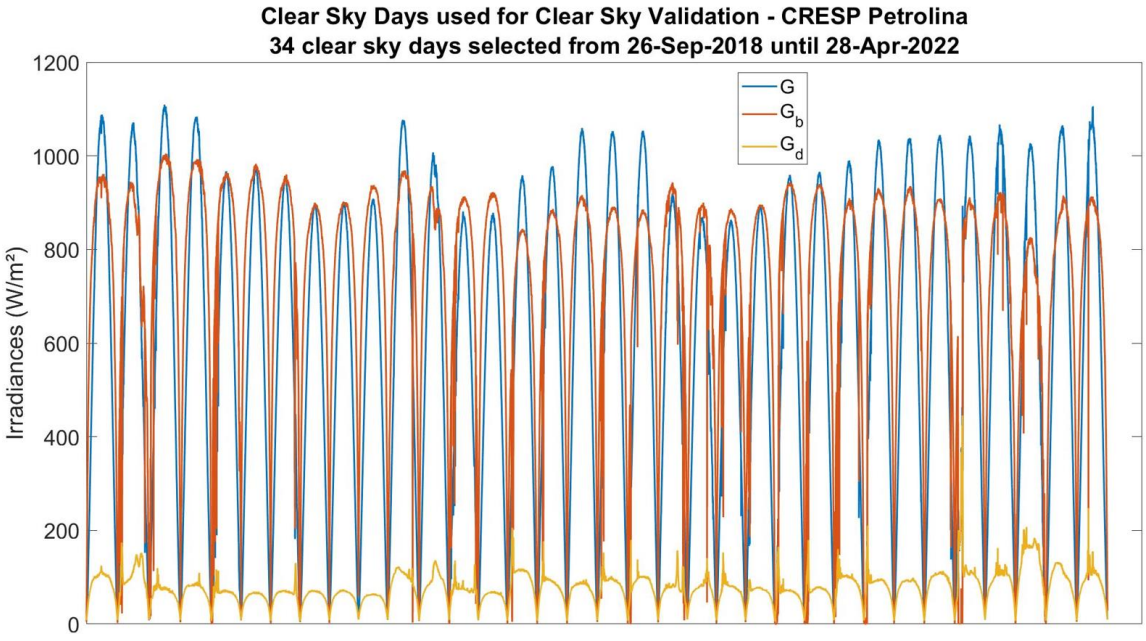


Fig. 3: Clear sky days used for the clear sky validation in CRESP Petrolina station.

Tab. 5: Results of clear sky validation.

CRESP Petrolina Station							
Clear Sky Model	BIAS (W/m²)	MBE (%)	STDRatio	RMSE (W/m²)	nRMSE (%)	Corr (.)	SS4 (.)
G_{cs} (McClear)	-13,37	-2,25%	1,00	32,35	5,45%	1,00	0,99
$G_{b,cs}$ (McClear)	-53,53	-7,33%	0,97	106,12	14,53%	0,93	0,86
SONDA Petrolina Station							
G_{cs} (McClear)	-21,38	-3,62%	1,03	40,94	6,93%	0,99	0,99
$G_{b,cs}$ (McClear)	-40,21	-5,68%	0,96	131,17	18,52%	0,87	0,77
Araripina Station							
G_{cs} (McClear)	-12,02	-1,92%	1,01	26,98	4,31%	1,00	0,99
$G_{b,cs}$ (McClear)	-50,59	-6,65%	0,99	85,29	11,22%	0,96	0,92
Gobabeb Station							
G_{cs} (McClear)	-9,99	-1,99%	1,03	16,33	3,26%	1,00	1,00
$G_{b,cs}$ (McClear)	33,28	4,39%	1,06	63,13	8,33%	0,98	0,95

Finally, in the third step, the 45 days detected in step 1 are separated according to the months where they occur. If the particular month has no clear days selected by step 1, the data detected by the first two equations receives flag 2 for G and flag 3 for G_b and G_d. Otherwhise, when the particular month has clear sky days detected by step 1, the nMBE is checked to decide if the G or G_b measurements are coherent or not, as shown in Fig. 4. The idea is to compare the clear sky model with the measurements only on clear sky days. As the clear sky model was validated in step 2, the model is a reference for what the observations should measure. Then, if the absolute value of nMBE is higher than 8% when comparing the measurements with the clear sky model, it means the measures are incoherent. If the measures are incoherent on the clear sky days, it is considered incoherent on all data detected by the two equations of coherence test for that particular month (Tests 7i and 7ii). For example, in September 2019, just one clear sky day was identified. On this clear sky day, there is a difference of 26% in the nMBE comparing the G_b and $G_{b,cs}$, and a difference of 3% comparing G and G_{cs}. Then, the measurements of G are coherent and the measurements of G_b are incoherent on this particular clear sky day. The data detected by the first two equations of coherence test (Tab. 2 – Tests 7i and 7ii) will receive anomalous data for G_b, good data for G and suspicious data for G_d on September 2019. Fig. 5 shows the clear sky day of September 2019 and the next day, which was a day with high solar radiation variability. It can be noted that G_b is below the clear sky model for direct normal radiation on September 1st, while both global radiation measurements (G_1 and G_2) match with the clear sky model for G and the diffuse horizontal irradiance also corresponds with the clear sky model for G_d on mean, with some difference on the morning. However, the direct normal irradiance presents a nMBE of 26% when compared with the clear sky model for G_b, as it is possible to see in Fig. 4 for September 2019. Then, the component with measurement problems is the direct normal irradiance, causing the difference between the global horizontal measurements (G_1 and G_2) and the sum $G_d + G_b cos(\Theta_Z)$, marked with black points in the plot. The major outliers found by the coherence test in CRESP Petrolina and Araripina stations are related to pyrheliometer underestimates, probably due to soiling. A possible explanation for the pyrheliometer underestimating is that the cylindric geometry of the pyrheliometer with the flat glass in front of the sensor and the surrounding rain screen can contribute to a greater soiling accumulation than in the pyranometer sensor, whose glass is semi-spherical (Geuder and Quaschning, 2006). Furthermore, dust affects the narrow pyrheliometer field of view more than on the pyranometer with its all-sky view. Fig. 6 shows the soiling problem on two stations: CRESP Petrolina and Araripina. The soiling layer stays on the pyrheliometer flat glass and works as a filter from the sunlight photons of direct normal irradiance causing the difference detected by the coherence test with clear sky validation.

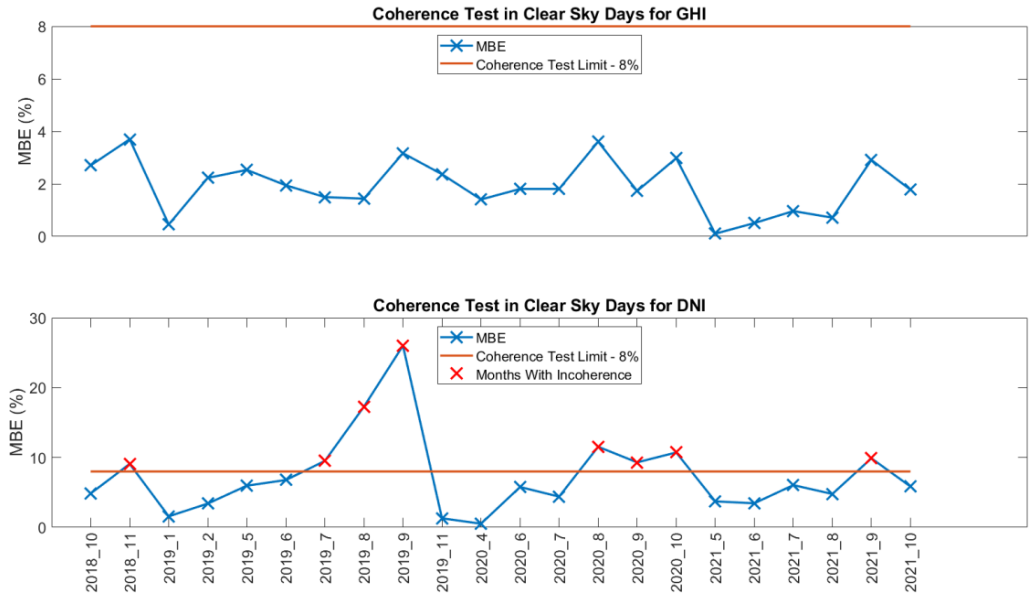


Fig. 4: Monthly analysis to check coherence of clear sky days occurring in each month.

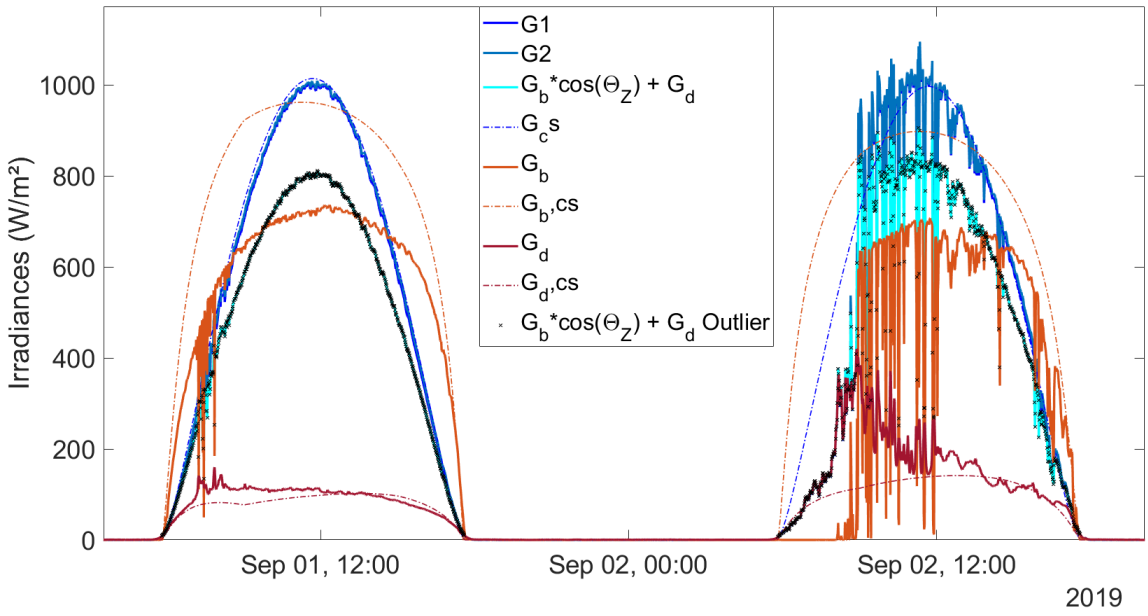
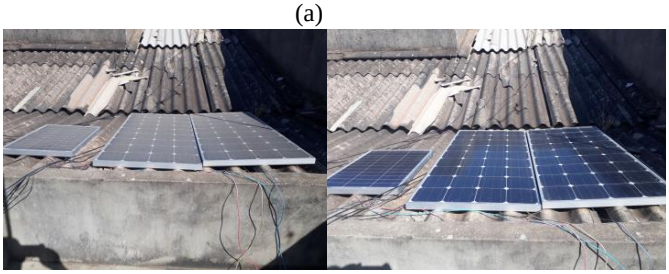


Fig. 5: Results of coherence test in September 2019 for CRESP Station.

The tracker off test presents interesting results for the daily test. For Araripina station, Fig. 7 shows the moments when the tracker did not work. The red points indicate days when the tracker did not work during all-day hours (Tab. 1 – Test 6i) while the black points indicate specific moments of days when the tracker was off (Tab. 1 – Test 6ii). In the detailed plot, two days were detected by the daily test. On 05 and 07 of July, the tracker was off because G_d is similar to G, and G_b is close to zero. It should be noted that the daily test did not detect any problem on July 6th, a day where the tracker was off too. Then, the QA procedure results analysis needs to be complemented to manually put some flags, such as the anomalous flag data in the morning of July 6th for G_d and G_b.

(a)



(b)

Fig. 6: a) Pyrheliometer of CRESP Station before cleaning (left) and after cleaning maintenance (right); b) Soiling on the modules that power the meteorological station of Araripina station before (left) and after (right) cleaning.

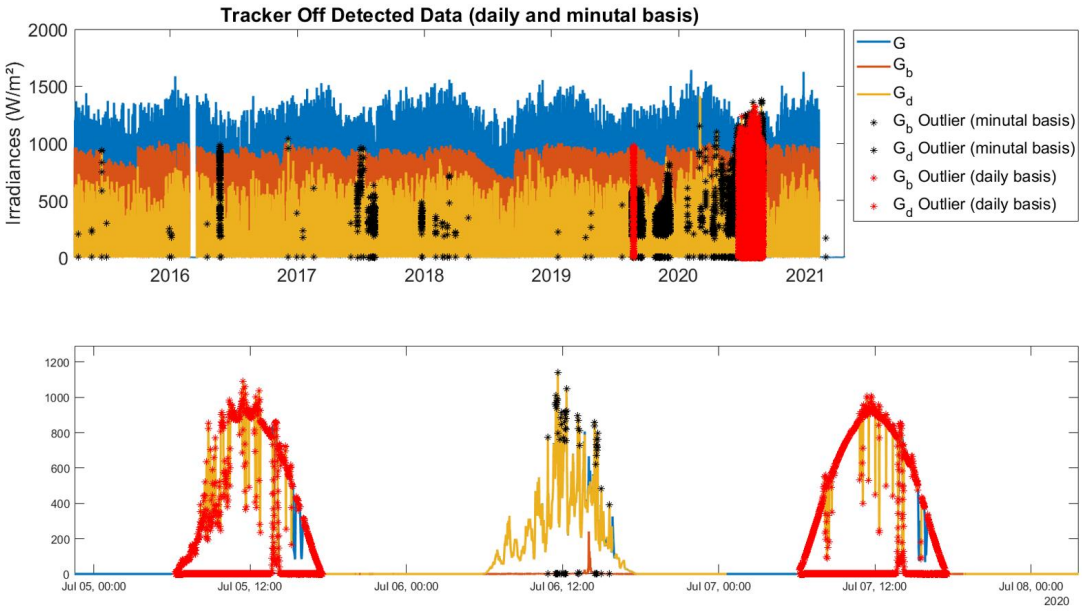


Fig. 7: Results of tracker off test for Araripina station.

In particular, the tracker off test could be helpful to compare the measurements of G and G_d pyranometers, because when the tracker is off, G and G_d should be equal. For example, Fig. 8 shows one tracker off day for SONDA Petrolina station and another for Araripina station, both clear days. The G and G_d pyranometers present close measurements for Araripina station, while in SONDA Petrolina station, G is measuring below G_d and G_{cs}. One linear regression was fitted in SONDA Petrolina, using exclusively the days identified by tracker off test on daily basis. The target of linear regression was G_d, while G was the variable regressor. After this correction, the G and G_d measurements of SONDA Petrolina station match in tracker off days, as shown by the third plot in Fig. 8. This correction is important because the clear sky model for global horizontal irradiance would not have been validated for SONDA Petrolina without the correction of G. The need for a correction was identified by analysing the results of tracker off test on a daily basis (Tab. 1 – Test 6i). Then, for SONDA Petrolina station, first the QA Procedure applies this correction in G component and, after, starts all the data qualification procedures. The performed correction for SONDA Petrolina station is shown in Fig. 9. In the set of days when the tracker was off, it is possible to observe that the diffuse pyranometer (current measuring global irradiance because the tracker is off) presents

higher measurements than the global pyranometer (left plot of Fig. 9). Therefore, a correction using linear regression is applied to the global pyranometer to fit the two global measurements in a line x = y trend, as shown by the right plot of Fig. 9.

Finally, Fig. 10 presents the scatterings k_d x k_t and k_n x k_t to all stations. These scatterings are helpful tools to identify if the QA procedure is filtering the data well. If some points are outliers in these scatterings, measuring problems remain on the dataset. If this occurs, an analysis of the outliers in these scatterings is recommended to complement the QA procedure.

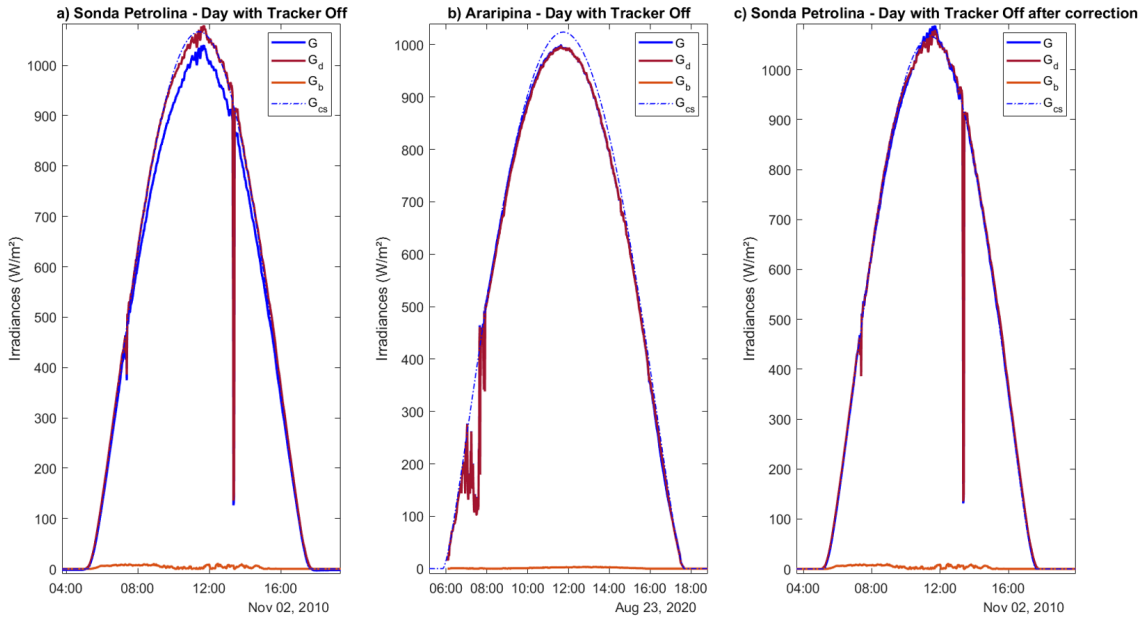


Fig. 8: Comparing the results of tracker off on daily basis for Araripina and SONDA Petrolina stations.

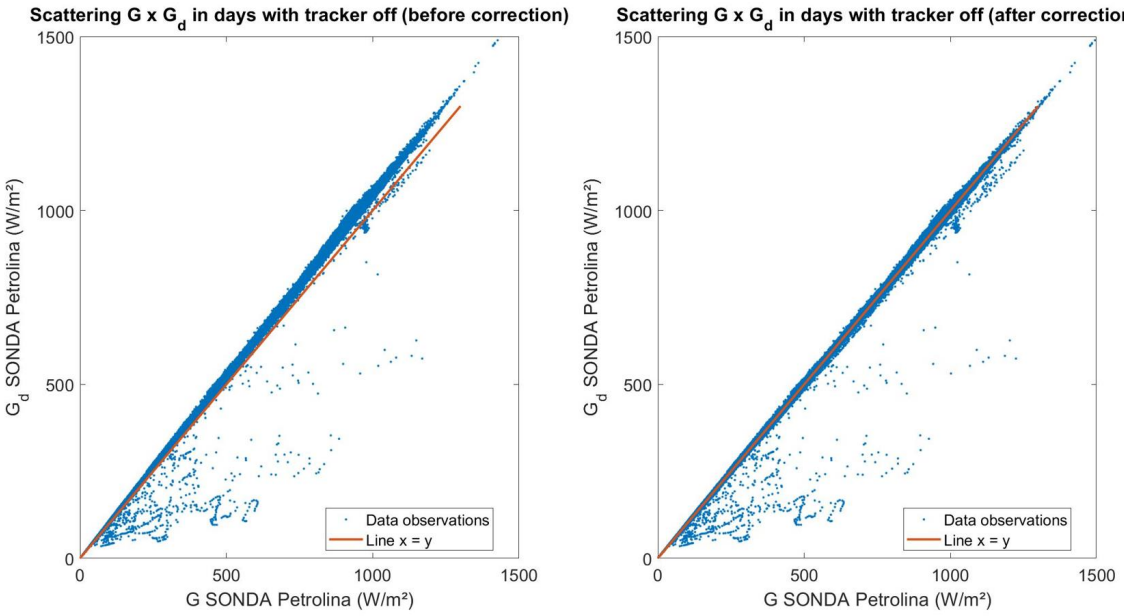


Fig. 9: Scattering G_d *x* G before (a) and after (b) G correction in SONDA Petrolina station.

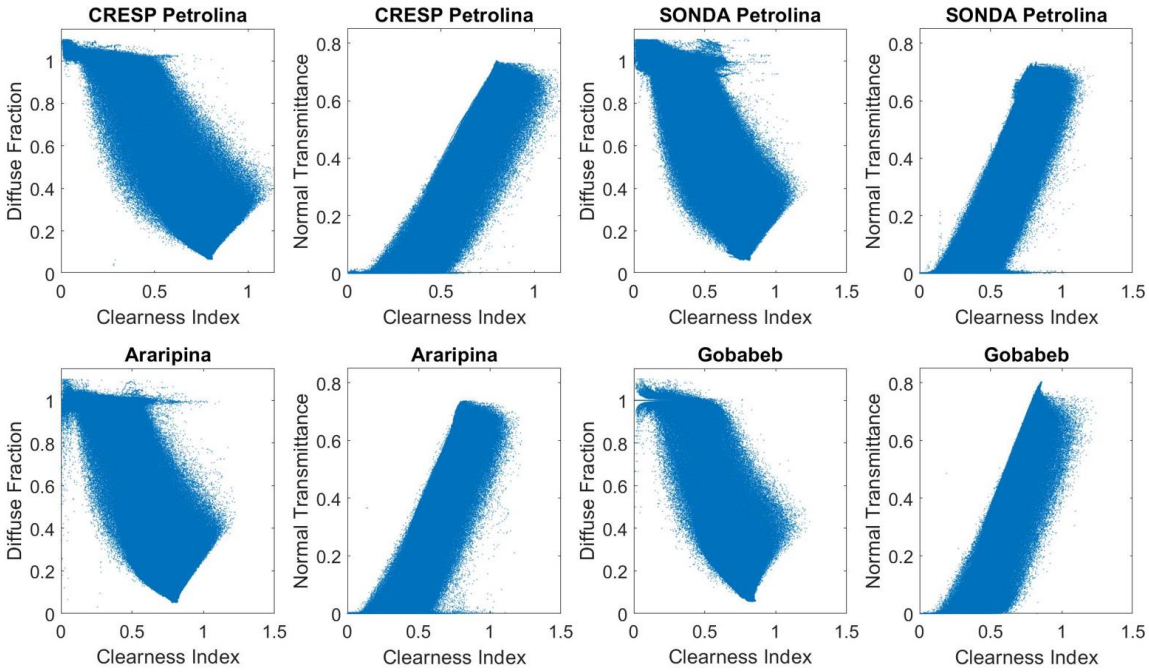


Fig. 10: Scattering k_d x k_t and k_n x k_t for all stations.

4. Conclusion

Quality assurance of the measured data is essential to increase its reliability to apply them in radiation models (separation models, transposition models and photovoltaic power output models) or to evaluate the solar irradiance for a site. The proposed QA procedure applied to the meteorological stations allowed the detection of anomalous data on the irradiances time series. The use of the coherence test with a clear sky validation to define which of the three radiation components are presenting problems detected by the first two equations of the coherence test and the tracker off methodology split into two different tests, one on daily basis and another using minutely measurements, are the two main innovations of the proposed QA procedure.

It was identified that the anomalous data detected by the coherence test in stations CRESP Petrolina and Araripina happened due to problems in the direct normal irradiance component, which presents lower amplitudes than expected. This issue with G_b was attributed to soiling on the pyrheliometers, which was validated empirically according to the maintenance of the stations. Another comparative test, an envelope test, is under development to complement the quality assurance methodology. The envelope test highlights a trend in the k_d x k_t and k_n x k_t scatterings, seeking to draw envelope curves for the upper and lower limits of these scatterings, similar to what was proposed by Yunnes et al. (2005).

5. Acknowledgments

The authors thank the financial support from the National Agency of Electric Energy (ANEEL) and Companhia Hidro Elétrica do São Francisco (Eletrobras CHESF) under the ANEEL R&D Program (CVI 23076.009704/2020-56). Likewise, the authors acknowledge the support of the World Radiation Monitoring Center - Baseline Surface Radiation Network (WRMC – BSRN), Instituto Federal de Pernambuco (IFPE) and Instituto Nacional de Pesquisas Espaciais (INPE) for providing part of the data used in this work and the Coordination for the Improvement of Higher Education Personnel (CAPES) in the scope of the graduate program of the UFPE members.

6. References

Geuder, N., Quaschning, V., 2006. Soiling of irradiation sensors and methods for soiling correction. Solar Energy, v. 80, p. 1402-1409, doi: https://doi.org/10.1016/j.solener.2006.06.001.

Gueymard, C. A., 2018. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Solar Energy, v. 168, p. 2-9, doi: https://doi.org/10.1016/j.solener.2018.04.001.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., Morcrette, E.J.J., 2013. McClear: A New Model Estimating Downwelling Solar Radiation at Ground Level in Clear-Sky Conditions. Atmospheric Measurement Techniques, 6(9):2403–18, doi: https://doi.org/10.5194/amt-6-2403-2013.

Long, C.N., Dutton, E.G., 2002. BSRN Global Network recommended QC tests, V2.0.

Long, C. N. e Shi, Y., 2006. The QCRad value added product: Surface radiation measurement quality control testing, including climatology configurable limits. Atmospheric Radiation Measurement Program Technical Report, https://doi.org/10.2172/1019540.

Maxwell, E., Wilcox, S. e Rymes, M., 1993. Users manual for SERI QC software, assessing the quality of solar radiation data. Solar Energy Research Institute, Golden, CO.NREL, 1993. User's Manual for SERI_QC Software - Assessing the Quality of Solar Radiation Data.

Perez-Astudillo, D., Bachour, D., & Martin-Pomares, L., 2018. Improved quality control protocols on solar radiation measurements. Solar Energy, v. 169, p. 425–433, doi: https://doi.org/10.1016/j.solener.2018.05.028.

Petribú, L. B., Sabino, E., Barros, H., Costa, A., Barbosa, E., Vilela, O.C., 2017. Procedimento objetivo para a garantia de qualidade de dados de radiação solar. Avances en Energías Renovables y Medio Ambiente, vol. 21, p. 67-78.

Raichijk, C., 2012. Control de calidad de mediciones de radiación solar. Presentado en XXXV Reunión de Trabajo de la Asociación Argentina de Energías Renovables y Ambiente (ASADES), Rosario, Argentina.

Rigollier, C., Bauer, O., Wald, L., 2000. On the clear sky model of the ESRA - European Solar Radiation Atlas with respect to the Heliosat method. Solar Energy 68 (1), 33-48. , doi:10.1016/S0038- 092X(99)00055-9.

Ruiz-Arias, J., Alsamamra, H., Tovar-Pescador, J. e Pozo-Vázquez, D., 2010. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. Energy Conversion and Management 51(5): 881-893.

Stein, J. S.; Hansen, C. W.; Reno, M. J., 2012. The Variability Index: A New and Novel Metric for Quantifying Irradiance and Pv Output Variability. World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) Annual Conferen 4(May): 2764–70.

Sönmez, I., 2013. Quality control tests for western Turkey Mesonet. Meteorological Applications 20: 330-337.

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

Younes, S., Claywell, R. e Muneer, T., 2005. Quality control of solar radiation data: present status and proposed new approaches. Energy 30(9): 1533-1549