Type of the Paper (Article)

Quality control of solar radiation data within the South African Weather Service solar radiometric network

Lucky Ntsangwane^{1,2*}, Venkataraman Sivakumar², Brighton Mabasa¹, Nosipho Zwane¹, Katlego Ncongwane¹ and Joel Botai¹

¹ South African Weather Service, 01 Eco park Drive, Eco glades Block B, Cnr Olievenhoutbosch and Ribbon Grass Street, Centurion, Pretoria, South Africa; lucky.ntsangwane@weathersa.co.za

² Discipline of Physics, School of Chemistry and Physics, University of KwaZulu Natal, Private Bag X54001, Durban 4001, South Africa; Venkataramans@ukzn.ac.za

* Correspondence: lucky.ntsangwane@weathersa.co.za; Tel.: +27-012-367-6239

Abstract: Quality control (QC) may be a lengthy and tedious process. As a result, most data users use data from meteorological services without performing data quality checks. The South African Weather Service (SAWS) re-established the national solar radiometric network comprising of 13 new stations within the six climatic zones of the country. This study reports on the performance results of the Baseline Surface Radiation Network (BSRN) QC procedures applied to the solar radiation data within the SAWS radiometric network. The overall percentage performance of the SAWS solar radiation network based on BSRN QC methodology is 97.79%, 93.64%, 91.6% and 92.23% for Long Wave Downward Irradiance (LWD), Global Horizontal Irradiance (GHI), Diffuse Horizontal Irradiance (DHI) and Direct Normal Irradiance (DNI) respectively with operational problems largely dominating the percentage of bad data. The overall average performance of the Surface Solar Radiation Dataset – Heliosat (SARAH) data records for the GHI estimation for all the stations showed a Mean Bias Deviation (MBD) of -8.28 Wm⁻², a Mean Absolute Deviation (MAD) of 9.06 Wm⁻² and the Root Mean Square Deviation (RMSD) of 11.02 Wm⁻². The correlation (quantified by R²) between ground-based and SARAH-derived GHI time series was ~ 0.98. The established network has the potential of providing high quality minute solar radiation data sets (GHI, DHI, DNI and LWD) and auxiliary hourly meteorological parameters vital for scientific and practical applications in renewable energy technologies in South Africa.

Keywords: Quality control, BSRN, solar radiation, satellite-retrieve irradiance, ground stations, validation

1. Introduction

Knowledge of the local solar radiation arriving at the surface of Earth is very important for many different applications; such as crop growth models, architectural designs, planning, designing and sizing of solar energy systems [1, 2, 3]. To be successful in these applications solar radiation measurements are required at strategic sites [1]. Historically, solar radiation data has been measured and recorded by the national meteorological services [4]. Until recently, the South African Weather Service (SAWS) has been the primary source of ground-based solar radiation data in South Africa [5, 6]. The old solar radiometric network collapsed due to technical difficulties and lack of maintenance [5, 7]. Owing to the rapid development of solar-based renewable energy technologies and projects, the demand for reliable and accurate data for site-specific solar resource assessment has increased [4].

Quality control (QC) may be a lengthy and tedious process, as a result, most users are keen to use data directly from meteorological services in confidence without performing an additional and fine data check [8]. To this end, in 2013, the SAWS re-established the national solar radiometric network comprising of 13 new stations within the six climatic zones of the country as shown in Table 1. These stations are equipped with robust and reliable instruments which are suitable for Baseline Surface Radiation Network (BSRN) solar radiation measurements [9]. Climatic zones are the regions with similar climatic conditions [10] and according to Conradie [11] climatic zones were established in order to classify different areas based on their maximum energy demand and maximum energy consumption. Where each radiometric station is located there is an automatic weather station (AWS) measuring hourly temperature, rainfall, pressure, humidity, cloud cover, wind speed and wind direction providing auxiliary meteorological parameters.

Measurements of solar radiation are more susceptible to errors than other meteorological parameters [12]. According to Urraca et al. [12], there are two major sources of these errors related to ground-based solar radiation measurements. These are equipment and operational errors. Equipment errors are inherent to the type and construction of the sensors used in the measuring campaign [4, 8]. Solar measuring sensors produce electric current when reacting with radiation, which is converted into measurement of solar radiation. Solar radiation measurement instruments are also prone to change in sensitivity, thermal offsets, spectral effects, geometry and the environment [4]. On the other hand, operational errors are independent of the type of sensor and involve different factors such as shading by nearby objects or dust covering the dome of the sensor, incorrect levelling, station shut-downs, electric fields in the vicinity of cables or a malfunction in the data-logger, among others. An adequate selection of the place to install the station, as well as a regular maintenance, can ameliorate most of these operational errors.

In order to identify and quantify different all types of errors in the measurements of solar radiation, applying a QC procedure becomes an essential step before using ground-based datasets [4, 8]. According to Huld et al. [13], accurate and reliable solar radiation measurements provide investment grade bankable solar radiation data to the solar energy industry, project developers, decision makers in the financing institutions and policy and also to the scientific community. The accurate ground-based solar radiation is also important for the improvement and validation of satellite-derived solar radiation data and scientific models. In this study, the BSRN QC procedures are applied to the solar radiation data within the SAWS radiometric network.

The paper is organised as follows: Section 2 describes the ground-based solar radiation measurements subjected to the BSRN QC procedures and the satellite-derived solar radiation data used for validation against the ground measurements. Section 3 presents the methods used in applying the BSRN QC procedures to the ground-based data, as well as for the validation of the satellite-derived data. Section 4 discusses the results of the QC and the validation. Conclusions are given in Section 5.

2. Materials and Methods

2.1. Ground-based solar radiation data

The datasets used in this study comprise of ground measurements registered at thirteen stations in South Africa owned and maintained by the South African Weather Service. They are evenly distributed in six different climatic regions [10] over an area bounded by latitudes 23° to 34° south and longitudes 18° to 31° east, see Figure 1. The elevation of the stations ranges from 80 m to almost 1700 m, as described in Table 1.

Table 1. Coordinates of the considered stations in South Africa. The length of the time series, in months, is also indicated, along with the starting date of the dataset registered in every station.

Station	Latitude (S, positive)	Longitude (E, positive)	Altitude (m)	Data Coverage	Climatic Zone
SAWS current					
Prieska	-29.68	22.71	989	2013-09 to 2015-08	Arid
Upington	-28.48	21.12	848	2014-02 to 2017-12	Arid
De Aar	-30.67	23.99	1284	2014-05 to 2017-12	Cold
Irene	-25.91	28.21	1524	2014-08 to 2017-12	Temperature
Nelspruit	-25.39	31.1	870	2014-05 to 2017-12	Hot Interior
Mahikeng	-25.81	25.54	1289	2014-06 to 2017-12	Temperature
Mthatha	-31.55	28.67	744	2014-08 to 2017-12	Subtropical
Bethlehem	-28.25	28.33	1688	2015-01 to 2017-12	Cold Interior
Cape Point	-34.35	18.48	86	2015-02 to 2017-12	Temperature
George	-34.01	22.38	192	2015-01 to 2017-12	Temperature
Durban	-29.61	31.11	91	2016-02 to 2017-12	Subtropical
Polokwane	-23.86	29.45	1233	2015-03 to 2017-12	Temperature
Thohoyandou	-23.08	30.38	619	2015-03 to 2017-10	Hot Interior

Most stations were installed during 2014, although five started registering measurements only in the first quarter of 2015. The present study used data recorded up to December 2017, except for Prieska and Thohoyandou, where the last available ground measurement is from September 2015 and October 2017 respectively. Therefore, the length of the time series depends on the location, ranging from the 23 months in Prieska to the 46 months in Upington. One year of recordings is, nonetheless, always considered at every station in order to guarantee analysing a representative dataset for every location and climate.

Figure 1: Location of the thirteen considered stations and the six climatic regions (Hot interior, Temperate interior, Arid interior, Cold interior, Sub-tropical coastal and Temperate coastal).

The solar radiation stations share the following configuration: a SOLYS 2 sun tracker with shading ball assembly and one first class ISO - 9060 classified CHP1 pyr heliometer (Kipp & Zonen) for measuring the Direct (or beam) Normal Irradiance (DNI). All stations, except De Aar which is part of the BSRN network, are equipped with two high performance pyranometers CMP11 (Kipp & Zonen) for registering the Global and the Diffuse Horizontal Irradiances (hereafter GHI & DHI respectively). The CMP11 pyranometer is fully compliant with all ISO 9060:1990 secondary standard instrument performance criteria. De Aar, being a BSRN station, employs two ventilated CMP21 pyranometers by Kipp & Zonen, which are rated in the highest possible ISO pyranometer performance category. The ventilation units keep the pyranometer's domes clean from frost and water. All stations are equipped with an automatic weather station that registers basic weather parameters such as rainfall, temperature, cloud cover, humidity wind speed, wind direction, pressure and sunshine. Periodical maintenance procedures are applied to the various instruments in order to satisfy the BSRN quality requirements. The ground-based solar radiation database contains one-minute values of all the measured parameters at each station.

2.2. Satellite-derived solar radiation data

The Surface Solar Radiation Dataset – Heliosat (SARAH) [18] is part of the climate data records produced by Satellite Application Facility on Climate Monitoring (CMSAF), where the objective is to produce a temporally homogeneous data record for long time periods suitable for climate analysis, i.e. assessment of anomalies and trends. The SARAH data records are derived using data from the Meteosat Visible Infra-Red Imager instruments of the MFG satellites (Meteosat 2-7) up to the end of 2005, and from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) instruments on the Meteosat Second Generation (MSG) (Meteosat 8-10) satellites after then. SARAH provides data for the GHI and DNI irradiance at the earth surface from 1983 to 2015 at high temporal (down to 30 minutes, but also daily and monthly averages) and spatial ($0.05^{\circ} \times 0.05^{\circ}$) resolutions.

Surface solar radiation is obtained using a modified Heliosat method to calculate the effective cloud albedo (CAL), and the SPECMAGIC clear-sky model [14], which is an extension to spectral bands of the MAGIC model (Mesoscale Atmospheric Irradiance Code) [15]. The SPECMAGIC uses monthly average values of atmospheric water vapour content from the European Centre for Medium-Range Weather Forecasts ERA-interim product and long-term monthly climatologies of aerosol optical depth based on monitoring atmospheric composition and climate [16, 17].

Validation of SARAH, using high-quality ground stations from international networks (e.g., Baseline Surface Radiation Network, BSRN, <http://bsrn.awi.de>) as well as from national networks, has been reported in a number of publications [18-21].

At present the SARAH dataset provided by CMSAF exists in two versions. The dataset used in the present work is based on version 1 of SARAH, with one difference: the hourly data used here are calculated from one satellite image per hour. In contrast, the SARAH version 1 data available from CMSAF use a weighted average of three half-hourly satellite images to calculate the hourly solar radiation values.

2.3. Methods for quality control and validation

2.3.1. Quality control solar radiation data

The schematic diagram illustrating the methodology considered in the present study is given in Figure 2.

Figure 2: Flow chart of the quality control applied to SAWS Solar Radiation data

According to Urraca et al. [20], there are several and diverse QC methods applied to solar radiation data by different meteorological services and independent researchers. The SAWS has preferred to use well-known QC procedures from the BSRN [22]. These QC procedures flag those samples identified out of the normal test limits of data and usually leave the decision of removing flagged cases to the user. In this study, the BSRN QC procedures with three levels of testing was applied on the archived monthly minute data stored in a central database. Each point minute data is associated with its own quality code after the test. These tests can be classified in three major categories: physical possible limits, extremely rare limits, and coherence between measurements or across quantities relationships.

161 • **Physical possible limits**

162 Physical possible limits check for possible physical reasonable maximum and minimum values.
163 These extremal values are assigned codes 1 and 2 corresponding to less than a minimum and greater
164 than maximum reasonable values respectively [12, 22]. Data that did not pass this test is regarded as
165 highly suspect and it is strongly recommended to exclude it from further analysis.

166 • **Extremely rare limits**

167 Extremely rare limits check the data that is in the physical possible limit range for random errors
168 often associated to unusual weather conditions like multi reflexion between broken clouds and the
169 snow surface or a tracking problem (hardware). Data that is beyond the extremely rare cases should
170 at least be visually inspected and if no physical reasons are found it is excluded in the analysis.

171 • **Coherence between measurements**

172 Compared measurements or across quantities relationships is based on the relationship between
173 all three main solar radiation parameters (GHI, DHI, DNI). Data that does not pass this test is difficult
174 to make a decision to exclude them in the analysis because it is unclear which parameter, two
175 parameters or all parameters that caused the inconsistency.

177 **Table 2. Description of QC flags for SAWS solar radiation data. The flag numbers are a**
178 **summation of all 3 QC stages and it indicates a stage where the data failed the QC test.**

Flag number	Description
0	Good quality, passed all three QC tests
5	Missing data (placeholder -999 when archiving)
8	Data is above extremely rare limits
10	Data is above physical possible and extremely rare limit
16	Data is below compared measurement
24	Data is above extremely rare limit (coded as 8)
26	Data is above physical possible and extremely rare limit (coded as 10)
32	Data is above comparison measurement
40	Data is above extremely rare limit and compared measurement
42	Data is above physical possible, extremely rare limit and compared measurement

179
180
181 The minute ground-based solar radiation data of GHI, DHI, DNI and LWD from all 13 SAWS
182 solar radiometric stations as shown in Table 1 above, was subjected to quality check procedures based
183 on BSRN QC standards [22, 23, 24], before the validation was performed. Only the data that passed
184 the first two quality check tests (physically possible limits and extremely rare limits) was used in the
185 validation. Files containing missing values and the data that did not pass the first two BSRN QC tests
186 were flagged and later discarded (replaced by NaN) and was considered for the validation [22, 23,
187 24]. After quality check, months that have less than 10 % of missing data plus data that failed the first
188 two quality checks were used (only a month with 90% or more of available data plus the data that
189 passed the first two quality check tests was used). Moreover, the minute values were averaged to 15
190 minutes and then 4 slots of 15 minute averages were averaged to get hourly mean values [23, 24].
191 Furthermore, all night values, values between sunset (20:00) and sunrise (05:00) based on South
192 African standard time (SAST), i.e., when the solar zenith angle is less than 90 degrees ($SZA < 90$
193 degrees) were replaced by 0. Hourly mean values were then averaged to get daily mean values and
194 subsequently monthly mean values calculated from the daily mean values.

195 **2.3.2. Validation of the satellite-based solar radiation**

196 **Quality controlled irradiance values in the validation**

197 In order to assess the quality of the satellite-based models, the computed monthly mean satellite-
198 retrieved estimates were compared to the monthly averaged ground-based solar radiation data. The

CMSAF-SARAH monthly mean surface incoming shortwave radiation (SIS) data with a spatial resolution of (0.05°x 0.05°) from Meteosat Second Generation (MSG) was used to validate concurrent quality checked monthly average GHI values calculated from minute GHI values measured from 13 SAWS solar radiometric network. According to Schulz et al [25], the CMSAF-SARAH products are accurate enough to be used for solar energy applications and to support meteorological organization with diurnal, sub-seasonal and seasonal solar radiation data sets.

Validation metrics

The validation metrics including the Mean Bias Deviation (MBD), Mean Absolute Deviation (MAD), Root Mean Square Deviation (RMSD), square of correlation coefficient (R^2) were calculated from all the months with 90% or more good data. In addition, Diffuse Fraction (DHI/GHI) and clearness index (GHI/ToA), hereafter referred to as DF and K_T respectively were also calculated from all 13 stations from the months with 90% or more of both GHI and DHI good data. Annual average temperature and humidity of each station from 2013 to 2017 were calculated for each radiometric station using hourly data from AWS

Satellite-retrieved and ground-based solar radiation values of GHI were compared at the different stations for every year and month independently. The mean bias deviation (MBD), the root mean square deviation (RMSD), and mean absolute deviation (MAD) in absolute (Wm^{-2}) values were computed according to Equations 1 to 3 [3, 26, 27]. In addition to these, the R^2 correlation coefficient was also calculated using Equation 4 [26, 27].

$$MBD = \frac{1}{N} \sum_{i=1}^N (G_i^s - G_i^m) \quad (1)$$

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (G_i^s - G_i^m)^2} \quad (2)$$

$$MAD = \frac{1}{N} \sum_{i=1}^N |G_i^s - G_i^m| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (G_i^m - \bar{G}^m)^2}{\sum_{i=1}^N (G_i^m - \bar{G}^m)^2} \quad (4)$$

Where G_i^s is the satellite-retrieved irradiance value at the i^{th} time point and G_i^m is the ground-based solar radiation value for that timestamp, N is the total number of points considered in the period of time analysed, year or month, and $\bar{G}^m$ is the average ground-based solar radiation value during the considered time period.

3. Results and Discussion

3.1 BSRN QC results

Considering the BSRN QC results shown in Table 3 and Table 4 below, the overall percentage performance of SAWS' solar radiation network based on BSRN QC procedures is 97.79%, 93.64%, 91.6% and 92.23% for LWD, GHI, DHI and DNI respectively. Operational problems dominated the percentage of bad data as follows LWD 2.21%, GHI 1.6%, DHI 3.57% and DNI 3.57%. Only data represented by code 0 is regarded as the data that passed all the three BSRN QC tests, and thus representing the overall percentage performance. Code 5 represents missing data or data that was not recorded, which is indicative of overall percentage of the operational problems and errors. On the other hand, code 8 and 10 represent data that failed first and second BSRN QC tests respectively. All the irradiance data sets passed the first QC test and only 0.3% of GHI and DHI data failed the second QC test. Code 16 and 32 represent data that failed the third BSRN QC test, with the results

249 showing that at least 4% of GHI, 4.3% of DHI and 4.2% of DNI failed the measurement coherence test
250 between them. Code 40 represent data that failed both the second and third BSRN QC test and 42
251 represent data that failed all three BSRN QC tests. For the validation, only data coded 0, 16 and 32
252 was used. This data was regarded as good quality data. On the other hand, data bearing the codes 5,
253 8, 10, 40 and 42 was discarded and replaced by NaN as they were regarded as bad data. The final
254 number of monthly data used at every station depends on the QC results of the measured data. In
255 general, only the data with more than (90 % sum of the data coded 0, 16, 32) was used in the validation
256 study.

Table 3. Weighted averages of all BSRN QC codes for GHI, DHI and DNI per station from when the station started recording a full month of data to end of December 2017.

Parameter	Global Horizontal Irradiance (GHI)								Diffuse Horizontal Irradiance (DHI)								Direct Normal Irradiance (DNI)							
Code/Station	0	5	8	10	16	32	40	42	0	5	8	10	16	32	40	42	0	5	8	10	16	32	40	42
Prieska	96.91	2.79	0	0	0	0.27	0	0	96.77	2.79	0.15	0.01	0.17	0.03	0.01	0	96.85	2.79	0.07	0.01	0.26	0.02	0	0
Upington	97.29	1.61	0	0	0.7	0.41	0	0	97.2	1.61	0.06	0.06	0.39	0.68	0.01	0	97.23	1.57	0.17	0	0.29	0.54	0	0
De Aar	97.02	2.13	0	0	0	0.8	0	0	97.02	2.24	0.21	0.01	0.79	0.05	0.01	0	96.79	2.13	0.21	0	0.79	0.06	0	0
Irene	96.89	0.02	0.1	1.2	0.8	0.54	0.14	0	96.93	0.02	0.2	1.04	0.89	0.79	0	0	96.86	0.02	1.39	0	0.87	0.79	0	0
Nelspruit	93.45	0.01	0.1	0.8	0.1	5.58	0.02	0	93.42	0.01	0.08	1.24	5.14	0.08	0	0.01	93.64	0.01	0.7	0	5.56	0.1	0	0
Mahikeng	94.86	0.04	0.2	0.8	0.4	3.62	0.05	0	94.8	0.04	0.27	0.83	3.72	0.4	0	0	95.1	0.04	0.83	0	3.74	0.38	0	0
Mthatha	99.16	0.01	0	0	0.2	0.6	0.02	0	99.16	0.06	0.02	0	0.6	0.21	0	0	99.15	0	0.02	0	0.62	0.21	0	0
Bethlehem	94.7	0.12	0.1	0.9	0.7	3.31	0.14	0	92.12	2.12	1.51	1.06	3.45	0.65	0.07	0.01	94.74	0.12	1	0.01	3.45	0.71	0	0
Cape Point	92.37	3.13	0	0	0.3	4.17	0.03	0	90.56	3.13	1.77	0.02	4.19	0.28	0.02	0	92.38	3.13	0.04	0	3.6	0.26	0	0
George	84.07	2.85	0.1	0	2.5	11.32	0.04	0	83.81	5.47	0.38	0.01	11.6	1.23	1.12	0.01	84	5.63	0.09	0	11.45	2.35	0	0
Durban	95.59	0.01	0	0	1.5	2.87	0	0	74.98	20.64	0.08	0	2.87	1.51	0	0.01	75.7	20	0	0	2.86	1.43	0	0
Polokwane	84.12	2.32	0	0.2	0.5	7.26	0.1	0	83.64	2.32	0.66	0.09	7.44	0.46	0	0	84.06	2.32	0.25	0	7.44	2.28	0	0
Thohoyandou	90.84	5.61	0	0	4.9	1.16	0	0	90.41	6	0.01	0	1.6	4.46	0	0	92.7	6	0	0	1.21	1.64	0	0.01
Overall	93.6	1.59	0	0.3	1	3.22	0.04	0	91.6	3.57	0.42	0.34	3.29	0.83	0.1	0.01	92.23	3.37	0.37	0	3.24	0.83	0	0.01

Parameter	Long-Wave Downward Irradiance (LWD)							
Code/Station	0	5	8	10	16	32	40	42
De Aar	97.79	2.21	0	0	0	0	0	0
Overall	97.79	2.21	0	0	0	0	0	0

Table 4. Weighted averages of all BSRN QC codes for LWD at De Aar from when the station started recording a full month of data to end of December 2017.

3.2 Validation of the satellite-derived solar radiation product

Considering the results obtained in all the stations, the overall average performance of the SARAH data record for the GHI estimation showed an MBD of -8.28 Wm⁻², MAD of 9.06 Wm⁻² and RMSD of 11.02 Wm⁻². Analysing the correlation between ground-based and satellite-derived GHI time series with the R² coefficient, the average performance of the SARAH satellite product in the estimation of the GHI values was ~ 0.98.

Table 5 contains the absolute average MBD, RMSD and MAD values obtained from the validation of the complete valid time series of the SARAH GHI estimates at every station. Besides, the R² coefficient, the number of months used at every location are also indicated.

From the validation of the global irradiance estimates, it can be observed that the SARAH product provided accurate estimates of the monthly average GHI values in every location other than Durban and Cape Point, where it also showed the highest overestimation. This overestimation from at these locations could indicate some problems with the measuring instruments. The lowest levels of the global irradiance, were registered at these coastal stations, where the satellite products tended to overestimate the solar resource to a greater extent. This could indicate either a problem with the ground measurements or a misinterpretation of the input parameters used by the satellite methods. The deviations observed in these stations could indicate a possible underestimation of the attenuation suffered by solar radiation in the atmosphere.

Months with 90% or more of minute good data were used to calculate monthly mean GHI and DHI. Table 5 provides information on the number of months (for each station) that were used for validation analysis. Furthermore, the validation results show that almost all the sites had a negative MBD (observation minus satellite) values implying that SARAH derived irradiance parameters were consistently overestimated compared to ground measurements. According to Posselt et al [28], the validation accuracy threshold for MAD of monthly mean GHI against SARAH monthly mean GHI ought to be 15 Wm⁻², target accuracy threshold is 10 Wm⁻² and optimal accuracy threshold is 8 Wm⁻². The comparison between concurrent SARAH GHI against SAWS GHI monthly means showed a great similarity with MAD of less than 15 Wm⁻² in 11 of 13 stations. Figure 3 shows only Cape Point (temperature coastal) and Durban (subtropical coastal) stations had an MAD greater than the validation threshold accuracy of 15 Wm⁻², recording 18.9 Wm⁻² and 19.0 Wm⁻² respectively.

Prieska, Upington, De Aar, Irene, Mafikeng, Bethlehem, Polokwane (stations located in the arid climatic regions, cold interior and temperature interior) reached an optimal threshold with a MAD of less than 8 Wm⁻² when compared to the SARAH GHI. On the other hand, Nelspruit (hot interior) reached a target accuracy threshold with MAD less than 10 Wm⁻², while Thohoyandou (hot interior), George (temperature coastal) and Mthatha (Subtropical coastal) meet the validation threshold with MAD less than 15 Wm⁻².

Table 5. Results from the validation of the global horizontal irradiance (GHI) estimates retrieved from the Surface Solar Radiation Dataset – Heliosat (SARAH), which are comparable to other countries.

Station	Latitude (S, positive)	Longitude (E, positive)	Altitude (m)	N (Months)	MBD (W/m ²)	MAD (W/m ²)	RMSD (W/m ²)	(r ²)
SAWS current								
Prieska	-29.68	22.71	989	21	-4.2	4.2	5.4	0.998
Upington	-28.48	21.12	848	44	0.9	3.8	5.2	0.995
De Aar	-30.67	23.99	1284	40	-0.8	3.6	4.6	0.997
Irene	-25.91	28.21	1524	41	-6.9	7.8	10.0	0.981
Nelspruit	-25.39	31.1	870	42	-8.9	9.0	11.4	0.958
Mafikeng	-25.81	25.54	1289	23	-4.1	5.1	6.9	0.989
Mthatha	-31.55	28.67	744	41	-10.3	10.3	12.1	0.982
Bethlehem	-28.25	28.33	1688	30	-5.2	5.6	7.7	0.989
Cape Point	-34.35	18.48	86	17	-18.9	18.9	21.5	0.991
George	-34.01	22.38	192	36	-12.5	12.5	13.5	0.997
Durban	-29.61	31.11	91	23	-19.0	19.0	20.1	0.981
Polokwane	-23.86	29.45	1233	29	-7.4	7.6	11.6	0.964
Thohoyandou	-23.08	30.38	619	30	-10.3	10.4	13.2	0.964
Spain								
Almeria	36.85	-2.39		12	-5.4	7.5	9.8	0.990
Malaga	36.72	-4.48		12	-5.9	7.6	9.7	0.990
Bilbao	43.17	-2.91		12	6.5	10.2	13.9	0.980
Sweden								
Visby	57.67	18.35		212	-4.7	5.8	7.7	0.890
Lund	55.71	13.21		222	0.8	5.0	7.1	0.900
Karlstad	59.36	13.47		214	-3.3	4.9	7.0	0.900
Finland								
Helsinki-Kumpu	60.2	24.96		35	-5.5	6.4	8.0	0.900
Uto	59.78	21.37		97	-8.1	8.2	10.1	0.920

Figure 3: Monthly MAD values for the validation of the GHI estimates derived from the SARAH product for all locations.

3.3 Diffuse Fraction and Clearness index

The clearness index and diffuse fraction given in Figure 4 depict higher values (near 1) of clearness index suggesting a clear sky and calm atmosphere. In case of Upington and De Aar, clearness index averages > 0.6 throughout the year. The diffuse fraction of solar radiation was also calculated for all locations. It also varies from 0 to 1. Higher values indicate more aerosols and clouds. The values for diffuse fraction for Upington and De Aar were always < 0.3 in all of the months. This indicates suitability of the locations for solar energy prospecting.

Figure 3: Monthly Diffuse Fraction and Clearness Index for one station per climatic region.

4. Conclusion

The BSRN QC tests proved to be the effective and efficient in detecting errors at different stations. Furthermore, the validation of the CMSAF-SARAH GHI estimates against ground-based solar radiation with the SAWS radiometric network showed that:

- The overall average performance of the SARAH data record for the GHI estimation for all the stations exhibits MBD of -8.28 Wm^{-2} , MAD of 9.06 Wm^{-2} and RMSD of 11.02 Wm^{-2} . These values compare favorably to the recommended validation accuracy threshold for MAD of monthly mean GHI of 15 Wm^{-2} , target accuracy threshold of 10 Wm^{-2} and optimal accuracy threshold of 8 Wm^{-2} . High values of the correlation ($R \sim 0.98$) between ground and SARAH derived irradiance parameters were also obtained. The higher values of clearness index observed in Upington and De Aar indicate the clear sky and atmosphere condition.
- The SARAH estimates can provide the basis for further analysis, such as the one presented in this paper on annual PV electricity production.

The overall percentage performance of SAWS' solar radiation network based on BSRN QC procedures is 97.79%, 93.64%, 91.6% and 92.23% for LWD, GHI, DHI and DNI respectively demonstrating the potential value of SAWS solar resources database's practical and scientific applications in South Africa.

5. Acknowledgements

Part of the work presented here was carried out within the 'Solar resource mapping for South Africa' project funded by the Department of Science and Technology. SAWS staff from regional offices for the continuous efforts in helping with station inspection, cleaning and maintenance. BSRN staff who helped us with some answers related to the BSRN toolbox software operation. Dr Roland Vogt from the University of Basel for providing us with the IDL BSRN programme, Mr Louis van Hemert from SAWS who helped with the FORTRAN code to convert data from station-to-archive format and SAWS staff in the Renewable Energy Application group for their inputs in the daily operation of the radiometric network.

6. References

-
- [1] Jäger-Waldau, A. PV Status Report 2016. EUR 28159 EN; **2016**, doi:10.2790/682995.
- [2] International Energy Agency. Analysis and forecast to 2022. *IEA Publications* **2017**, ISSN 2520-2774.
- [3] Sengupta, M.; Habte, A.; Gueymard, C.; Wilbert, S.; Renné, D. Best practices handbook for the collection and use of solar resource data for solar energy applications, Second edition. National Renewable Energy Laboratory (NREL): Colorado, United States, **2017**, NREL/TP-5D00-68886.
- [4] Antoñanzas-Torres, F.; Sanz-Garcia, A.; Martinez-de-Pison-Ascacibar, F.J.; Perpiñan-Lamiguero, O. Evaluation and improvement of empirical models of global solar irradiation: Case study northern Spain. *Renewable Energy* **2013**, 60, 604-614.
- [5] Meyer, A.J.; van Niekerk, J.L. Towards a Verified DNI Map for South Africa. In Third Southern African Solar Energy Conference, Kruger National Park, South Africa, 11 – 13 May **2015**.
- [6] Brooks, M.J.; du Clou, S.; van Niekerk, W.L.; Gauche, P.; Leonard, C.; Mouzouris, M.J.; Meyer, R.; van der Westhuizen, N.; van Dyk, E.E.; Vorster, F.J. SAURAN: A new resource for solar radiometric data in Southern Africa. *Journal of Energy in Southern Africa* **2015**, 26, 2-10.
- [7] Ciolkosz, D. SASRAD. An Hourly-Timestep Solar Radiation Database for South Africa. *Journal of Energy in Southern Africa* **2009**, 20, 25–34.
- [8] Ineichen, P.; Barroso, C.S.; Geiger, B.; Hollmann, R.; Marsouin, A.; Müller, R. Satellite application facilities irradiance products: Hourly time step comparison and validation over Europe. *International Journal of Remote Sensing* **2009**, 30, 5549-5571.
- [9] Lee, C. Baseline Surface Radiation Network, World Climate Research Programme, An overview of the requirements and operational issues. Derived from the BSRN Operations Manual Version 2.1, Customer Services Specialist, Kipp & Zonen B.V, 2014.
- [10] SANS 204. South African National Standard. Energy efficiency in buildings, Part 1: General requirements. SABS Standards Division. Available online: https://www.sabs.co.za/Sectors-and-Services/Services/Energy/energy_sp.asp (Accessed on 15 September 2013).
- [11] Conradie, D.C.U. South Africa's Climatic Zones: Today, Tomorrow, International Green Building Conference and Exhibition Future Trends and Issues Impacting on the Built Environment, Sandton, South Africa, July 25-26 July **2012**.
- [12] Urraca, R.; Gracia-Amillo, A.M.; Huld, T.; Martinez-de-Pisona, F.J.; Trentmanne, J.; Lindfors, A.V.; Riihelä, A.; Sanz-Garcia, A. Quality control of global solar radiation data with satellite-based products. *Solar Energy* **2017**, 158, 49–62.
- [13] Huld, T.; Mueller, R.; Gambardella, A. A new solar radiation database for estimating PV performance in Europe and Africa. *Solar Energy* **2012**, 86, 1803-1815.
- [14] Gracia-Amillo, A.; Huld, T.; Müller, R. A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation. *Remote Sens.* **2014**, 6, 8165-8189.
- [15] Müller, R.; Matcoukas, C.; Gratzki, A.; Behr, H.D.; Hollmann, R. The CM-SAF Operational scheme for the satellite based retrieval of solar Surface irradiance – A LUT based eigenvector hybrid approach. *Remote Sens. and Environ.* **2009**, 113, 1012-1024.

[16] Inness, A.; Baier, F.; Benedetti, A.; Bouarar, I.; Chabrillat, S.; Clark, H.; Clerbaux, C.; Coheur, P.; Engelen, R.J.; Errera, Q.; *et al.* The MACC reanalysis: An 8-year data set of atmospheric composition. *Atmos. Chem. Phys.* **2013**, *13*, 4073-4109.

[17] Müller, R.; Pfeifroth, U.; Träger-Chatterjee, C. Towards optimal aerosol information for the retrieval of solar surface radiation using Heliosat. *Atmosphere* **2015**, *6*, 863-878.

[18] Müller, R.; Pfeifroth, U.; Träger-Chatterjee, C.; Trentmann, J.; Cremer, R. Digging the METEOSAT Treasure – 3 Decades of Solar Surface Radiation. *Remote Sens.* **2015**, *7*, 8067-8101.

[19] Riihelä, A.; Carlund, T.; Trentmann, J.; Müller, R.; Lindfors, A.V. Validation of CM SAF Surface Solar Radiation Datasets over Finland and Sweden. *Remote Sens.* **2015**, *7*, 6663-6682.

[20] Urraca, R.; Gracia-Amillo, A.M.; Koubli, E.; Huld, T.; Trentmann, J.; Riihelä, A.; Lindfors, A.V.; Palmer, D.; Gottschalg, R.; Antonanzas-Torres, F. Extensive validation of CM SAF surface radiation products over Europe. *Remote Sens. Environ.* **2017**, *199*, 171-186.

[21] Žák, M.; Mikšovský, J.; Pišoft, P. CMSAF Radiation Data: New Possibilities for Climatological Applications in the Czech Republic. *Remote Sens.* **2015**, *7*, 14445-14457.

[22] Long, C.N.; Dutton, E.G. BSRN Global Network recommended QC tests, V2.0. Available online: http://epic.awi.de/30083/1/BSRN_recommended_QC_tests_V2.pdf (Accessed on 02 December 2017).

[23] Roesch, A.; Wild, M.; Ohmura, A.; Dutton, E.G.; Long, C.N.; Zhang, T. Assessment of BSRN Radiation Records for the Computation of Monthly Means. *Atmos. Meas. Tech.* **2011**, *4*, 339-354.

[24] Thomas, C.; Saboret, L.; Wey, E.; Blanc, P.; Wald, L. Validation of the New Helioclim-3 Version 4 Real-Time and Short-Term Forecast Service Using 14 BSRN Stations. *Adv. Sci. Res.* **2016**, *13*, 129-136.

[25] Schulz, J.; Albert, P.; Behr, H.D.; Caprion, D.; Deneke, H.; Dewitte, S.; Dürr, B.; Fuchs, P.; Gratzki, A.; Hechler, P.; *et al.* Operational climate monitoring from space: THE EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF). *Atmos. Chem. Phys.* **2009**, *9*, 1687-1709.

[26] 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. Best practices handbook for the collection and use of solar resource data for solar energy applications, National Renewable Energy Laboratory (NREL): Colorado, United States, **2015**, NREL/TP-5D00-63112.

[27] Lorenz, E.; Kühnert, J.; Heinemann, D.; Nielsen, K.P.; Remund, J.; Müller, S.C. Comparison of irradiance forecast based on numerical weather prediction models with different spatio-temporal resolutions. In 31st EU PVSEC in Hamburg, Germany 2015.

[28] Posselt, R.; Mueller, R.W.; Stockli, R.; Trentmann, J. Remote Sensing of Solar Surface Radiation for Climate Monitoring -The CM-SAF Retrieval in International Comparison, *Remote Sens. Environ.* **2012**, *118*, 186-198.

Conflicts of Interest: The authors declare no conflict of interest.