Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface

Alexandre Boilley, Lucien Wald

To cite this version:

Alexandre Boilley, Lucien Wald. Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface. Renewable Energy, Elsevier, 2015, 75, pp.135-143. <10.1016/j.renene.2014.09.042>. <hal-01074107>

HAL Id: hal-01074107
https://hal-mines-paristech.archives-ouvertes.fr/hal-01074107

Submitted on 24 Oct 2014

Renewable Energy 75 (2015) 135–143

Contents lists available at ScienceDirect

Renewable Energy

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

Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface

Alexandre Boilley [a], [b], Lucien Wald [a], *

[a] MINES ParisTech, Center for Observation, Impacts, Energy, BP 207, 06904 Sophia Antipolis Cedex, France
[b] Transvalor, Mougins, France

ARTICLE INFO

Article history:
Received 16 April 2014
Accepted 8 September 2014
Available online

Keywords:
Meteorology
Solar radiation
Solar energy
Re-analyses
Measurements
Validation

ABSTRACT

This paper compares the daily solar irradiation available at surface estimated by the MERRA (Modern-Era Retrospective Analysis for Research and Applications) re-analysis of the NASA and the ERA-Interim re-analysis of the European Center for Medium-range Weather Forecasts (ECMWF) against qualified ground measurements made in stations located in Europe, Africa and Atlantic Ocean. Using the clearness index, also known as atmospheric transmissivity or transmittance, this study evidences that the re-analyses often predict clear sky conditions while actual conditions are cloudy. The opposite is also true though less pronounced: actual clear sky conditions are predicted as cloudy. This overestimation of occurrence of clear sky conditions leads to an overestimation of the irradiation and clearness index by MERRA. The overall overestimation is less pronounced for ERA-Interim because the overestimation observed in clear sky conditions is counter-balanced by underestimation in cloudy conditions. The squared correlation coefficient for clearness index ranges between 0.38 and 0.53, showing that a very large part of the variability in irradiation is not captured by the re-analyses. Within an irradiation homogeneous area, the variability of the bias, root mean square error and correlation coefficient are surprisingly large. MERRA and ERA-Interim should only be used in solar energy with proper understanding of the limitations and uncertainties. In regions where clouds are rare, e.g. North Africa, MERRA or ERA-Interim may be used to provide a gross estimate of monthly or yearly irradiation. Satellite-derived data sets offer less uncertainty and should be preferred.

1. Introduction

The solar radiation impinging at ground level is an essential variable in solar energy. It is often called surface solar irradiance or irradiation (SSI) in solar energy, solar flux or solar exposure when dealing with measurements, and downwelling shortwave flux, or downwelling surface shortwave flux in numerical weather modelling. The present article deals with the surface daily solar irradiation, i.e. the energy received per surface unit during a day. Applications under concern are construction of time-series or maps for locating favourable areas for solar plants, pre-feasibility studies or monitoring of existing plants.

There are several means to assess the daily irradiation [10]. Ground measuring stations and satellite observations are two of them, sometimes in combination [2,29]. Re-analyses are the third means. Models for weather forecasts are used in a re-analysis mode to reproduce what was effectively observed. The SSI in re-analyses is diagnostic. It is computed by a radiative transfer model and hence depends on the representation of the whole set of radiatively active variables of the atmospheric column above the point. There are several available re-analyses. Of interest here are the MERRA (Modern-Era Retrospective Analysis for Research and Applications) re-analysis proposed by the NASA Global Modeling and Assimilation Office and the ERA-Interim re-analysis of the ECMWF (European Center for Medium-range Weather Forecasts).

Advantages of re-analyses for companies and practitioners in solar energy are the worldwide coverage, the multi-decadal temporal coverage, and their availability at no cost. Re-analysis estimates should not be mistaken with observed data in SSI because while the re-analysis method assimilates state variables such as temperature, moisture and wind, physics variables such as radiation and cloud properties derive from a model and include the uncertainty of this model. However, because of the advantages listed above in coverage, availability and costs, re-analyses are appealing to companies and several are using re-analyses in their daily work. This paper aims at establishing the quality of re-

---

* Corresponding author. Tel.: +33 6 86 40 06 23; fax: +33 4 93 67 89 08.
E-mail address: Lucien.wald@mines-paristech.fr (L. Wald).

http://dx.doi.org/10.1016/j.renene.2014.09.042

analyses when compared to qualified ground measurements. Ref. [15] published a similar work using the MERRA re-analysis in order to estimate the climatological variability of the photovoltaic power production over Czech Republic. Our study complements their work and other similar studies in meteorology for limited areas and adds more evidence by comparing MERRA re-analyses to a large number of ground stations located in Europe, Africa and Atlantic Ocean. The limitation in geographical coverage of the present study is explained by the expertise of the authors who are dealing with this area for long. In addition, our study reports on the ERA-Interim re-analysis.

To better apprehend the possible benefits of re-analyses, a comparison is also performed with the database HelioClim-1 of daily irradiation created by MINES ParisTech within the HelioClim project [3] and extensively validated against ground measurements. HelioClim-1 is a well-known database of easy access on the Web at no cost with many usages and approximately 400 requests per working day [17]. It will be seen whether the re-analyses offer better accuracy than HelioClim-1.

Refs. [4,24] have compared measurements made at several ground stations located in the same area: Mozambique, to HelioClim-1 and have found that though Mozambique is fairly homogeneous regarding SSI, the differences between HelioClim-1 and ground measurements are spatially variable. Similar conclusions were reported by Ref. [1] for North Africa (Algeria, Egypt and Tunisia) and Ref. [20] for Southern Africa. This paper examines whether re-analyses exhibit less variable errors in homogeneous climatic areas.

2. Material and methods

2.1. MERRA re-analysis

The MERRA data set [22] has a resolution of 0.5° × 0.65° with 72 vertical levels from ground to 0.01 hPa. The radiative transfer model for shortwave radiation (CLIRAD-SW) is described in Ref. [9]. MERRA includes an odd-oxygen family transport model providing the ozone concentration necessary for solar absorption. Production and loss of ozone as well as other optically active species are specified from climatology of the Goddard two-dimensional chemistry and transport model [12]. The hourly SSI estimates are horizontally interpolated using a bilinear interpolation technique, to the measurement site from the closest four surrounding grid cells with a weighting factor that is inversely proportional to the distance. Daily irradiation is computed by summing the hourly SSI estimates after multiplying them by the number of seconds in 1 h.

2.2. ERA-interim re-analysis

The ERA-Interim data set [11] has a resolution of 0.75° × 0.75° and counts 60 vertical levels from ground to 0.1 hPa. Ref. [11] noted an overestimation of 2 W m^{-2} of the incoming radiation at the top of the atmosphere. The radiative transfer model uses the prognostic water vapour and cloud variables (cloud cover, cloud condensed water) from the meteorological model and climatologic values for aerosols, carbon dioxide, trace gases and ozone. The SSI is estimated every 3 h. Similarly to MERRA, the SSI is bi-linearly interpolated to the measurement site. Daily irradiation is computed by summing the eight available SSI estimates after multiplying them by the number of seconds in 3 h.

2.3. HelioClim-1 database

The HelioClim Project is an ambitious initiative of MINES ParisTech launched in 1997 after preliminary works in 80's [3] to increase knowledge on the SSI and to offer SSI values for any site, any instant over a large geographical area and long period of time, to a wide audience. The project comprises several databases that cover Europe, Africa and the Atlantic Ocean. These databases use satellite images as inputs for their creation and updating. The HelioClim-1 database offers daily means of the SSI for the period 1985–2005.

The accuracy of the HelioClim-1 data has been assessed by comparison with ground measurements made by high-quality pyranometers on a daily basis. If well-calibrated and well-maintained, these pyranometers exhibit a relative uncertainty of 10% of the daily mean of SSI at a 95 per cent confidence level [27]. Ref. [18] compared 55 sites in Europe for the period June 1994–July 1995 and 35 sites in Africa for the period 1994–1997. Ref. [1] compared HelioClim-1 data with ground measurements in Algeria, Egypt, and Tunisia. Refs. [4,24] performed a similar study for Mozambique, while Ref. [20] focused on Southern Africa. These works demonstrated that the HelioClim-1 database offers good quality for Africa, the Mediterranean Basin, and more generally for latitudes comprised between −45° and +45°. Outside these limits, the quality may decrease because of the characteristics of the satellite images used for the construction of HelioClim-1 [3] though this is not a systematic effect and local conditions may prevail. The HelioClim-1 has many usages as illustrated by published works in various domains: oceanography, climate, energy production, life cycle analysis, agriculture, ecology, human health, and air quality [17].

The Global Earth Observation System of Systems (GEOSS) is a project aiming at proactively linking together existing and planned observing systems around the world and supporting the development of new systems where gaps currently exist. The GEOSS Data-CORE (GEOSS Data Collection of Open Resources for Everyone) is a distributed pool of documented data sets with full, open and unrestricted access at no more than the cost of reproduction and distribution. The HelioClim-1 database has been identified as a Data-CORE by the GEOSS in November 2011 [14]. Previously, HelioClim-1 was open to researchers and students at no cost on a case-by-case basis. HelioClim-1 can easily be accessed at no cost on the Web (www.soda-is.com).

2.4. Ground measurements

National meteorological services (NMS) usually measure solar radiation at a few sites. Data are sent to the World Radiation Data Center (WRDC), a laboratory of the Voeikov Main Geophysical Observatory in Saint-Petersburg, Russia, under the control of the World Meteorological Organization (WMO). There, the data are archived and published. They are available only for research and educational communities of the countries participating to WMO for non-commercial activities. Quality of measurement is difficult to assess from the WRDC archives. All data are scrutinized at WRDC and quality-flagged before entering archives. No information on uncertainty other than the flag is provided with the radiation data. It is considered that these data meet the requirements set by WMO for international exchange: relative uncertainty is 5%–10% for good to moderate quality [27].

Efforts were made and are being made by the WRDC to publish data on the Web. For data prior to 1994, a joint effort by the WRDC and the National Renewable Energy Laboratory (NREL) of the USA resulted in an automatic delivery system based on e-mail (wrdc-mgo.nrel.gov). This system is very convenient though it has a few drawbacks. The major one deals with the format of data which are returned in ASCII format. Sometimes spaces between successive values are replaced by the digit 1, yielding large incorrect numbers that must be separated accordingly. Thus, one has to scrutinize the data returned by the automatic system to detect these cases and correct them. This may be an additional cause of error in data. For data in 1994 and after, the WRDC has set up a Java-based interface (wrdc.mgo.rssi.ru) which is very convenient to display the data but does not allow downloading data. Consequently, data have to be copied by hand or other solutions such as optical character recognition which are not fully satisfactory. Whatever the solution, it requests manual handling of numbers which may be another source of error. Getting data from the WRDC has some risk and burden. However, it is a much better situation than if one has to request data separately to each NMS.

Several stations belong to the Global Atmosphere Watch (GAW) program of the WMO and exhibit good quality, including a quality flag. Data may be downloaded from the Web site (wrdc.mgo.rssi.ru) in HTML format. Other stations belong to the Baseline Solar Radiation Network (BSRN) and exhibit high to good quality, including quality flags. Data may be downloaded from the Web site (wrdc.mgo.rssi.ru).

In the present work, only ground stations that are known to the authors for the reliability of their measurements were kept.

In addition, the authors have also collected data set from the PIRATA network of buoys in the Tropical Atlantic Ocean [7]. These sites do not suffer any orographic effect and have been selected for the sake of the demonstration. Corrections are brought to measurements to take into account exposure of instruments to elements such as sea-spray, natural and anthropogenic aerosols. The corrected PIRATA data sets have been downloaded from the Pacific Marine Environmental Laboratory (PMEL) of the National Oceanic and Atmospheric Administration (NOAA) of the USA (www.pmel.noaa.gov/tao/data_deliv). Pyranometers installed on the buoys may experience accumulation of African dust, potentially leading to significant underestimation of the SSI [13]. These authors indicate that good quality is offered by buoys located in the area $-10°$ to $4°$. Only these buoys were kept in this study.

Only stations having more than 1000 days of valid measurements were kept in this study. 135 stations were studied.

2.5. Method for comparison

Comparison was carried on the SSI and the clearness index (KT). If E denotes the daily SSI and $E0$ denotes the daily irradiation received on a horizontal surface at the top of atmosphere, KT is defined as:

$$\mathrm{KT} = E\ /\ E0 \tag{1}$$

The clearness index is also called global transmissivity of the atmosphere, or atmospheric transmittance, or atmospheric transmission. The greater KT, the clearer the atmosphere. Values of KT around 0.7 denote clear sky conditions. The changes in solar radiation at the top of the atmosphere due to changes in geometry, namely the daily course of the sun and seasonal effects, are usually well reproduced by models and lead to a de facto correlation between observations and estimates hiding potential weakness of a model. KT is a stricter indicator of the performances of a model regarding its ability to estimate the optical state of the atmosphere. E0 is a function of the day in the year and of the solar constant which is the mean yearly value of the solar radiation received by a plane normal to the sun rays located at the top of atmosphere. E0 in Eq. (1) is estimated by the model of Ref. [5]. The solar constant in this model is 1367 W m^{-2}, equal to that used in HelioClim-1 and very close to those in MERRA (1365 W m^{-2}) and ERA-Interim (1370 W m^{-2}).

The deviations were computed by subtracting measurements from re-analyses estimations and HelioClim-1. These deviations are summarized by the bias, i.e. the mean value of the deviations, the standard-deviation, the root mean square error (RMSE), and the correlation coefficient. Relative bias and RMSE are also computed by dividing the bias and the RMSE by the mean value of the observations for the station under concern. The deviations are computed separately for each re-analysis and HelioClim-1. Consequently, the number of samples and the mean of the observations may vary slightly.

3. Results

Figs. 1–4 are examples of scatter density plots between the in situ measurements and the re-analyses estimates for the SSI and KT, respectively, for Mersa Matruh in Egypt and Maputo in Mozambique. These stations were selected for their contrast in climate. Mersa Matruh is on the Mediterranean coast and experiences a rather Mediterranean climate with a mild rainy boreal winter and a dry, warm and rainless summer. The soil is generally sandy. The sky is very clear in boreal summer: KT is larger than 0.65. Maputo (Mozambique) is located on the Southeast coast of Africa, in the Limpopo plain. The climate is

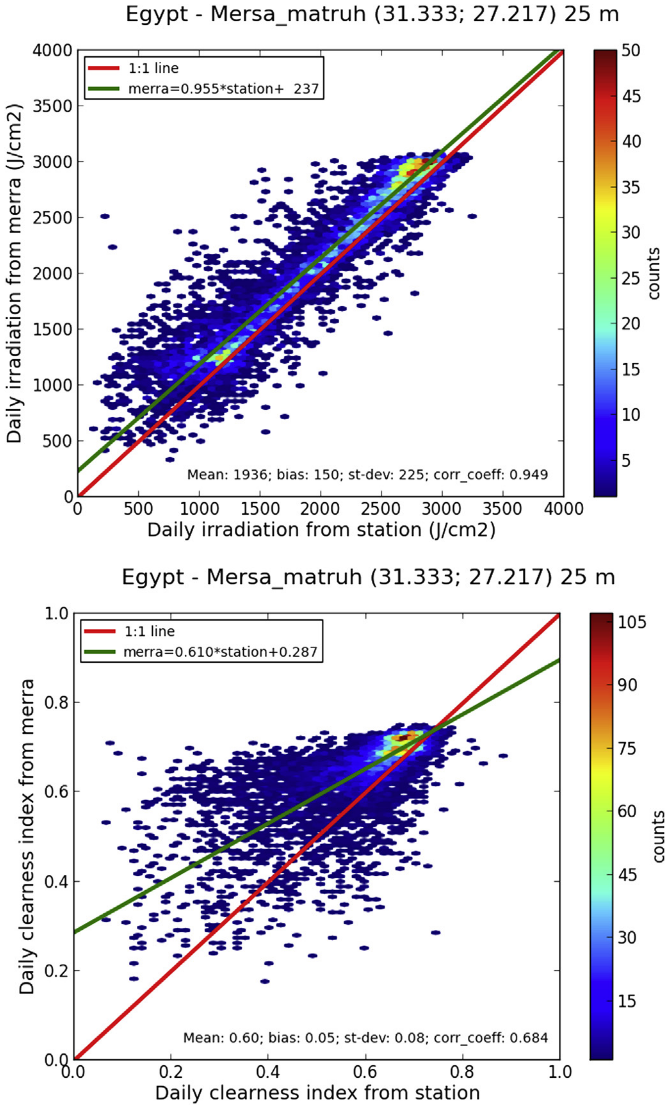

Fig. 1. Scatter density plot between measurements and MERRA estimates for Mersa Matruh. Daily irradiation (top), daily clearness index (bottom).

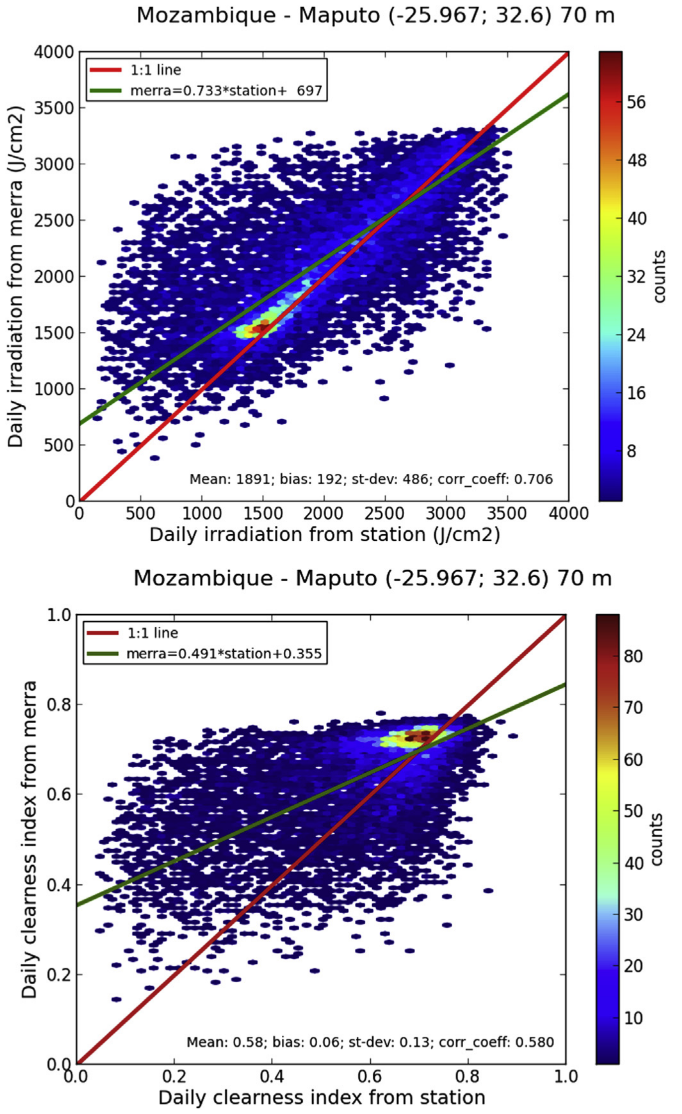

Fig. 2. Scatter density plot between measurements and MERRA estimates for Maputo. Daily irradiation (top), daily clearness index (bottom).

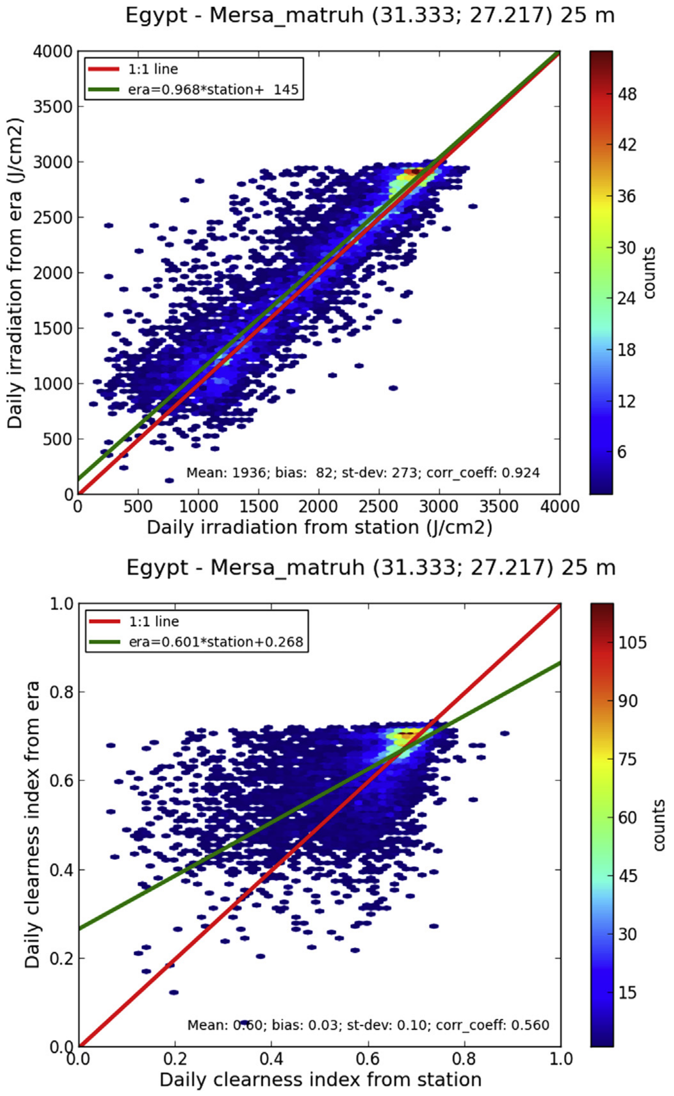

Fig. 3. Scatter density plot between measurements and ERA-Interim estimates for Mersa Matruh. Daily irradiation (top), daily clearness index (bottom).

rainy with no winter and the driest months are in austral winter (June–August). It is characterized by KT almost constant during the year, ranging between 0.53 and 0.59, with a maximum in austral winter.

In Fig. 1 (left), the cloud of points for the daily irradiation is elongated along a line which is fairly close to the 1:1 line. The correlation coefficient is large: 0.95. Most of the points lie above the 1:1 line, denoting an overestimation of SSI by MERRA. Bias is 150 J cm^{-2}, i.e. 8% of the mean observed value. The cloud of points is scattered. The standard-deviation is large: 225 J cm^{-2}, i.e. 12% of the mean observed value. One notes that the extreme irradiation for clear sky is not accurately reproduced.

Fig. 1 (right) for daily clearness index clearly shows a large discrepancy between MERRA and the measurements. The cloud of points does not follow the 1:1 line. The correlation coefficient is 0.68, meaning that only 46% of the variance, i.e. the quantity of information contained in observations, is explained by MERRA. MERRA overestimates KT with a bias of 0.05 (8%). MERRA overestimates KT for almost all values of KT but underestimates the greatest KT. The occurrence of KT comprised between 0.6 and 0.7 is much greater in MERRA than in observations; MERRA overestimates the frequency of clear sky conditions. The standard deviation is large: 0.08 (13%).

The performance of MERRA is worse for Maputo (Fig. 2) than for Mersa Matruh. The points are not elongated along the 1:1 line. There is a clear overestimation of both the daily irradiation (Fig. 2 left) and the clearness index (Fig. 2 right) though highest values in SSI and KT are underestimated. The bias is respectively 192 J cm^{-2}, i.e. 10% of the mean observed value, and 0.06 (10%). The standard-deviation is respectively 486 J cm^{-2} (26%) and 0.13 (22%), denoting the large scattering of the points. The correlation coefficient is low: 0.71 and 0.58. It means that MERRA explains only 50% and 34% of the variance in respectively daily irradiation and clearness index. Fig. 2 (right) shows that MERRA exhibits very often KT equal to 0.7 while observed KT are less than 0.6. MERRA often predicts clear sky conditions while actual conditions are cloudy. The opposite is also true though less pronounced: actual clear sky conditions are predicted as cloudy by MERRA.

Fig. 3 (left) exhibits the SSI for Mersa Matruh and ERA-Interim. The cloud of points for the daily irradiation in is elongated along a line which is very close to the 1:1 line. The correlation coefficient is large: 0.92. Many points lie above the 1:1 line, denoting an

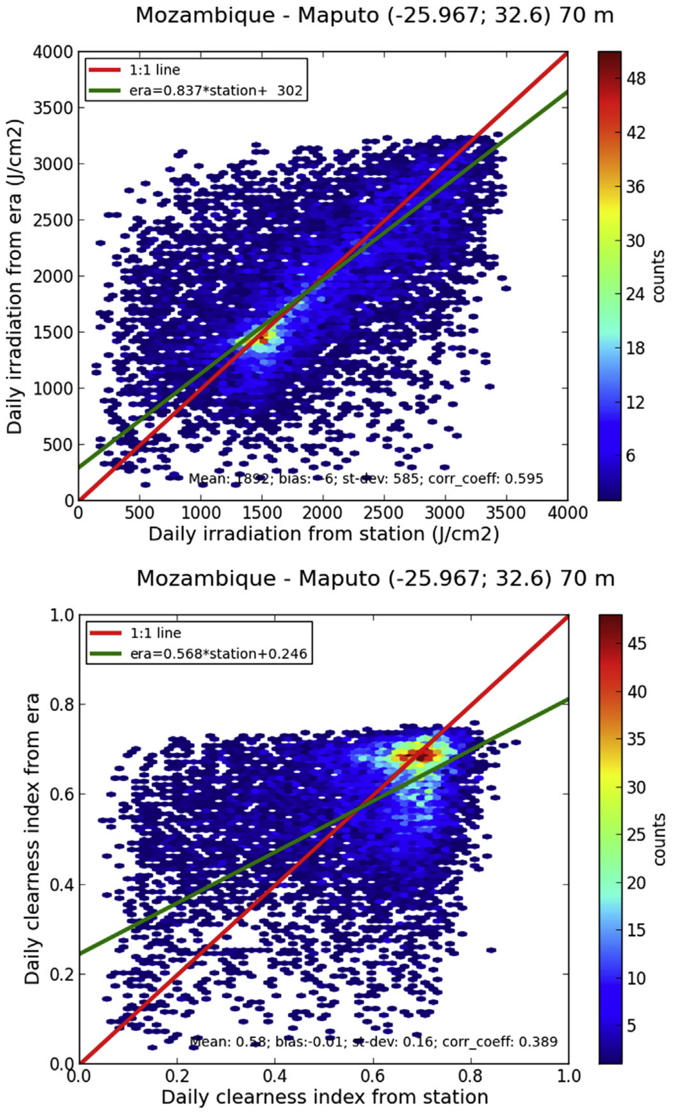

Fig. 4. Scatter density plot between measurements and ERA-Interim estimates for Maputo. Daily irradiation (top), daily clearness index (bottom).

overestimation. Bias is 82 J cm^{-2}, i.e. 4% of the mean observed value. The cloud of points is scattered with a large standard-deviation: 273 J cm^{-2}, i.e. 14% of the mean observed value. The greatest irradiations for clear sky are not accurately estimated by ERA-Interim. There is a large discrepancy in daily clearness index between ERA-Interim and the measurements (Fig. 3 right). Points are not located along the 1:1 line. The correlation coefficient is low: 0.56, meaning that only 31% of the variance is explained by ERA-Interim. ERA-Interim overestimates KT as a whole though it underestimates the greatest KT. The bias is 0.03 (5%). The standard deviation is large: 0.10 (17%). A striking feature in this graph is that ERA-Interim exhibits very often KT equal to 0.7 while observed KT are less than 0.7. ERA-Interim predicts clear sky conditions while actual conditions are cloudy.

Similarly to MERRA, the performance of ERA-Interim is worse for Maputo (Fig. 4) than for Mersa Matruh. Though the fitting line is fairly close to the 1:1 line for daily irradiation, the scattering of the points is very large for both irradiation and clearness index. The bias is small: respectively −6 J cm^{-2} (0%) and 0.01 (2%). The standard-deviation is very large: respectively 585 J cm^{-2} (31%) and 0.16 (28%). The correlation coefficient is low: 0.59 and 0.39. ERA-Interim explains only 35% and 15% of the variance in respectively SSI and KT. Fig. 4 (right) exhibits a striking feature: the scattering of points is almost rectangular which denotes a great level of uncertainty in the estimate of KT. Large KT are predicted by ERA-Interim while actual values are low. Conversely, low KT situations are predicted while actual values are large. In other words, ERA-Interim often predicts clear sky conditions while actual conditions are cloudy or cloudy conditions while the sky is clear. This overestimation in cloudy conditions compensates the underestimation in clear sky conditions, yielding a small bias overall.

As a whole, the re-analyses for the 135 stations exhibit similar behaviour with respect to ground measurements than the two examples of Mersa Matruh and Maputo, especially in clearness index. To better illustrate the findings and for the sake of clarity, a limited number of stations is discussed from now on. Six homogeneous climatic areas have been selected that offer a variety of the conditions encountered in Europe and Africa. Each contains a sufficient number of stations to assess whether the variability in error is less than that observed in HelioClim-1. These six climatic areas are: *1*) Baltic Area, *2*) France, *3*) Eastern Europe, *4*) North Africa, *5*) Mozambique, and *6*) Equatorial Atlantic Ocean. The 42 stations retained are listed in Table 1.

Tables 2 and 3 present the statistical results for MERRA and ERA-Interim for respectively the SSI and KT for the six areas. The results for HelioClim-1 have been added. For each climatic area, each table reports the mean of the observations for this area, the ranges of bias, relative bias, RMSE, relative RMSE and correlation coefficient observed for the set of stations located in the area. Figs. 5 and 6 exhibit the relative bias and relative RMSE for respectively the SSI and KT for the six areas and permit a visualisation of the differences between the data sets and the variability of performances within an area.

Except in few occasions, MERRA overestimates the SSI and KT. The RMSE is large in all cases, except for desert areas, when clear sky conditions prevail and as a consequence, the influence of false prediction of clear sky conditions by MERRA is of lesser importance. The correlation coefficient for daily irradiation is usually large and greater than 0.85. This is not the case at all for Mozambique or the Equatorial Atlantic where the correlation coefficient is much lower and the RMSE much greater. One may suspect the measurements. However, results of HelioClim-1 for these two areas are similar to the others. Consequently, there are other reasons for the large uncertainty of MERRA in these areas. MERRA often exhibits correlation coefficient less than 0.7 in KT (Table 3). This means that MERRA explains less than 50% of the variance in atmospheric transmissivity. A striking feature is the variability of the bias, RMSE and correlation coefficient within each homogeneous area. The worst cases are Mozambique and Equatorial Atlantic among those presented.

ERA-Interim tends also to overestimate the SSI and the KT but in a less pronounced manner than MERRA. The same features than those mentioned for MERRA can be observed for ERA-Interim. However, as a whole, the RMSE is greater for ERA-Interim than for MERRA (Fig. 6) and the correlation coefficient is lower. The errors in predicting cloudy situations are greater for ERA-Interim than for MERRA.

Finally, these tables report the performance for HelioClim-1 in order to situate those of MERRA and ERA-Interim. This is well illustrated in Figs. 5 and 6. Looking at these Figures and Tables 2 and 3, one may observe that as a whole, HelioClim-1 exhibits less bias, less RMSE and greater correlation coefficient than MERRA and ERA-Interim. In addition, HelioClim-1 offers less variability in uncertainty than MERRA and ERA-Interim within a given area. There are exceptions, such as Equatorial Tropical Ocean (area 6) where HelioClim-1 exhibits more bias than MERRA and ERA but less RMSE

Table 1
List of stations.

Area	Country	Name	Latitude N (deg)	Longitude E (deg)	Elevation a.s.l. (m)	Period
Baltic Area	Denmark	Copenhagen-Taastrup	55.67	12.30	28	1985–1993
	Latvia	Rucana	56.15	21.17	18	1994–2010
France	–	Auxerre	47.80	03.55	207	1985–1993
	–	Biscarosse	44.43	−01.25	33	1985–1993
	–	Carpentras	44.08	05.06	100	1985–2011
	–	La Roche sur Yon	46.70	−01.38	90	1985–1993
	–	Nice	43.65	07.20	4	1985–2010
	–	Strasbourg	48.55	07.63	153	1985–1993
Eastern Europe	Romania	Bucuresti	44.50	26.13	90	1985–1993
	Romania	Cluj Napoca	46.78	23.57	410	1985–1993
	Romania	Constanta	44.22	28.63	13	1985–1993
	Romania	Craiova	44.23	23.87	192	1985–1993
	Romania	Iasi	47.17	27.63	102	1985–1993
	Romania	Timisoara	45.77	21.25	86	1985–1993
	Ukraine	Kiev	50.40	30.45	179	1985–1992
	Ukraine	Odessa	46.48	30.63	64	1985–1992
North Africa	Algeria	Tamanrasset	22.78	05.52	1378	1995–2010
	Egypt	Aswan	23.97	32.78	192	1985–2009
	Egypt	Asyut	27.20	31.17	52	1985–2009
	Egypt	Cairo	30.08	31.28	33	1985–1998
	Egypt	El Arish	31.08	33.75	31	1986–2009
	Egypt	El Kharga	25.45	30.53	78	1985–1998
	Egypt	Mersa Matruh	31.33	27.22	25	1985–2009
	Egypt	Rafah	31.20	34.20	73	1994–1998
	Egypt	Sidi Barrani	31.62	25.90	24	1985–2008
	Egypt	Tahrir	30.65	30.70	16	1994–1998
	Tunisia	Sidi Bou Said	36.87	10.35	127	1985–1999
Mozambique	–	Beira	−19.80	34.90	10	1985–1997
	–	Chimoio	−19.12	33.47	731	1985–1995
	–	Chokwe	−24.52	33.00	33	1985–1998
	–	Maniquenique	−24.73	33.53	13	1985–1997
	–	Maputo	−25.97	32.60	70	1985–2010
	–	Pemba	−12.97	40.50	49	1985–1998
	–	Tete	−16.18	33.58	123	1985–1998
Equatorial Atlantic Ocean	–	Pirata1	00.00	00.00	0	1998–2011
	–	Pirata2	00.00	−10.00	0	1999–2011
	–	Pirata3	00.00	−23.00	0	1999–2011
	–	Pirata4	00.00	−35.00	0	1998–2011
	–	Pirata8	04.00	−38.00	0	1999–2011
		Pirata10	−06.00	−10.00	0	2000–2011
		Pirata14	−10.00	−10.00	0	1997–2011

and greater correlation coefficient. The variability in bias within an area is quite often greater for HelioClim-1 than for MERRA and ERA-Interim. This is not the case for RMSE, where there is no “best” data set.

4. Discussion

The overestimation of the SSI by re-analyses has been already documented. Refs. [25,26] found a tendency of a majority of re-analyses to overestimate the SSI and wrote that deficiencies in clear-sky radiative transfer calculations are major contributors to the excessive SSI.

Ref. [28] report on MERRA compared to ground daily irradiation measurements made in the U.S.A. MERRA exhibits a positive bias of 143 J cm^{-2} and a RMSE of 400 J cm^{-2}. The uncertainty is less for arid sites, i.e. where clear sky conditions prevail, than in northern sites: the bias is 20–100 J cm^{-2} and 30–150 J cm^{-2} respectively, and the RMSE is 400–500 J cm^{-2} and 300–700 J cm^{-2}. Ref. [30] compare MERRA against the FLUXNET measurements in Canada and U.S.A. for monthly averages of daily irradiation. They report an overestimation of 175 J cm^{-2} for all sky conditions. The bias is larger for cloudy skies than for clear skies. These results are similar to the present findings. If all stations are merged for all sky conditions (Table 2), the overall bias is 119 J cm^{-2} for MERRA and 57 J cm^{-2} for ERA-Interim (respectively 0.04 and 0.02 for KT in Table 3). If only the clearest conditions are retained by selecting the largest KT (percentile 90) for each station, the bias for the clear sky conditions is less in SSI for MERRA (59 J cm^{-2}) and is negative for ERA-Interim (−33 J cm^{-2}). The bias for KT is 0.00 for MERRA and slightly negative for ERA-Interim: −0.02. This can be seen in Figs. 1–4 where one may also note that the scatter for large values of the horizontal axis is limited.

Ref. [16] find that MERRA captures the seasonal variations of the monthly means of cloud fraction (CF) observed in the Atmospheric Radiation Measurement (ARM) Climate Research Facility in U.S.A. but with a negative bias and a low correlation (0.78). A negative bias in CF means overestimation in SSI and KT. Similarly to the present study, the bias is smaller in clear-sky conditions: approx. 35 J cm^{-2} versus 160 J cm^{-2} in all sky conditions. Ref. [8] report an underestimation of the cloud fraction by both MERRA and ERA-Interim in the Arctic region, yielding overestimation of the SSI. Ref. [6] analyse the global energy and water budgets in MERRA. They conclude that on a global scale, cloud effects in MERRA may be generally weak, leading to excess shortwave radiation reaching the ocean surface.

The present study does not study the cloud fraction but KT. It evidences that the re-analyses often predict clear sky conditions while actual conditions are cloudy. The opposite is also true though less pronounced: actual clear sky conditions are predicted as cloudy. Deficiencies by MERRA and ERA-Interim in prediction of the cloud amount would explain the low

Table 2
Comparison between MERRA, ERA-Interim, HelioClim-1 and in situ measurements. Relative values are expressed relatively to the mean observed value. Daily irradiation (J cm^{-2}).

		MERRA	ERA-I	HelioClim-1
Baltic Area	Mean obs.	1179	1209	1247
	Bias	88–173	26–195	−12–41
	Relative bias	7–16 %	2–17 %	−1–3 %
	RMSE	332–367	524–572	243–264
	Relative RMSE	27–33 %	41–51 %	19–22 %
	Correl. coeff.	0.922–0.930	0.785–0.801	0.944–0.956
France	Mean obs.	1340	1352	1355
	Bias	81–240	−69–177	−150–46
	Relative bias	5–21 %	−4–14 %	−12–3 %
	RMSE	297–439	474–518	203–279
	Relative RMSE	19–36 %	31–43 %	15–23 %
	Correl. coeff	0.899–0.944	0.795–0.838	0.950–0.968
Eastern Europe	Mean obs.	1324	1338	1324
	Bias	−12–199	−70–162	−166–68
	Relative bias	−1–16 %	−5–−13 %	−12–5 %
	RMSE	317–389	456–534	245–362
	Relative RMSE	24–31 %	31–39 %	17–27 %
	Correl. coeff	0.892–0.914	0.794–0.848	0.925–0.960
North Africa	Mean obs.	2007	2006	2005
	Bias	52–283	26–230	−139–62
	Relative bias	2–15 %	1–12 %	−7–4 %
	RMSE	193–353	214–427	164–236
	Relative RMSE	9–20 %	10–25 %	8–14 %
	Correl. coeff	0.803–0.957	0.786–0.944	0.907–0.977
Mozambique	Mean obs.	2013	2014	2023
	Bias	42–345	−215–139	−294–5
	Relative bias	2–19 %	−10–8 %	−13–0 %
	RMSE	461–530	484–585	270–455
	Relative RMSE	21–29 %	22–31 %	14–21 %
	Correl. coeff	0.532–0.774	0.562–0.651	0.817–0.931
Equatorial Atlantic	Mean obs.	2022	2022	2015
	Bias	−80–126	−85–197	61–379
	Relative bias	−4–7 %	−4–11 %	3–21 %
	RMSE	364–496	391–538	254–463
	Relative RMSE	18–26 %	18–28 %	12–25 %
	Correl. coeff	0.274–0.563	0.234–0.569	0.823–0.927

Table 3
Comparison between MERRA, ERA-Interim, HelioClim-1 and in situ measurements. Relative values are expressed relatively to the mean observed value. Daily clearness index.

		MERRA	ERA-I	HelioClim-1
Baltic Area	Mean obs.	0.443	0.443	0.440
	Bias	0.034–0.053	0.004–0.059	−0.005–0.016
	Relative bias	7–12 %	1–14 %	−1–4 %
	RMSE	0.128–0.130	0.184–0.200	0.094–0.097
	Relative RMSE	28–30 %	40–47 %	21–23 %
	Correl. coeff	0.754–0.757	0.331–0.422	0.852–0.870
France	Mean obs.	0.479	0.481	0.484
	Bias	0.031–0.088	−0.024–0.064	−0.060–0.015
	Relative bias	6–21 %	−4–14 %	−13–3 %
	RMSE	0.107–0.152	0.175–0.191	0.077–0.105
	Relative RMSE	19–37 %	32–46 %	15–25 %
	Correl. coeff	0.706–0.818	0.378 to0.458	0.851–0.907
Eastern Europe	Mean obs.	0.482	0.485	0.483
	Bias	0.002–0.082	−0.020–0.063	−0.066–0.023
	Relative bias	0–18 %	−4–14 %	−13–5 %
	RMSE	0.128–0.154	0.175–0.196	0.097–0.142
	Relative RMSE	27–34 %	34–41 %	19–29 %
	Correl. coeff	0.628–0.712	0.331–0.469	0.761–0.869
North Africa	Mean obs.	0.620	0.620	0.620
	Bias	0.014–0.062	0.008–0.074	−0.048–0.018
	Relative bias	2–16 %	1–13 %	−8–3 %
	RMSE	0.060–0.118	0.073–0.149	0.053–0.081
	Relative RMSE	9–21 %	10–27 %	8–15 %
	Correl. coeff	0.530–0.712	0.277–0.564	0.631–0.877
Mozambique	Mean obs.	0.597	0.597	0.598
	Bias	0.015–0.103	−0.063–0.039	−0.083–0.004
	Relative bias	2–19 %	−10–7 %	−13–1 %
	RMSE	0.127–0.151	0.136–0.165	0.077–0.126
	Relative RMSE	20–28 %	22–28 %	13–20 %
	Correl. coeff	0.437–0.587	0.371–0.465	0.804–0.903
Equatorial Atlantic	Mean obs.	0.562	0.562	0.561
	Bias	−0.022–0.035	0.024–0.054	0.017–0.105
	Relative bias	−4–7 %	−4–11 %	3–21 %
	RMSE	0.101–0.138	0.107–0.150	0.070–0.129
	Relative RMSE	18–26 %	18–29 %	12–25 %
	Correl. coeff	0.278–0.535	0.169–473	0.792–0.925

correlation coefficient in KT and the high bias and RMSE in both SSI and KT.

Ref. [31] focus on the Arctic region where very few measuring stations are available. They report that ERA-Interim has a low bias. They found that the bias in MERRA is positive in one BSRN station (415 J cm^{-2}) and negative in the other one (−85 J cm^{-2}). No other work was found documenting the variability of the bias, root mean square error and correlation coefficient within an irradiation homogeneous area. The present study finds large variability in these quantities. Practically, it means that a post-processing correction of MERRA or ERA-Interim may not be simple. Ref. [30]. propose an empirical model exploiting 14 stations in Canada and U.S.A. for monthly averages of daily irradiation. This post-processing algorithm consists of empirical relationships between model bias, clearness index, and site elevation. Ten other stations in the same region are used for validation. As a whole, the bias is strongly reduced as well as the RMSE, without a degradation of the correlation coefficient. The two sites in Florida used for calibration exhibit a large increase in RMSE and a large decrease in correlation coefficient after correction. When dealing with annual mean of SSI, Ref. [30] reports correlation coefficient between MERRA and observations ranging from −0.50 to 0.95 before correction. The empirical correction has a tendency to improve them.

5. Conclusion

The major finding in this study is that the MERRA and ERA-Interim re-analyses often predict clear sky conditions while actual conditions are cloudy. The opposite is also true though less pronounced: actual clear sky conditions are predicted as cloudy. This overestimation of occurrence of clear sky conditions leads to an overestimation of the SSI and KT by MERRA. Overestimation is less pronounced for ERA-Interim. Indeed, actual cloud-free conditions may be predicted as cloudy conditions as well, i.e. yielding an underestimation, and this compensates the overestimation of cloud free conditions with a slight positive bias as a result. The squared correlation coefficient for clearness index ranges between 0.38 and 0.53, showing that a very large part of the variability in irradiation is not captured by the MERRA or ERA-Interim re-analyses. Finally, a striking feature is the variability of the bias, RMSE and correlation coefficient within a same area though each area is fairly homogeneous for the SSI.

In clear sky conditions MERRA, and to a lesser extent ERA-Interim, is fairly accurate though underestimation is observed. These re-analyses do not have the same accuracy than models taking into account the dynamics of aerosols such as McClear [19] because they use climatology of aerosols. This may not be important if one uses the re-analysis to obtain an overview of the SSI in clear sky conditions. MERRA is more accurate than ERA-Interim for other sky conditions and should be preferred. It can be noted that the radiative scheme used at ECMWF for forecasts has changed from that used in ERA-Interim for a better one in June 2007 (cycle 32R2 at the ECMWF Integrated Forecasting System) [21,23].

The present results bring more evidence on the overestimation observed in several re-analyses as discussed in the previous section, the dependency of this overestimation with the frequency of

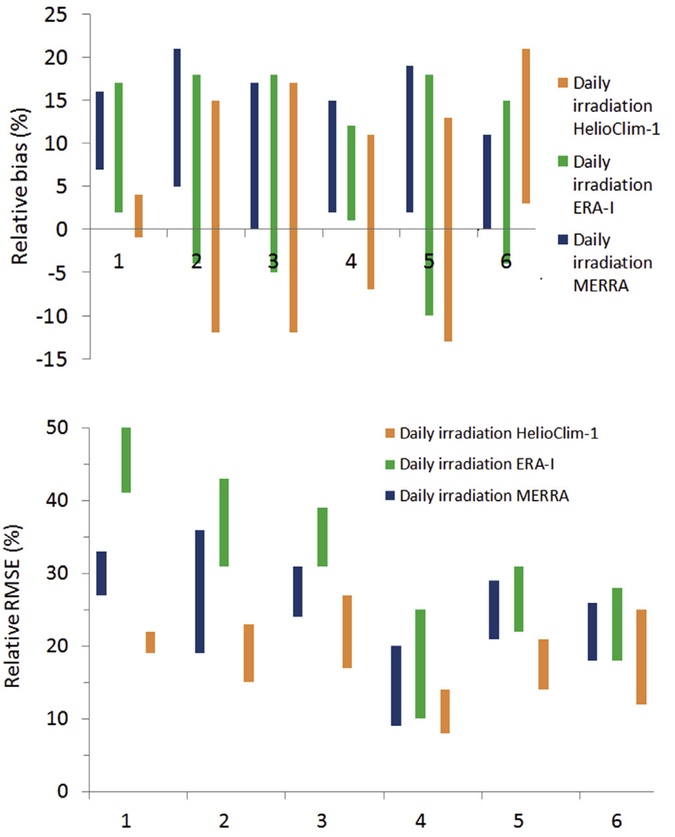

Fig. 5. Ranges of relative bias (top) and relative RMSE (bottom) for each climatic area and each database. Daily irradiation. Area *1*) Baltic Area, *2*) France, *3*) Eastern Europe, *4*) North Africa, *5*) Mozambique, and *6*) Equatorial Atlantic Ocean.

clear-sky conditions, and the necessity to take into account the dynamics of the aerosols. Cases such as Mozambique and Equatorial Atlantic Ocean have not been studied in literature relating to MERRA or ERA-Interim. The present study shows that these reanalyses present serious drawbacks for these areas in tropical humid climate. It is further observed that stations in Ghana and in Guiana also located in tropical humid climate exhibit similar results than those in Mozambique. The BSNR site at De Aar in South Africa is located in a dry climate and has similar results that those found in Northern Africa. The number of reliable data from stations collected in the present study is very limited in the tropical humid climate affecting Central Africa, Equatorial Atlantic Ocean, and Northern South America and no clear explanation may be provided on why these areas exhibit greater uncertainty that the areas in dry climate or rainy climate with mild winters.

To conclude, MERRA and ERA-Interim should only be used in solar energy with proper understanding of the limitations and uncertainties. In regions where clouds are rare, e.g. North Africa, MERRA or ERA-Interim may be used to provide a gross estimate of monthly or yearly irradiation. In all cases, a correction of MERRA or ERA-Interim by a function fitted on available ground measurements as shown in Figs. 1–4 will reduce the bias but not significantly the scattering of the estimates, i.e. the standard-deviation of the errors, and will not increase the correlation coefficient. There is no simple means to correct a posteriori for the errors made in MERRA or ERA-Interim mistaking cloudy hours as cloud-free ones. Uncertainty in MERRA or ERA-Interim is greater than that observed in HelioClim-1. If Europe or Africa is at stake, HelioClim-1 should be preferred to ERA-Interim and MERRA, though limited to the period 1985–2005.

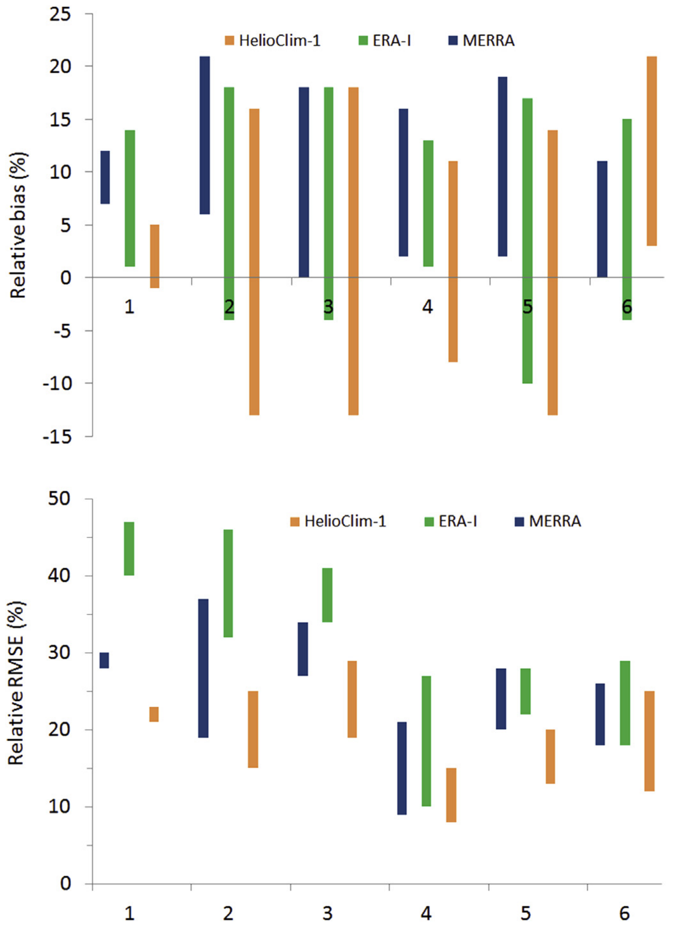

Fig. 6. Ranges of relative bias (top) and relative RMSE (bottom) for each climatic area and each database. Daily clearness index. Area *1*) Baltic Area, *2*) France, *3*) Eastern Europe, *4*) North Africa, *5*) Mozambique, and *6*) Equatorial Atlantic Ocean.

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 262892 (ENDORSE project). The authors thank all ground station operators of the WMO network, GAW, BSRN and PIRATA networks for their valuable measurements. They thank the World Radiation Data Centre, the National Renewable Energy Laboratory (USA), the Alfred-Wegener Institute and the PMEL for hosting the websites for downloading data. The authors are indebted to Francesca Di Giuseppe and Michael G. Bosilovich for their help in understanding ERA-Interim and MERRA respectively and improving the content of the paper. The authors thank the company Transvalor which is taking care of the SoDa Service for the common good, therefore permitting an efficient access to the HelioClim databases.

References

[1] Abdel Wahab M, El-Metwally M, Hassan R, Lefevre M, Oumbe A, Wald L. Assessing surface solar irradiance in Northern Africa desert climate and its long-term variations from Meteosat images. Int J Remote Sens 2009;31:261–80. http://dx.doi.org/10.1080/01431160902882645.

[2] Beyer H-G, Czeplak G, Terzenbach U, Wald L. Assessment of the method used to construct clearness index maps for the new European Solar Radiation Atlas (ESRA). Sol Energy 1997;61:389–97. http://dx.doi.org/10.1016/S0038-092X(97)00084-4.

[3] Blanc P, Gschwind B, Lefevre M, Wald L. The HelioClim project: surface solar irradiance data for climate applications. Remote Sens 2011a;3:343–61. http://dx.doi.org/10.3390/rs3020343.

[4] Blanc P, Gschwind B, Lefevre M, Wald L. Validating Meteosat-derived surface solar irradiance in Mozambique. In: Halounova L, editor. Proc. 30th EARSeL symposium. Thermal remote sensing; 2011. p. 561–8. held in Prague, Czech Republic, 30 May – 2 June 2011. Available at: http://www.earsel.org/symposia//2011-symposium-Prague/Proceedings/.

[5] Blanc P, Wald L. The SG2 algorithm for a fast and accurate computation of the position of the Sun. Sol Energy 2012;86:3072–83. http://dx.doi.org/10.1016/j.solener.2012.07.018.

[6] Bosilovich MG, Robertson FR, Chen J. Global energy and water budgets in MERRA. J Clim 2011;24:5721–39. http://dx.doi.org/10.1175/2011JCLI4175.1.

[7] Bourles B, Lumpkin R, McPhaden MJ, Hernandez F, Nobre P, Campos E, et al. The Pirata program: history, accomplishments, and future directions. Bull Am Meteorol Soc 2008;89:1111–25. http://dx.doi.org/10.1175/2008BAMS2462.1.

[8] Chernokuslky A, Mokhov II. Climatology of total cloudiness in the Arctic: an intercomparison of observations and reanalyses. Adv Meteorol 2012:15. http://dx.doi.org/10.1155/2012/542093. Article ID 542093.

[9] Chou M-D, Suarez MJ. A solar radiation parameterization for atmospheric studies. 1999 [online] Available from: http://ntrs.nasa.gov/search.jsp?R=19990060930 [accessed 21.12.13].

[10] Cros S, Mayer D, Wald L. The availability of irradiation data. Report IEA-PVPS T2-04: 2004. Vienna, Austria: International Energy Agency; 2004. p. 29.

[11] Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 2011;137:553–97. http://dx.doi.org/10.1002/qj.828.

[12] Fleming EL, Jackman CH, Considine DB, Stolarski RS. Sensitivity of tracers and a stratospheric aircraft perturbation to two-dimensional model transport variations. J Geophys Res 2001;106:14245–63.

[13] Foltz GR, Evan AT, Freitag HP, Brown S, McPhaden MJ. Dust accumulation biases in PIRATA shortwave radiation records. J Atmos Ocean Technol 2013;30:1414–32. http://dx.doi.org/10.1175/JTECH-D-12-00169.1.

[14] GEOSS. SoDa service releases 21 years of data on surface solar radiation. CORE; 2011. Available at: www.earthobservations.org/art_012_004.shtml.

[15] Jurus P, Eben K, Resler J, Krc P, Kasanicky I, Pelikan E, et al. Estimating climatological variability of solar energy production. Sol Energy 2013;98:255–64. http://dx.doi.org/10.1016/j.solener.2013.10.007.

[16] Kennedy AD, Dong X, Xi B, Xie S, Zhang Y, Chen J. A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data. J Clim 2011;24:4541–57. http://dx.doi.org/10.1175/2011JCLI3978.1.

[17] Lefevre M, Blanc P, Espinar B, Gschwind B, Menard L, Ranchin T, et al. The HelioClim-1 database of daily solar radiation at earth surface: an example of the benefits of GEOSS data-CORE. IEEE J Sel Top Appl Earth Obs Remote Sens 2014;7(5):1745–53. http://dx.doi.org/10.1109/JSTARS.2013.2283791.

[18] Lefevre M, Diabate L, Wald L. Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance. Sol Energy 2007;81:240–53. http://dx.doi.org/10.1016/j.solener.2006.03.008.

[19] Lefevre M, Oumbe A, Blanc P, Espinar B, Gschwind B, Qu Z, et al. McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmos Meas Tech 2013;6:2403–18. http://dx.doi.org/10.5194/amt-6-2403-2013.

[20] Lefevre M, Wald L. Solar radiation in Southern Africa: assessing the quality of the HelioClim-1 database. In: 9th EMS conference, Lodz, Poland, 10–14 September 2012; 2012.

[21] Morcrette J-J, Baker HW, Cole JNS, Iacono MJ, Pincus R. Impact of a new radiation package, McRad, in the ECMWF integrated forecasting system. Mon Weather Rev 2008;136:4773–98. http://dx.doi.org/10.1175/2008MWR2363.1.

[22] Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, et al. MERRA: NASA's modern-era retrospective analysis for research and applications. J Clim 2011;24:3624–48. http://dx.doi.org/10.1175/JCLI-D-11-00015.1.

[23] Troccoli A, Morcrette J-J. 2012: forecast assessment of surface solar radiation over Australia. Denver, CO, USA: World Renewable Energy Forum; May 2012. p. 13–8.

[24] Wald L, Blanc P, Lefevre M, Gschwind B. The performances of the HelioClim databases in Mozambique. In: Proceedings ISES solar world congress 2011, 28 August – 2 September 2011, Kassel, Germany. Resource assessment; 2011. p. 268–75.

[25] Wild M. Short-wave and long-wave surface radiation budgets in GCMs: a review based on the IPCC-AR4/CMIP3 models. Tellus 2008;60A:932–45. http://dx.doi.org/10.1111/j.1600-0870.2008.00342.x.

[26] Wild M, Long CN, Ohmura A. Evaluation of clear-sky solar fluxes in GCMs participating in AMIP and IPCC-AR4 from a surface perspective. J Geophys Res 2006;111:D01104. http://dx.doi.org/10.1029/2005JD006118.

[27] WMO. Guide to meteorological instruments and methods of observation. 7th ed. Geneva, Switzerland: World Meteorological Organization; 2008. WMO-No 8.

[28] Yi Y, Kimball J, Jones L, Reichle R, McDonald K. Evaluation of MERRA land surface estimates in preparation for the soil moisture active passive mission. J Clim 2011;24:3797–816. http://dx.doi.org/10.1175/2011JCLI4034.1.

[29] Zarzalejo L, Polo J, Martin L, Ramirez L, Espinar B. A new statistical approach for deriving global solar radiation from satellite images. Sol Energy 2009;83:480–4. http://dx.doi.org/10.1016/j.solener.2008.09.006.

[30] Zhao L, Lee X, Liu S. Correcting surface solar radiation of two data assimilation systems against FLUXNET observations in North America. J Geophys Res Atmos 2013;118:9552–64. http://dx.doi.org/10.1002/jgrd.50697.

[31] Zib BJ, Dong X, Xi B, Kennedy A. Evaluation and intercomparison of cloud fraction and radiative fluxes in recent reanalyses over the Arctic using BSRN surface observations. J Clim 2012;25:2291–305. http://dx.doi.org/10.1175/JCLI-D-11-00147.1.