Estimating downwelling solar irradiance at the surface of the tropical Atlantic Ocean: A comparison of PIRATA measurements against several re-analyses and satellite-derived data sets

Mélodie Trolliet[1], Jakub P. Walawender[2], Bernard Bourlès[3], Alexandre Boilley[4], Jörg Trentmann[2], Philippe Blanc[1], Mireille Lefèvre[1], Lucien Wald[1]

[1] MINES ParisTech, PSL Research University, O.I.E. - Center for Observation, Impacts, Energy, Sophia Antipolis, France
[2] Deutscher Wetterdienst, Offenbach, Germany
[3] IRD/LEGOS, Brest, France
[4] Transvalor, Mougins, France

Correspondence to: Mélodie Trolliet (melodie.trolliet@mines-paristech.fr)

Running title: M. Trolliet et al.: Comparing PIRATA measurements, re-analyses and satellite-derived data.

Abstract.

This paper assesses the merits and drawbacks of several data sets of solar downwelling radiation received at the horizontal surface of the tropical Atlantic Ocean where the magnitude of this radiation and its spatial and temporal variability are hardly known. The data sets are compared to quality-controlled measurements of hourly irradiance made at five buoys of the PIRATA network for the period 2012-2013. The data sets comprise the re-analyses MERRA-2 and ERA5, and three satellite-derived data sets: HelioClim-3v5, SARAH 2 and CAMS Radiation Service v2. The two re-analyses often report cloud-free conditions while actual conditions are cloudy and reciprocally, actual cloud-free conditions as cloudy. They exhibit more bias in irradiance in cases of medium and high level clouds than for low level clouds. The re-analyses poorly correlate with the clearness index derived from the measurements, a quantity that describes the optical state of the atmosphere: the correlation coefficient ranges between 0.49 and 0.56 for MERRA-2, and between 0.57 and 0.64 for ERA5. The irradiance pattern at both hourly and daily time scales is spatially distorted by re-analyses, especially for MERRA-2. The three satellite-derived data sets exhibit similar performances between them. They reproduce well the dynamics of the irradiance and clearness index at both hourly and daily time scales with correlation coefficients greater than 0.95 for irradiance and 0.80 for clearness index in most cases. They exhibit overestimation, with the lowest biases reached by CAMS Radiation Service v2 and ranging between 10 and 36 W m^{-2} depending on site. The bias of HelioClim-3v5 is almost the same for the five sites: the systematic error is fairly constant in space for this data set, meaning that the irradiance pattern may not be noticeably distorted and that the spatial gradients are well reproduced.

Introduction

Solar radiation reaching the ocean surface is an essential variable in the ocean-climate system (Budyko, 1969; Manabe, 1969; Siegel et al., 1995; Lean and Rind, 1998). The density of power received from the sun on a horizontal surface at sea level per unit surface is called the downwelling solar irradiance at the surface and is here abbreviated as DSIS. Other terms may be found in literature, such as solar exposure, solar insolation, solar flux, surface solar irradiance, downwelling shortwave flux or surface incoming shortwave radiation. The DSIS intensity is large over the tropical Atlantic Ocean and influences the sea surface temperature. The net downward surface energy is positive and accumulates within the ocean, resulting in a northward meridional transport of heat in the Atlantic Ocean (Liu et al., 2017). The DSIS influences the vertical structure of the ocean at more rapid time scales with local impacts on physics and plankton (Siegel et al., 1995).

Currently, the field of DSIS is hardly known in the Atlantic Ocean. One of the means of assessing the DSIS is to use measuring stations such as pyranometers aboard ship or on buoys (Cros et al., 2004). Such measurements are usually accurate though the stations are sparse. They cannot offer a synoptic view of the DSIS field. Images acquired by satellites observing the ocean surface are a second means of getting a synoptic view of the temporal variations of the DSIS field. For example, the series of geostationary Meteosat satellites offers synoptic views of the tropical and equatorial Atlantic Ocean every 15 min with a spatial resolution of between 3 and 5 km. Several data sets of DSIS have been constructed from these images, such as the HelioClim-3, SARAH-2 (Surface Radiation Data Set – Heliosat, version 2), and CAMS (Copernicus Atmosphere Monitoring Service) Radiation Service v2 data sets which are dealt with in this paper (see Section 1 for further details).

Re-analyses are a third means. They derive from weather forecast models used in a re-analysis mode to reproduce what was actually observed. They assimilate state variables such as temperature, moisture and wind. On the contrary, the DSIS is diagnostic i.e. it is derived from a radiative transfer model and depends on the representation of the whole set of radiatively active variables of the atmospheric column above the point. Hence, re-analysis estimates should not be mistaken with measurements of DSIS, because they include the uncertainty of the models. Of interest here are the ERA5 developed at the ECMWF (European Center for Medium-range Weather Forecasts) and MERRA-2 (Modern-Era Retrospective Analysis for Research and Applications, version 2) of the NASA (National Aeronautics and Space Administration) data sets.

Despite the fairly recent availability of gridded data sets, their use is spreading outside the climate community. However, for a more informed usage of these data in ocean sciences as a whole, greater validation effort is needed. This paper aims at establishing the merits and drawbacks of each of the five data sets, three satellite-derived data sets and two re-analyses data sets, when compared to quality-controlled hourly and daily measurements of DSIS recorded by the PIRATA (Prediction and Research Moored Array in the Tropical Atlantic) network of moored buoys in the tropical Atlantic Ocean, here considered as a reference. The data sets are briefly presented in Section 1. The protocol for validation is presented in Section 2. The merits and drawbacks of each data set are discussed in Section 3. The size of the grid cell is typically 3 km for satellite-derived data sets and 50 km for the re-analysis data; it is large compared to a single point and this difference is discussed in Section 3. How to access the data is described in the “data availability” section.

1 Data-sets

1.1 The PIRATA measurements

The PIRATA network comprises eighteen meteo-oceanic buoys (Atlas type, progressively replaced by T-FLEX systems from 2015; refer to https://www.pmel.noaa.gov/gtmba/pirata for more information) located in the Atlantic Ocean, between the latitudes 19° S and 21° N (Bourlès et al. 2008). Each PIRATA buoy is equipped with an Eppley pyranometer mounted at a height of 3.8 m that measures the DSIS. Values are recorded as 2 min averages. The sensors are deployed for about one year on average before replacement. Sensors are cleaned manually every trimester.

The measurements are subject to the same sources of uncertainty as their counterparts on firm ground, such as incorrect sensor levelling, shading caused by close structures, dust, dew, water-droplets, bird droppings, miscalibration of sensors, electronic failures, time shifts in data loggers, maintenance mishandling etc. (see e.g. Muneer and Fairooz, 2002). Some buoys experience on accumulation of African dust which potentially leads to a significant underestimation of the DSIS (Foltz et al., 2013). These authors have proposed corrections for the data from such buoys including corrections for sea-spray, natural and anthropogenic aerosols but limited to daily means of DSIS.

Pyranometers view a complete hemisphere and must be placed horizontally for accurate measurements. This is not the case within the PIRATA network where a pyranometer is affected by the motions of the buoy which change the portion of the sky seen, inducing errors in the measurements. The errors are very complex to estimate and correct (Katsaros and DeVault, 1986; MacWhorter and Weller, 1991). They depend on the relative sun-buoy geometry which may be expressed as the tilt angle, the angle between the plane of the pyranometer and the horizontal, and the difference in azimuth of the sun and tilt direction. This relative geometry is affected by wave action or strong surface current and depends on the time of the day, latitude and season. Since the downward radiation received from a portion of the sky depends on the sky conditions, the errors depend also on sky conditions. Errors are most apparent in conditions of high DSIS, in cloudless skies, and when solar zenithal angles are less than 60°. Katsaros and DeVault (1986) distinguished two main kinds of errors: the errors due to rocking motion caused by waves and the errors due to a mean tilt. The first one can be approached by the two following extreme cases: *(i)* the buoy motion is in the direction of the sun and *(ii)* the buoy motion is perpendicular to that direction. In the first situation, these authors expressed the error in irradiance measurement as a combination of losses produced by a motion away from the sun and gains by the tilting of the buoy toward the sun. By means of an analytical model and gross assumptions, they concluded that “the average error for a cycle of motion will not be zero but will not be large”. In the second situation, the effect of a perpendicular movement is always a loss, due to the loss of the sky portion seen by the pyranometer. Katsaros and DeVault (1986) calculated that the loss is of the order of 10 % in hourly mean of irradiance for 10° tilt and solar zenithal angle greater than 30°. For daily averages, the influence of the buoy movement is a combination of the two cases. As a consequence, compensating errors would often lead to smaller errors in measurement of daily means of irradiance. Wave action and a preferential tilt have the least effect in the tropics. However, diurnal variations in cloudiness, which are typical at low latitudes, make the compensating gains and losses uneven over the day, and therefore result in a larger net diurnal error than observed (Katsaros and DeVault, 1986). MacWhorter and Weller (1991) experimentally confirmed these calculations with simultaneous measurements of irradiance by gimbaled and un-gimballed pyranometers. Systematic tilts of 10° induced by strong surface currents or strong wind currents yield relative errors in excess of 40 %. Errors caused by wave motion are less severe and may amount to 10 %. Reynolds (2007) proposed an algorithm for correcting such errors. Inputs to this algorithm are the pitch, roll and heading of the sensor as well as the relative contributions of the beam and diffuse components of the global DSIS. Long et al. (2010) suggested using a combination of a specific pyranometer and algorithm to achieve an accuracy of 10 W m^{-2} in 90 % of the cases.

Currently, no correction is made to PIRATA measurements for errors due to buoy tilt or soiling. Measurements of 2 min DSIS for the period 2004-2016 were downloaded from the NOAA (National Oceanic and Atmospheric Administration) PMEL (Pacific Marine Environmental Laboratory) web site. Quality flags are provided together with the measurements. The NOAA procedure for quality checking rejects non-plausible values, i.e. values exceeding 1400 W m^{-2}. If any DSIS value, mean, standard deviation, or maximum, reads 0, all values are set to missing for that day. Flags are also raised if sensor outputs are zero or full scale throughout the day, or if the daily mean of the DSIS is outside the interval [50, 325] W m^{-2}. In a third pass, a visual inspection and comparison with time series plots from neighboring sites are performed.

An additional quality control was performed at MINES ParisTech on top of the NOAA screening since the PIRATA measurements serve as a reference in this comparison. The quality control used here is that of Korany et al. (2016) and comprises several tests of the 2 min DSIS data against extremely rare limits and physically possible limits. Values falling outside the limits were excluded from the time-series. Eventually, a visual analysis was performed to further remove suspicious values. A noticeable fraction of the data was removed. Only measurements that successfully passed all tests were kept. The hourly mean of DSIS was computed by averaging the 30 measurements within the hour only if all measurements were declared valid. Otherwise, the hourly mean of DSIS was excluded.

Following the recommendations by Foltz et al. (2013), the buoys located between 4° N and 8° N were discarded because of the possible large occurrence of significant tilt due to currents, and those located north of 8° N were discarded because of contamination by African dust. A further constraint in this study was the availability of enough measurements at each buoy without major gaps in a year in order to have an accurate description of the intra-year variability. In addition, the period for which data sets overlap start in 2010 as ERA5 was only available for the period 2010-2016 at the time of writing.

Eventually, five buoys on the eighteen PIRATA buoys have been selected: 0n0e, 0n10w, 0n23w, 6s10w and 19s34w (Fig. 1). These buoys have enough hourly means of DSIS for the period 2012-2013 to guarantee robust comparisons (Table 1). An additional ensemble of data was used to further controlling results and analyses: the year 2011 for the buoy 19s34w who offered approximately 4200 measurements, and the years 2010 and 2011 for the buoys 0n10w and 6s10w.

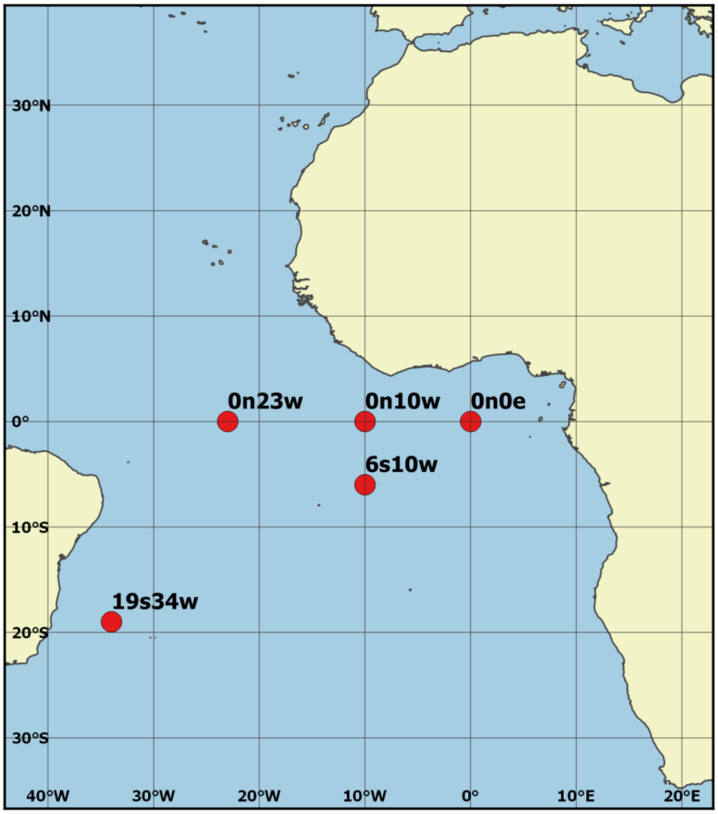


Figure 1: Map showing the location of the five PIRATA buoys used in this study.

Buoy	Latitude (positive North)	Longitude (positive East)	Number of valid hourly values	Hourly mean of DSIS (W m^{-2})	Hourly mean clearness index	Daily mean of DSIS (W m^{-2})	Daily mean clearness index
0n0e	0.0	0.0	8356	449	0.48	215	0.52
0n10w	0.0	-10.0	8417	480	0.52	232	0.56
0n23w	0.0	-23.0	7431	530	0.57	256	0.62
6s10w	-6.0	-10.0	8461	485	0.53	235	0.57
19s34w	-19.0	-34.0	8541	496	0.57	243	0.61

Table 1: Geographical coordinates of the PIRATA buoys used in this study, number of hourly values in each time-series, hourly and daily means of valid data in DSIS and clearness index for the period 2012-2013.

Table 1 reports the hourly and daily mean of DSIS as well as the means of the hourly and daily clearness indices (*KT*). *KT* characterizes the optical state of the atmosphere better than the DSIS. As a consequence, it allows comparing the modelling of the overall transparency of the atmosphere between models. Let *E* be the hourly mean of DSIS and *E0* the corresponding irradiance received on a horizontal plane located at the top of the atmosphere. The hourly clearness index *KT* is defined as the ratio of *E* to *E0*. *E0* was computed here using the SG2 algorithm (Blanc and Wald, 2012). Though *KT* is not completely independent of the position of the sun, the dependency of *KT* on the solar zenithal angle is much less pronounced than that of *E*. *KT* is typically close to 0.8 in cloud-free conditions, and close to 0.1 in overcast conditions with optically thick clouds.

The daily means of DSIS were computed by summing up the hourly means and dividing by 24 h by convention. The daily clearness index was also computed in the same way as the hourly *KT*. The means of the daily *KT* are greater than 0.5 denoting that the selected stations experience large occurrences of cloud-free conditions. Table 1 shows a tendency for an increase in *KT* from east to west.

1.2 The satellite-derived data set (HC3v5, SARAH-2, CRS)

HelioClim-3 v5 and CAMS Radiation Service v2 data sets, abbreviated respectively as HC3v5 and CRS, are constructed by processing images of the Meteosat Second Generation satellites by respectively the Heliosat-2 method (Rigollier et al., 2004; Lefèvre et al., 2007) modified by Qu et al. (2014), and the Heliosat 4 method (Qu et al., 2017). The SARAH-2 data record (Pfeifroth et al., 2017a) is obtained thanks to a retrieval approach based on Heliosat-2 (Müller et al., 2015; Pfeifroth et al., 2017b) exploiting the observations from Meteosat first and second generation. All data sets cover Europe, Africa, Middle East, parts of South America and Atlantic Ocean (full Meteosat disc).

HC3v5 and CRS are available from 2004 onwards with a 15 min time step. The spatial resolution depends on the pixel position and is approximately 3 km in the tropical Atlantic Ocean. Data can be accessed through a Web service at the SoDa Service (Gschwind et al., 2006, http://www.soda-pro.com/). This Web service performs the integration over time. It delivers the DSIS in all sky condition as well as in cloud-free conditions and the irradiance at the top of atmosphere *E0*. These three quantities were downloaded as hourly means and daily means for both HC3v5 and CRS for the locations of the selected buoys.

SARAH-2 is generated and distributed by EUMETSAT CM-SAF (Satellite Application Facility on Climate Monitoring). It provides information on the global and direct surface solar irradiance as well as the sunshine duration from 1983 to 2015. The data is provided on a regular grid with a grid spacing of 0.05° x 0.05° as instantaneous values of the DSIS every 30 min as well as daily and monthly averages. Here, the instantaneous values every 30 min were converted into instantaneous clearness index every 30 min. Assuming that *KT* is constant over 30 min, each instantaneous *KT* is multiplied by the corresponding *E0* integrated over 30 min, yielding 30 min irradiation. These 30 min irradiations are summed two by two to yield hourly irradiations, and then hourly means of irradiance.

1.3 The re-analysis data sets (MERRA-2, ERA5)

The MERRA-2 data set has many of the same basic features as the MERRA system (Rienecker at al., 2011) that has already been assessed against PIRATA daily means of DSIS by Boilley and Wald (2015), but includes a number of important updates (Gelaro et al., 2017). MERRA-2 offers 72 vertical levels of DSIS from ground to 0.01 hPa. The grid cell is 0.5°

(approx. 55 km) in latitude by 0.625° (approx. 71.5 km at the Equator) in longitude. MERRA-2 offers hourly means of DSIS starting from 1980.

ERA5 is the fifth generation of ECMWF atmospheric re-analyses of the global climate (Hersbach and Dee, 2016). It has several improvements compared to ERA. It has 137 levels from the surface up to 0.01 hPa. The size of the grid cell is 31 km. At the time of writing, the temporal coverage is 2010 to the present with 1 h time step. The period will be extended back to 1979 at the beginning of 2018.

The hourly means of DSIS in all sky conditions as well as in cloud-free conditions, and the irradiance at the top of atmosphere were downloaded from the MERRA web site for MERRA-2 and from the ECMWF MARS web site for ERA5. The time series for the location of each PIRATA buoy were constructed by firstly downloading the time series for the nearest four grid cells surrounding the buoy location, and then applying a spatial bilinear interpolation technique with a weighting factor that is inversely proportional to the distance to the PIRATA site. The daily means of DSIS were computed by summing up the hourly means and dividing by 24 h.

The characteristics of the five data sets are summarized in Table 2.

Data set		Start	End	Time resolution	Spatial resolution	Available from
Satellite-derived	HC3v5	2004	present	15 min	~ 3 x 3 km²	SoDa Service
	SARAH-2	1983	2015	30 min	~ 5 x 5 km²	EUMETSAT CM-SAF
	CRS	2004	present	15 min	~ 3 x 3 km²	SoDa Service
Re-analysis	MERRA-2	1980	present	1 h	55 x 71.5 km²	NASA
	ERA5	2010	present	1h	31 x 31 km²	ECMWF MARS

Table 2: Description of the data sets as on March 2018.

1.4 The CAMS cloud classification

In addition to these data sets, other variables have been downloaded to support the analyses of the errors for each data set. The CRS data set provides a classification of the clouds in four types as a function of altitude (Qu et al., 2017):

- low level cloud: water cloud at low altitude, with a base height of 1.5 km and a thickness of 1 km;
- medium level cloud: water cloud at medium altitude, with a base height of 4 km and a thickness of 2 km;
- high level cloud: deep cloud of large vertical extent from low altitude to medium altitude, with a base height of 2 km and a thickness of 6 km;
- thin ice cloud: ice cloud with a base height of 9 km and a thickness of 0.5 km.

A verbose mode is available in this service from which one may download these cloud types and other variables such as the fraction of pixel covered by cloud, solar zenithal angle and the aerosol optical properties.

2 Protocol for comparisons

The present work followed the protocol that was designed and is used in the framework of CAMS to perform quarterly validation of the CRS products against qualified ground measurements (see reports by Lefèvre and Wald at https://atmosphere.copernicus.eu/validation-supplementary-products). It comprises two parts.

The first part consists of the computation of differences between estimates and measurements. These differences are then summarized by classical statistical quantities. In this part, one more constraint applies on the PIRATA measurements: any measurement should be greater than a minimum significant value. This threshold is defined such that there is a 99.7 % chance that the irradiance is significantly different from 0 and that it can be used for the comparison. It is set to 30 W m^{-2}, i.e. 1.5 times the uncertainty (percentile 95) of measurements of good quality as reported by the WMO (World Meteorological Organization, 2012). Otherwise, the measurement, and therefore the corresponding estimate, is not included in the computation of the differences. Following the ISO standard (1995), the differences are computed by subtracting PIRATA measurements from the estimates. The set of differences is summarized by a few indicators namely the bias (mean of the differences), the standard deviation and the root mean square error. Relative values are computed from the mean of the corresponding PIRATA measurements at a given site. Correlation coefficients are computed. 2D histograms of PIRATA measurements and estimates, also called scatter density plots, are drawn as well as histograms of the differences.

Statistical properties of estimates and measurements are compared in the second part. Histograms of both the PIRATA measurements and the estimates are computed, and are superimposed into a single graph. Such graphs aim to assess the capability of a given data set to accurately reproduce the frequency distribution of the PIRATA measurements for the period. Monthly means and standard deviations within each calendar month of both the PIRATA measurements and the estimates are computed and displayed on a single graph.

In addition to the protocol for CRS validation, other graphs have been drawn to study the dependency of the statistical indicators on the irradiance or the clearness index, and other variables such as the month, year, solar zenithal angle, cloud types, cloud coverage, water vapour content, the aerosol optical properties or month. These graphs are not shown, except a few.

This enhanced protocol was applied to both *E* and *KT*. As for the performances of a model regarding its ability to estimate the optical state of the atmosphere, *KT* is a stricter indicator than *E* because it is less sensitive to changes induced by the daily and seasonal effects due to the Sun. These effects are usually well reproduced by models and lead to a de facto correlation between PIRATA measurements and estimates of *E*, hiding potential weaknesses of a model. The innovation of satellite-derived model and re-analyses is in the modelling of *KT* and not in *E0*, the latter being known from external astronomical models. Using the same *E0* for all data sets would create artificial distortions in results because of differences in *E0*. This is why *KT* was computed for each data sets using *E0* given by this data set. In any case, *E0* differs slightly from one data set to another by a few W m^{-2}, except for MERRA-2 as shown later.

The protocol was first applied to each data set for the five buoys for the period 2012-2013, using hourly values. In order to guarantee a better control and to support the conclusions, it was also applied to:

- the same data sample using daily values, with a threshold of 7.5 W m^{-2} instead of 30 W m^{-2},
- each data set for the buoys: 0n10w and 6s10w for 2010-2011, for both hourly and daily values,
- each data set for the buoy 19s34w in 2011, for both hourly and daily values.

The study has been conducted also on daily values for several reasons. One reason is that the performances may differ across these different time-scales. Another reason is that the daily values are the basis for constructed the monthly and yearly means, which are used in climatology, and thus it is important to assess the quality of the data sets at the daily time scale. In addition, dealing with daily values allows comparing the present results to already published works as it will be seen in Section “Discussion”.

It was found that the results are similar between the three satellite-derived data sets on the one hand, and between the two re-analyses on the other hand. Hence, the description of the results and the discussion are organized as follows. The three satellite-derived data sets are discussed altogether firstly for the hourly values, then for the daily values. Then, the two re-analyses are discussed for the hourly then daily values.

3 Results and discussion

3.1 The satellite-derived data sets (HC3v5, SARAH-2, CRS)

3.1.1 Hourly values

Table 3 shows the correlation coefficients, the biases and the standard deviations at each buoy, for the three satellite-derived data sets, for hourly means of E and KT (correlation coefficient only) in the period 2012-2013. These quantities are similar to those for 2010 and 2011 (not presented).

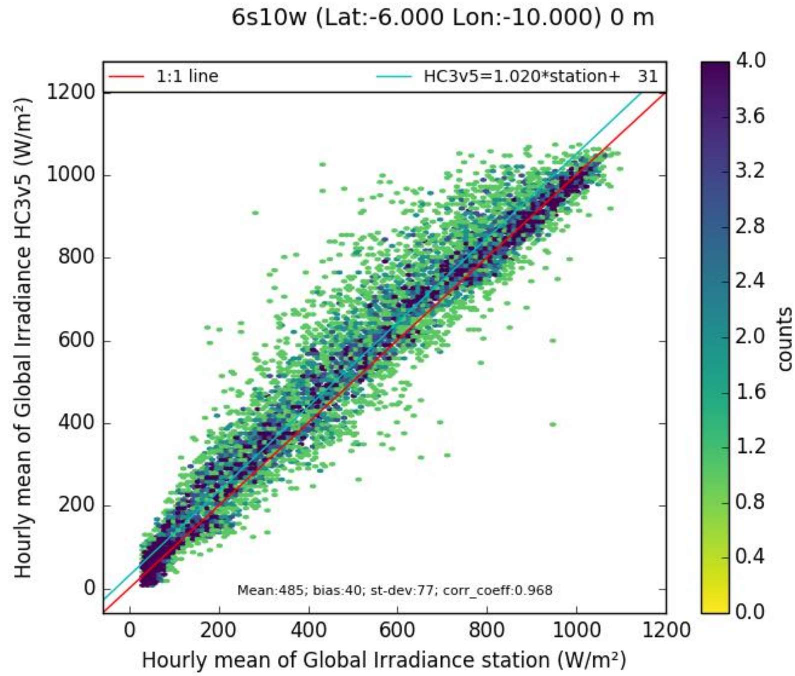
	Correlation coefficient			Bias			Standard deviation		
Buoy	**HC3v5**	**SARAH-2**	**CRS**	**HC3v5**	**SARAH-2**	**CRS**	**HC3v5**	**SARAH-2**	**CRS**
0n0e	0.964	0.970	0.965	48	55	31	80	72	77
	(0.865)	*(0.877)*	*(0.882)*	(10 %)	(12%)	(6 %)	(17 %)	(16 %)	(17 %)
0n10w	0.952	0.965	0.958	49	55	36	93	79	86
	(0.785)	*(0.827)*	*(0.823)*	(9 %)	(11 %)	(7 %)	(19 %)	(16 %)	(17 %)
0n23w	0.972	0.977	0.974	30	27	17	73	65	68
	(0.811)	*(0.832)*	*(0.829)*	(5 %)	(5 %)	(3 %)	(13 %)	(12 %)	(12 %)
6s10w	0.968	0.978	0.970	40	41	28	77	62	74
	(0.839)	*(0.879)*	*(0.875)*	(8 %)	(8 %)	(5 %)	(15 %)	(12 %)	(15 %)
19s34w	0.930	0.925	0.932	31	11	10	118	118	112
	(0.645)	*(0.627)*	*(0.662)*	(6 %)	(2 %)	(2 %)	(23 %)	(23 %)	(22 %)

Table 3: For each PIRATA buoy and each satellite-derived data set: Correlation coefficient between measurements and estimates from satellite-derived data sets for irradiance and clearness index *(in italics)*, bias and standard deviation ($W\ m^{-2}$) between measurements and estimates for irradiance with relative values (in brackets) for hourly means in the period 2012-2013.

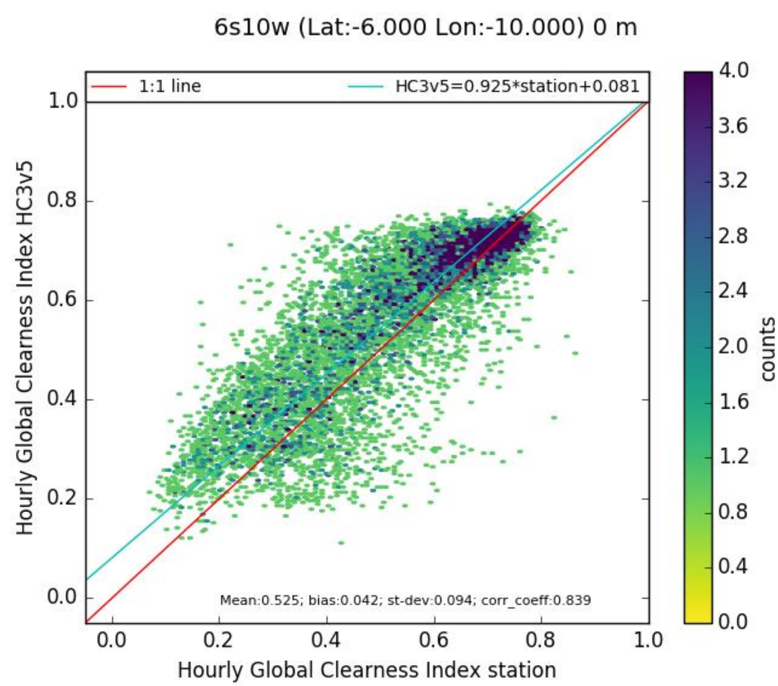
The three estimates correlate very well with the measurements. The correlation coefficients are very similar between data sets for both *E* and *KT*. They range between 0.93 and 0.98 for *E* (Table 3). They are slightly less for *KT* and range between 0.79 and 0.88, with lower values at 19s34w, respectively 0.65, 0.63 and 0.66. As a whole, the 19s34w buoy shows larger discrepancies in results than the others (Table 3). It could be partly explained by the finding of Foltz et al. (2013) who reported a significant bias at 19s34w compared to other moorings despite no apparent dust buildup.

The strong correlation is evidenced by the 2D histograms, illustrated in Figure 2 for 6s10w for the three data sets, which reveal well aligned distributions with small scattering for both *E* and *KT*. The 2D histograms for the other buoys, and more generally all plots, are quite similar to those of 6s10w presented here to illustrate, except at 19s34w. All plots are given in Appendix.

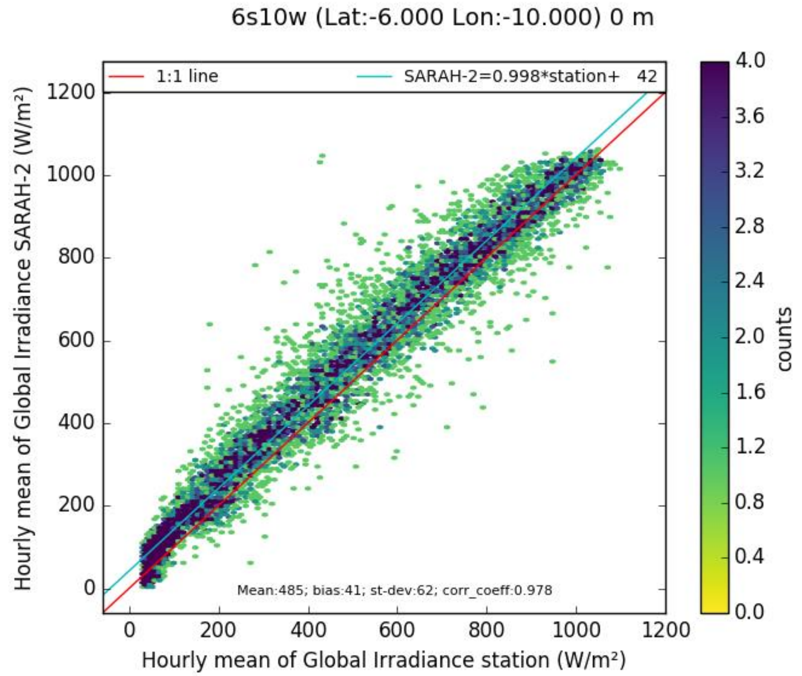
(a)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
HC3v5=1.020*station+ 31
Mean:485; bias:40; st-dev:77; corr_coeff:0.968
Hourly mean of Global Irradiance HC3v5 (W/m²)
Hourly mean of Global Irradiance station (W/m²)
counts
0
200
400
600
800
1000
1200
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

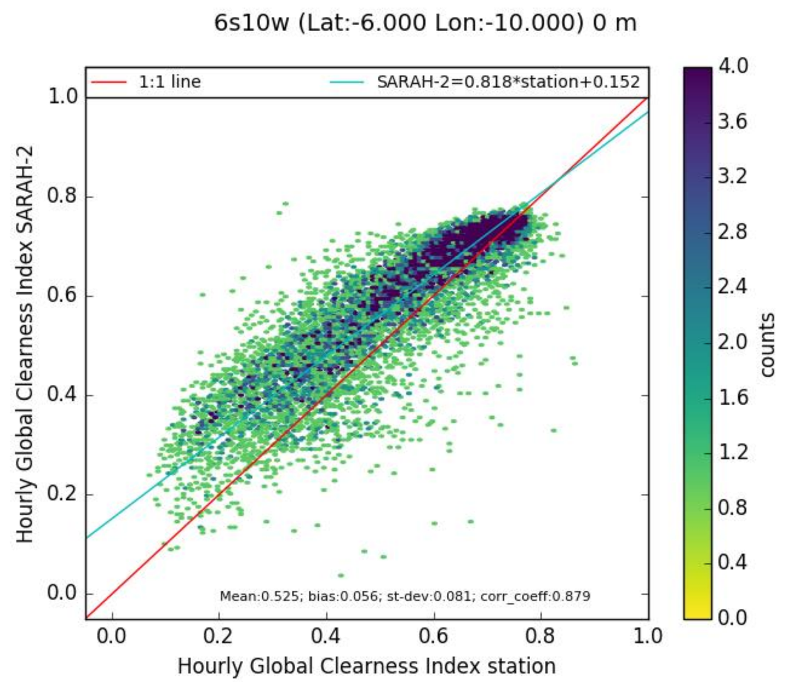
(b)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
HC3v5=0.925*station+0.081
Mean:0.525; bias:0.042; st-dev:0.094; corr_coeff:0.839
Hourly Global Clearness Index HC3v5
Hourly Global Clearness Index station
counts
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

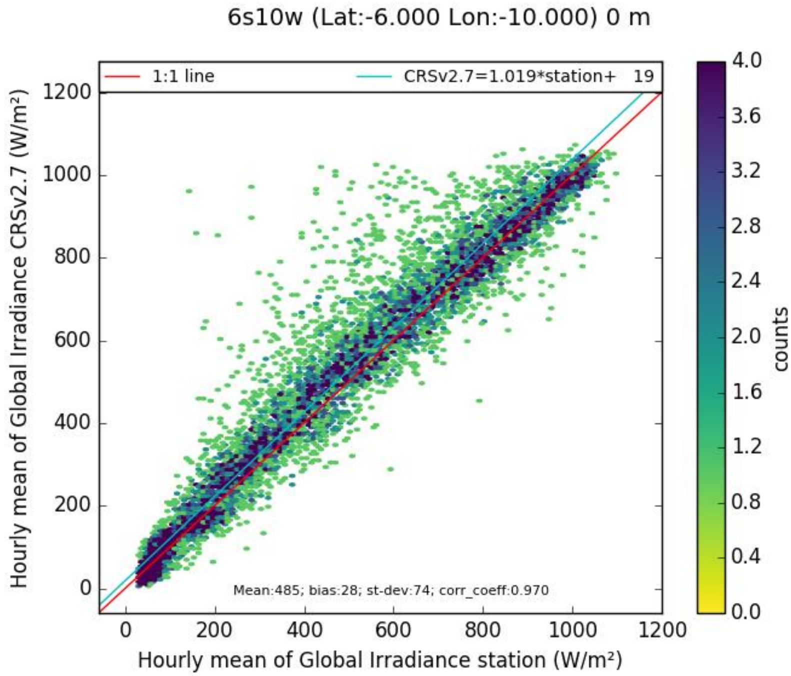
(c)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
SARAH-2=0.998*station+ 42
Hourly mean of Global Irradiance SARAH-2 (W/m²)
Hourly mean of Global Irradiance station (W/m²)
Mean:485; bias:41; st-dev:62; corr_coeff:0.978
counts
0
200
400
600
800
1000
1200
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

(d)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
SARAH-2=0.818*station+0.152
Hourly Global Clearness Index SARAH-2
Hourly Global Clearness Index station
Mean:0.525; bias:0.056; st-dev:0.081; corr_coeff:0.879
counts
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

(e)



(f)

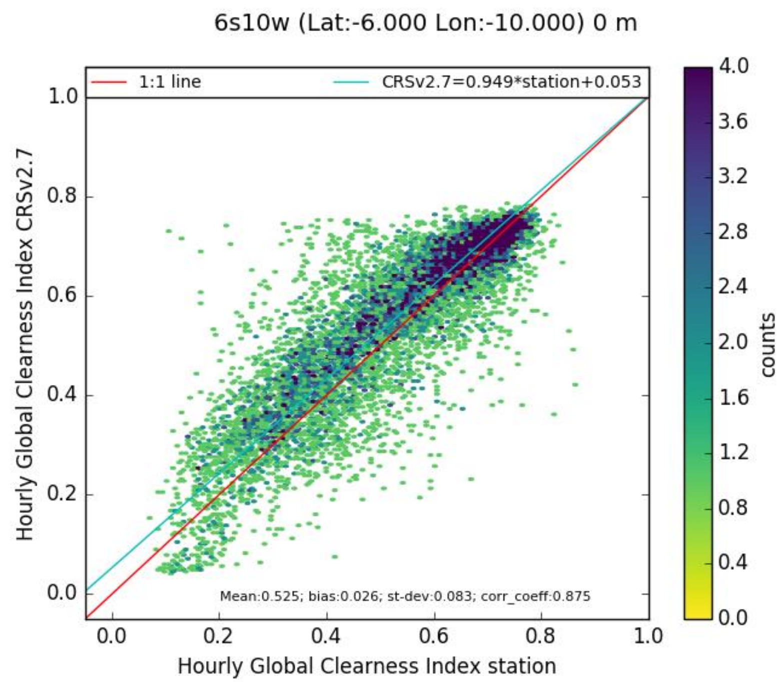


Figure 2: 2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at the buoy 6s10w for *E* and *KT*. HC3v5: (a) and (b); SARAH 2: (c) and (d); CRS: (e) and (f). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.

One may note an overall overestimation of *E* and *KT* for the three data sets in these graphs and Table 3. The bias in *E* is large and positive. Its range is slightly less for CRS than for the other data sets, with a range between 10 and 31 W m^{-2}, against a range between 30 and 49 W m^{-2} for HC3v5 and between 11 and 55 W m^{-2} for SARAH-2.

At a given buoy and a given data set, the errors may depend of the irradiance or clearness index, as shown in Fig. 2. The irradiances in the range [200, 800] W m^{-2} are overestimated while the smallest and greatest irradiances (>800 W m^{-2}) exhibit are correctly estimated. The greatest clearness indices are fairly well estimated in the three data sets. There is a tendency to an increasing overestimation with decreasing *KT*. The correct estimation of the greatest *E* and *KT* in HC3v5 and CRS can be related to the use of the McClear clear sky model that estimates the DSIS in cloud-free conditions. This model exploits the properties of the atmosphere delivered by CAMS. Several publications have underlined the good quality of the McClear estimates when compared to high-quality measurements performed at terrestrial stations (Eissa et al., 2015; Lefèvre et al., 2013; Lefèvre and Wald, 2016; Marchand et al., 2017). However, errors are possible in case of any gross errors in aerosol properties provided by CAMS. Other sources of errors found in cloud-free conditions originate from errors in the assessment of cloud properties (presence, coverage, optical depth).

The standard deviations of errors in *E* are fairly similar between the data sets (Table 3). SARAH-2 exhibits the smallest values, ranging from 69 to 79 W m^{-2}, with a much greater value at 19s34w (118 W m^{-2}). HC3v5 and CRS respectively exhibit ranges between 73 and 93 W m^{-2} (118 W m^{-2} at 19s34w), and 68 and 86 W m^{-2} (112 W m^{-2} at 19s34w). The scattering of errors in *KT* is greater than that in *E* for the three data sets (Fig. 2). There is no clear relationship between the standard deviation and the frequency of clouds or *KT* or the geographical location.

This first batch of results deals with pairs of coincident measurements and estimates. In the following, several statistical quantities (frequency distribution, monthly means and standard deviations) are computed on PIRATA data set and each satellite-derived data set independently and then compared.

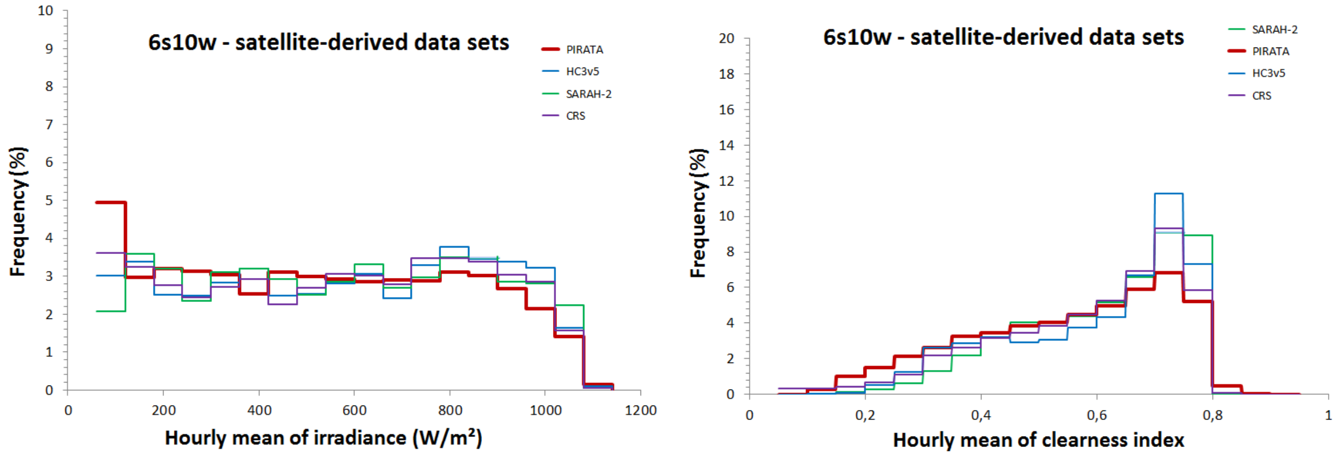


Figure 3: Frequency distributions of PIRATA measurements (red) and data sets (HC3v5: blue; SARAH 2: green; CRS: purple) for the station 6s10w for *E* (left) and *KT* (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

As a whole, the frequency distributions of estimates match well those of the measurements of *E* for each data set (Fig. 3 for the site 6s10w). The three data sets exhibit similarities. The frequencies for *E* less than 150 W m^{-2} are underestimated in the three data sets, this characteristic being more pronounced in SARAH-2. One may note an overestimation of the frequencies in the range [800, 1000] W m^{-2} which is slight for the three data sets and all buoys but more pronounced at 0n10w and 6s10w in CRS (see Appendix and Fig. 3). As for *KT*, the three data sets exhibit a more or less pronounced tendency to underestimate the frequencies for *KT* in the range [0.15, 0.5]. They also exhibit a noticeable overestimation of the frequencies for clearness indices greater than 0.6, i.e. there are too many cases of cloud-free conditions compared to the PIRATA measurements. Because this happens for the three data sets and at all sites, it is believed that this is mostly due to the physics in remote sensing and could be explained by the difficulty of accounting for the intra-pixel clouds.

The three satellite-derived data sets differ somehow between them regarding the estimation of the monthly means and standard deviation of *E* (see Fig. 4 for the site 6s10w). Each data set exhibits more or less the same performances at all sites. The monthly means of the estimated *E* from HC3v5 are very close to those from PIRATA in the period June-August (September-October at 19s34w) and overestimate those from PIRATA otherwise. As for SARAH-2, the estimated means overestimates those from PIRATA –quite often by large- all year long. The estimated means from CRS are very close to those from PIRATA from February to August, and overestimate them otherwise. At 19s34w, means from CRS estimates and PIRATA measurements are similar.

As for the standard deviation of *E* within a month, HC3v5 exhibits a tendency to overestimate those from PIRATA measurements though depending of the site there are periods in boreal summer where both are close. As a whole, Sara is close, with a tendency to overestimation in boreal winter at 0n0e and 0n10w. Standard deviations from CRS are very close from those from PIRATA, except in September-February when the former tend to overestimate the latter.

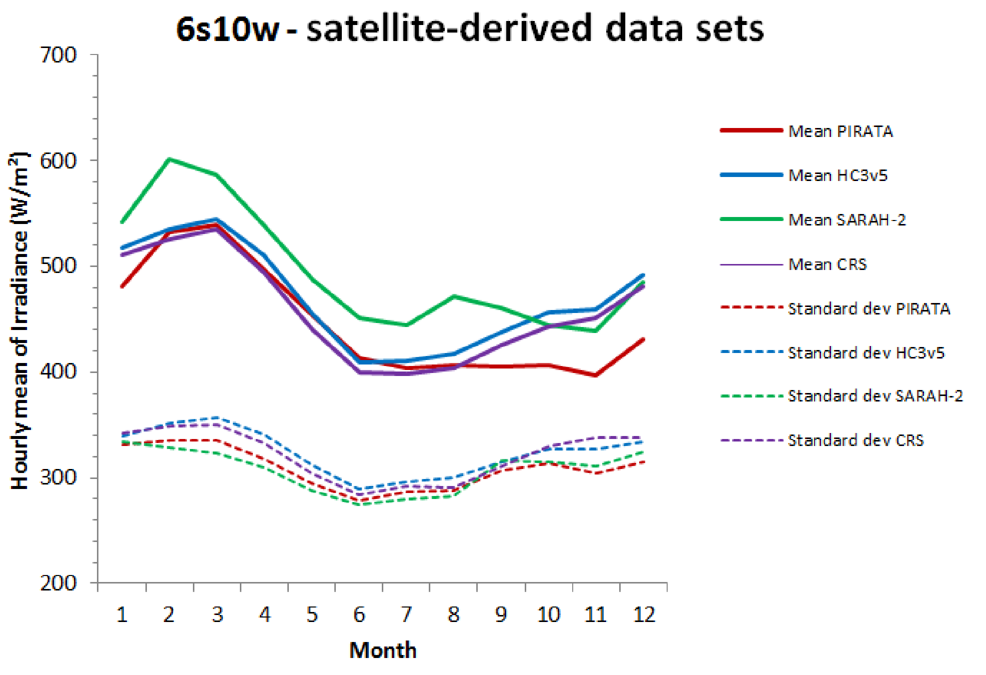


Figure 4: Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) at 6s10w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

3.1.2 Daily values

Table 4 shows the correlation coefficients, the biases and the standard deviations at each buoy for the three satellite-derived data sets, for daily means of *E* and *KT* (correlation coefficient only) in the period 2012-2013. These quantities are similar to those for 2010 and 2011 (not presented).

Buoy	Correlation coefficient HC3v5	Correlation coefficient SARAH-2	Correlation coefficient CRS	Bias HC3v5	Bias SARAH-2	Bias CRS	Standard deviation HC3v5	Standard deviation SARAH-2	Standard deviation CRS
0n0e	0.932	0.940	0.930	23	27	15	19	18	19
	(0.927)	*(0.935)*	*(0.925)*	(10 %)	(12 %)	(6 %)	(8 %)	(8 %)	(8 %)
0n10w	0.874	0.898	0.898	23	26	17	22	20	20
	(0.883)	*(0.904)*	*(0.906)*	(10 %)	(11 %)	(7 %)	(9 %)	(8 %)	(8 %)
0n23w	0.874	0.922	0.904	14	12	7	17	14	15
	(0.881)	*(0.926)*	*(0.908)*	(5 %)	(4 %)	(2 %)	(6 %)	(5 %)	(5 %)
6s10w	0.920	0.945	0.925	19	20	13	18	15	17
	(0.919)	*(0.943)*	*(0.929)*	(8 %)	(8 %)	(5 %)	(7 %)	(6 %)	(7 %)
19s34w	0.915	0.910	0.919	15	5	5	27	28	26
	(0.803)	*(0.795)*	*(0.813)*	(6 %)	(2 %)	(1 %)	(11 %)	(11 %)	(10 %)

Table 4: For each PIRATA buoy and each satellite-derived data set: Correlation coefficient between measurements and estimates from satellite-derived data sets for irradiance and clearness index *(in italics)*, bias and standard deviation (W m^{-2}) between measurements and estimates for irradiance with relative values (in brackets) for daily means in the period 2012-2013.

As expected, one may note that for a given data set and a given PIRATA site, the numbers in this Table are consistent between hourly and daily values: the relative biases are the same for the hourly and daily means, and the standard deviations of errors are less for daily values than for hourly values because of the averaging over the day (Tables 3, 4). The correlation coefficients for daily *E* are large and range between 0.93 and 0.98 for (Table 4).

The three estimates correlate very well with the measurements. The correlation coefficients are very similar between data sets for both *E* and *KT*. They range between 0.87 and 0.93 for HC3v5, 0.89 and 0.95 for SARAH-2, and 0.89 and 0.93 for CRS (Table 4). The correlation coefficients for daily *E* are less than for hourly *E* because of the strong influence of the solar zenithal angle on the correlation coefficient which creates a de facto correlation of hourly values between estimates and measurements. The correlation coefficients for daily *KT* are greater than for hourly *KT*, close to or greater than 0.9 for the three satellite-derived data sets (0.8 at 19s34w), and similar to those for the daily *E* (Table 4).

The results for daily values are close to the expectations given those for hourly values. As a consequence, the different 2D histograms, the frequency distributions and the monthly means and standard deviations of daily *E* and *KT* of PIRATA measurements and data sets are not discussed in detail. Only points of interest are exposed in the following paragraphs.

Contrary to the smallest hourly irradiances which are underestimated, the smallest daily irradiances (<300 W m^{-2}) are overestimated; otherwise they are well estimated (not shown). Compared to the PIRATA frequency distributions of daily *E* and *KT*, those in SARAH-2 are shifted towards the largest *E* and *KT*, except at 19s23w where the distributions are close. The three satellite-based data sets correctly estimate daily *E* and *KT* in cloud-free conditions though fairly cloudy conditions may be mistaken as cloud-free from time to time.

3.1.3 Discussions on the three satellite-derived data sets

The differences between the greatest and smallest bias are respectively 44 W m^{-2} and 26 W m^{-2} for SARAH-2 and CRS. The bias for both SARAH-2 and CRS shows a tendency to decrease from East to West though care should be taken given the small number of buoys. This could be related to the increase in the mean *KT* from East to West as shown in Table 1. For CRS, this tendency is in agreement with the CAMS validation results reported quarterly since 2014 by Lefèvre and Wald using terrestrial stations (https://atmosphere.copernicus.eu/validation-supplementary-products). Though it has not been discussed by these two authors, one may note that their reports indicate a tendency of the bias to decrease with an increase in the mean *KT*. The tendency is more visible at the terrestrial stations experiencing frequent cloud-free conditions similarly to the selected PIRATA buoys.

The bias for HC3v5 does not exhibit any trend and is almost the same at the five sites: the difference between the maximum and minimum is 19 W m^{-2}. This constant bias in space means that the spatial features of *E* are well reproduced by HC3v5 as a whole as the spatial variability, expressed e.g. by spatial gradients, is not artificially distorted by artefacts due to spatial variations in bias.

Thomas et al. (2016) have performed comparisons of hourly and daily measurements of *E* performed at a total of 42 Brazilian stations, i.e. at similar latitudes, against estimates from HC3v5 on the one hand, and CRS on the other hand. The reported performances are fairly similar to those of the present work for both hourly and daily *E*. One may note that the bias at terrestrial sites is closer to 0 than for the PIRATA buoys and that the standard deviations are a bit smaller. This may indicate some limitations in the accuracy of the PIRATA measurements.

The large correlation coefficients mean that the time-series of the actual field of DSIS are well reproduced by any of the three data sets at hourly and daily time scales though amplitudes of variation in time may be hampered by the large standard deviation of the errors (Tables 3 and 4). This finding is consistent with those of Bengulescu et al. (2017) who reported very high correlation coefficients between HC3v5 and in situ measurements at various temporal scales, from days to years. The same conclusion may be drawn for the clearness index.

The frequency distributions of *E* and *KT* are fairly well reproduced in each satellite-derived data set though one may note a noticeable overestimation of the number of cloud-free cases compared to the PIRATA measurements likely due to the difficulty of accounting for the intra-pixel clouds. The monthly standard deviations of *E* are also well reproduced while the situation on the monthly means is more mixed depending on the data set.

3.2 The re-analysis data sets (MERRA-2, ERA5)

3.2.1 Hourly values

Table 5 shows the correlation coefficients, the biases and the standard deviations at each buoy for the two re-analysis data sets, for hourly means of *E* and *KT* (correlation coefficient only) in the period 2012-2013. These quantities are similar to those for 2010 and 2011 (not presented).

	Correlation coefficient		Bias		Standard deviation	
Buoy	**MERRA-2**	**ERA5**	**MERRA-2**	**ERA5**	**MERRA-2**	**ERA5**
0n0e	0.821 *(0.486)*	0.875 *(0.606)*	-42 (-9 %)	25 (5 %)	166 (36 %)	141 (31 %)
0n10w	0.858 *(0.519)*	0.901 *(0.634)*	-17 (-3 %)	25 (5 %)	151 (31 %)	128 (26 %)
0n23w	0.906 *(0.562)*	0.929 *(0.641)*	-10 (-1 %)	-6 (-1 %)	129 (24 %)	112 (21 %)
6s10w	0.893 *(0.546)*	0.896 *(0.571)*	12 (2 %)	15 (3 %)	134 (27 %)	132 (27 %)
19s34w	0.910 *(0.560)*	0.927 *(0.640)*	23 (4 %)	-10 (-1 %)	128 (25 %)	115 (23 %)

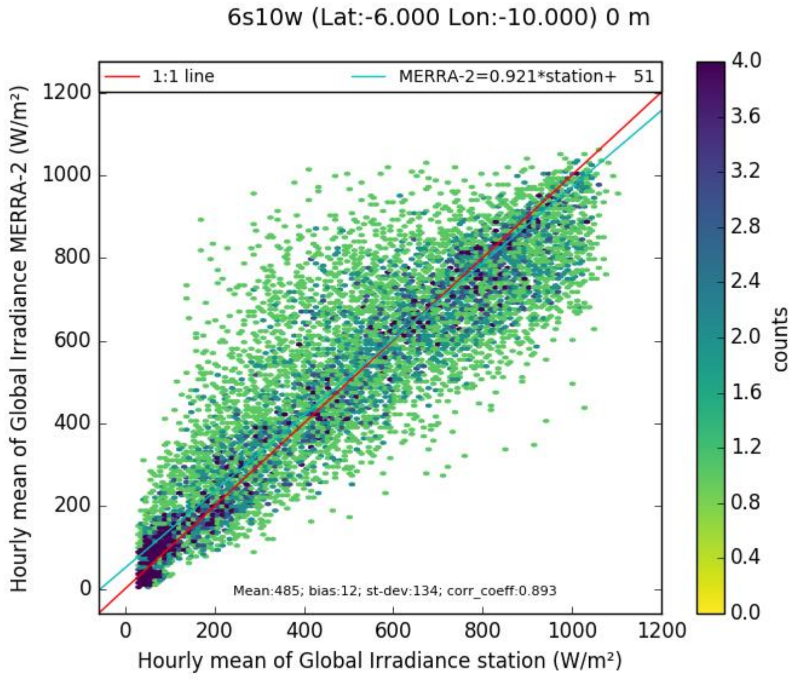
Table 5: For each PIRATA buoy and each re-analysis data set: Correlation coefficient between measurements and estimates from re-analysis for irradiance and clearness index *(in italics)*, bias and standard deviation (W m^{-2}) between measurements and estimates for irradiance with relative values (in brackets) for hourly means in the period 2012-2013.

The two estimates correlate well with the PIRATA measurements for *E* with correlation coefficients ranging between 0.82 and 0.91 for MERRA-2, and between 0.88 and 0.93 for ERA5. The correlation coefficients are weaker for *KT*: from 0.49 to 0.56 for MERRA-2, and from 0.57 to 0.64 for ERA5. It means that at most respectively 31 % and 41 % of the variance contained in the measured clearness indices is explained by MERRA-2 and ERA5. One may note that the correlation coefficients for MERRA-2 exhibit a tendency to increase from East to West, both in *E* and *KT*.

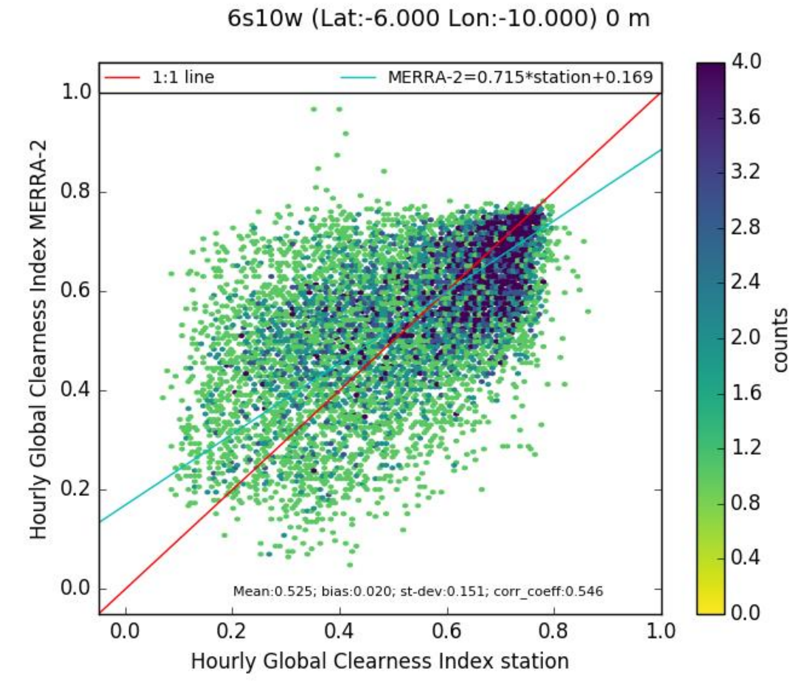
The strong correlation in *E* and the weak correlation in *KT* are evidenced in the 2D histograms, illustrated in Figure 5 for 6s10w for the two data sets. The 2D histograms at the other buoys, and more generally all plots, are quite similar to those of 6s10w presented here to illustrate, except at 19s34w. All plots are given in Appendix.

The 2D histograms show that the dots for *E* are fairly well aligned along the 1:1 line with a very large scattering (see e.g. Fig. 5a, c). One may note a large underestimation of the greatest DSIS, i.e. greater than 700-800 W m^{-2} for both data sets. As for *KT*, the 2D histograms for MERRA-2 show shapes that are not elongated and are more like a square or a triangle along lines whose slopes are less than 1 (Fig. 5b). The 2D histograms for ERA5 are more elongated along lines whose slopes are less than 1 (Fig. 5d) but still exhibit very large scattering.

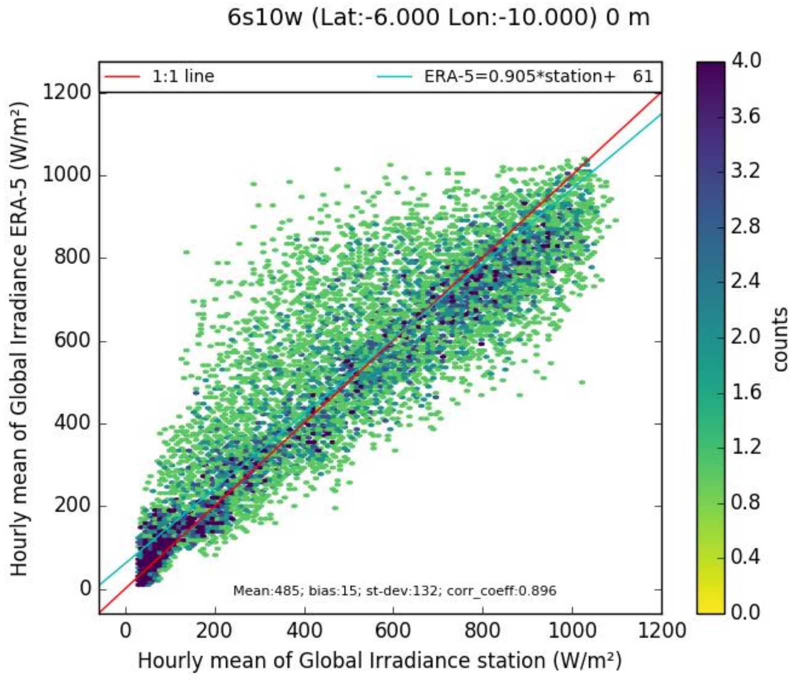
(a)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
MERRA-2=0.921*station+ 51
Mean:485; bias:12; st-dev:134; corr_coeff:0.893
Hourly mean of Global Irradiance MERRA-2 (W/m²)
Hourly mean of Global Irradiance station (W/m²)
counts
0
200
400
600
800
1000
1200
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

(b)


6s10w (Lat:-6.000 Lon:-10.000) 0 m
1:1 line
MERRA-2=0.715*station+0.169
Mean:0.525; bias:0.020; st-dev:0.151; corr_coeff:0.546
Hourly Global Clearness Index MERRA-2
Hourly Global Clearness Index station
counts
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
3.6
4.0

(c)



(d)

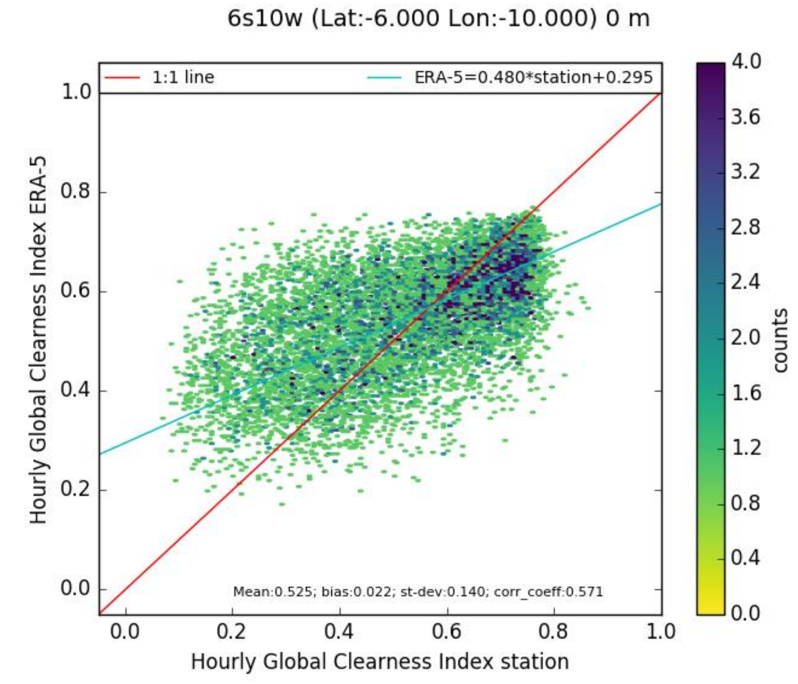


Figure 5: 2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at buoy 6s10w for *E* and *KT*. MERRA-2: (a) and (b); ERA5: (c) and (d). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.

One may ask if the underestimation of the greatest *E* originates from the model used in the re-analysis in cloud-free conditions. As MERRA-2 provides the hourly DSIS in cloud-free conditions, a comparison was made with the hourly DSIS from McClear. Both data sets reveal similar irradiances. It is concluded that the MERRA-2 cloud-free DSIS are likely accurate and that the underestimation of the greatest *E* is mostly due to errors in the assessment of cloud properties by MERRA-2. This is in agreement with the analysis of the 2D histograms for *KT* which exhibit a well-marked underestimation of the greatest clearness indices.

ERA5 does not provide estimates of the DSIS for cloud-free conditions contrary to MERRA-2. Hence, individual days have been selected randomly for which the daily profiles were similar to the expectations for cloud-free conditions based on a visual analysis. The daily profiles of the DSIS from McClear were superimposed. It was found that the DSIS for cloud-free conditions is underestimated by ERA5. This is supported by the analysis of the 2D histograms for *KT* (see e.g. Fig. 5d) which shows a tendency to overestimate for $KT<0.6$ and a well-marked underestimation for $KT>0.6$. These observations are clearly seen on the frequency distribution to be discussed later (Fig. 7).

The bias in *E* varies strongly with the site and with the data set. It ranges between -42 and 23 W m^{-2} for MERRA-2 and between -10 and 25 W m^{-2} for ERA5. The 2D histograms show a tendency for both data sets to overestimate the smallest irradiances and to underestimate the greatest ones. As for the clearness index, there is a tendency to overestimate *KT* whatever *KT* in both data sets but the situation is complex, especially in MERRA-2 where the error looks a bit random, i.e. there is as much chances of observing an overestimation of the actual clearness index as an underestimation.

The standard deviations of the errors for ERA5 are comprised between 112 (21 %) and 141 W m^{-2} (31 %). Those for MERRA-2 are greater: they range from 128 (25 %) to 166 W m^{-2} (36 %). Further exploration shows the dependency of the errors for MERRA-2 to the differences between the true solar time and the mean solar time. MERRA-2 does not account for this difference which is a function of the day in the year as a first approximation and is ranging from -17 to 17 min (Fig. 6, Wald, 2007). The true solar time is that needed for computing the solar zenithal angle accurately enough while it seems that MERRA-2 performs only a rough estimate of this angle by using the mean solar time. This angle intervenes twice: firstly to compute the irradiance impinging on the horizontal plane at the top of atmosphere and secondly as a major input to the radiative transfer model. Hence, an error in this angle yields an error in the estimated DSIS. This weakens the correlation between the PIRATA measurements and the MERRA-2 data set and it increases the standard deviation of errors for both hourly and daily values of the DSIS.

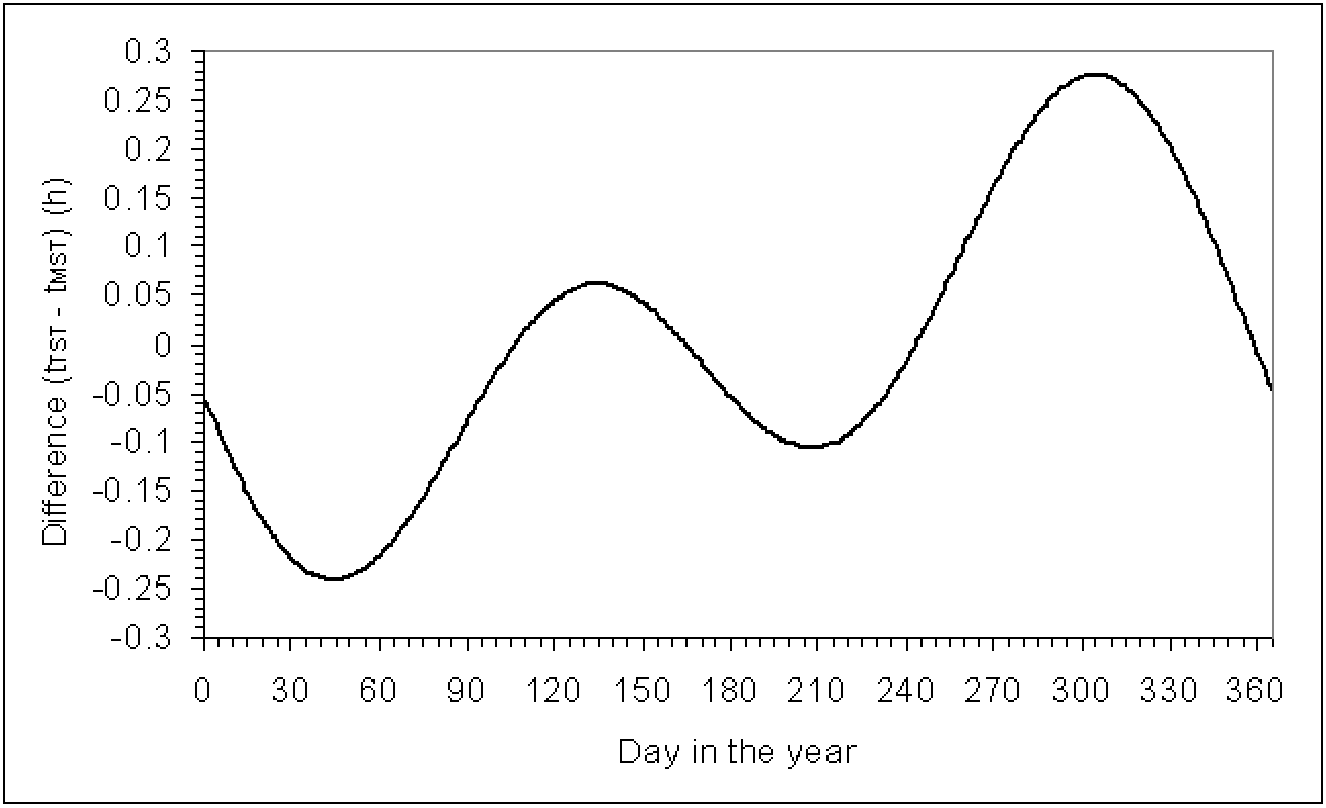


Figure 6: Difference between the true solar time (t_{TST}) and the mean solar time (t_{MST}). Excerpt from Wald (2007).

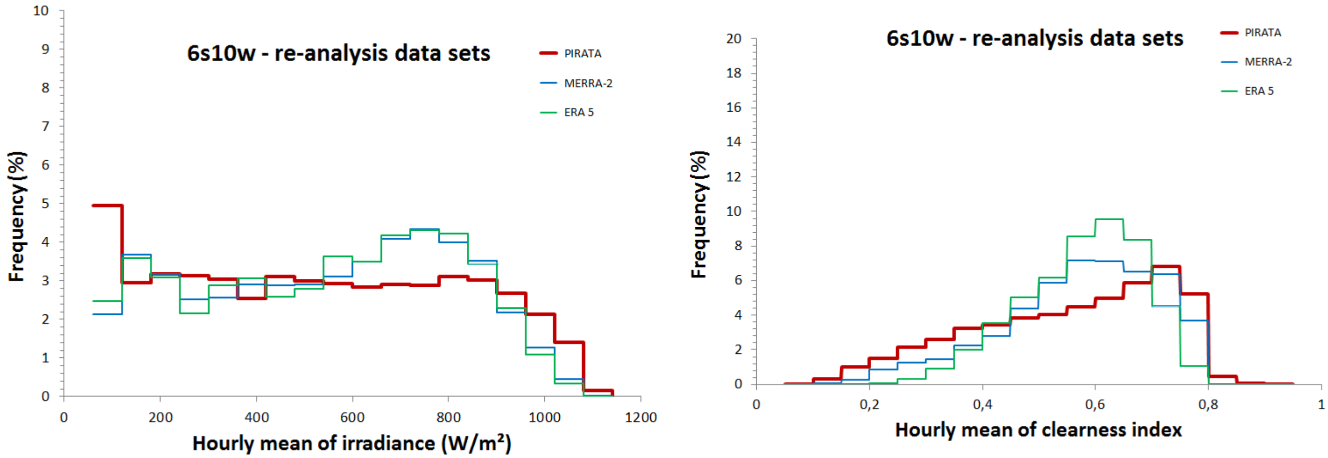


Figure 7: Frequency distributions of PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 6s10w for *E* (left) and *KT* (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

The frequency distributions of *E* and *KT* from measurements are not well reproduced by the two re-analyses data sets. Though there are some differences, the two re-analyses exhibit the same features between them and at all sites for *E*. There is an underestimation of the frequencies for *E* less than 150 W m^{-2} and of frequencies for *E* greater than approximately 800 W m^{-2} and an overestimates of estimation for *E* in the range [550, 800] W m^{-2}. As for the clearness index, the estimates by ERA5 underestimate the frequencies for KT<0.4 and KT>0.7, and overestimate the frequencies for clearness indices in-between. This means that there are not enough cases of overcast conditions as well as cloud-free conditions in the ERA5 data set compared to the PIRATA measurements. The situation is more complex for MERRA-2 at it depends upon the site. At 0n0e, 0n10w and 0n23w, the estimates by MERRA-2 overestimate the frequencies for clearness indices in the range [0.50, 0.65] and underestimate the frequencies when KT>0.65, i.e. there are not enough cases of cloud-free conditions. On the contrary, MERRA-2 underestimate the frequencies for KT<0.45 -not enough cases of overcast conditions- and overestimate the frequencies in the range [0.45, 0.7] at 6s10e and 19s34w.

The two re-analyses differ between them regarding the estimation of the monthly means and standard deviations of *E* and each data set does not exhibit the same features at all sites (see Fig. 8 for the site 6s10w). The monthly means of the estimated *E* from MERRA-2 tend to be close to those from PIRATA in the period August-December (January-March at 6s10w); during the other months, they underestimate those from PIRATA at 0n0e, 0n10w and 0n23w, and overestimate at 6s10w and 19s34w. As for ERA5, the estimated means tend to either overestimate or to be close, depending on the site and the period without clear scheme (see Appendix). For example, there is a large overestimation all year long at 0n0e and means are similar for all months at 19s34w.

MERRA-2 tends to underestimate the standard deviation of E within a month all year long, except at 0n10w and 0n23w where MERRA-2 is close in December and January, and at 19s34w where MERRA-2 is very close in February to August. ERA5 tends to underestimates the standard deviation during the year except at 0n10w and 6s10w where the standard deviations from ERA5 are close to those from PIRATA in the period from approximately October to January.

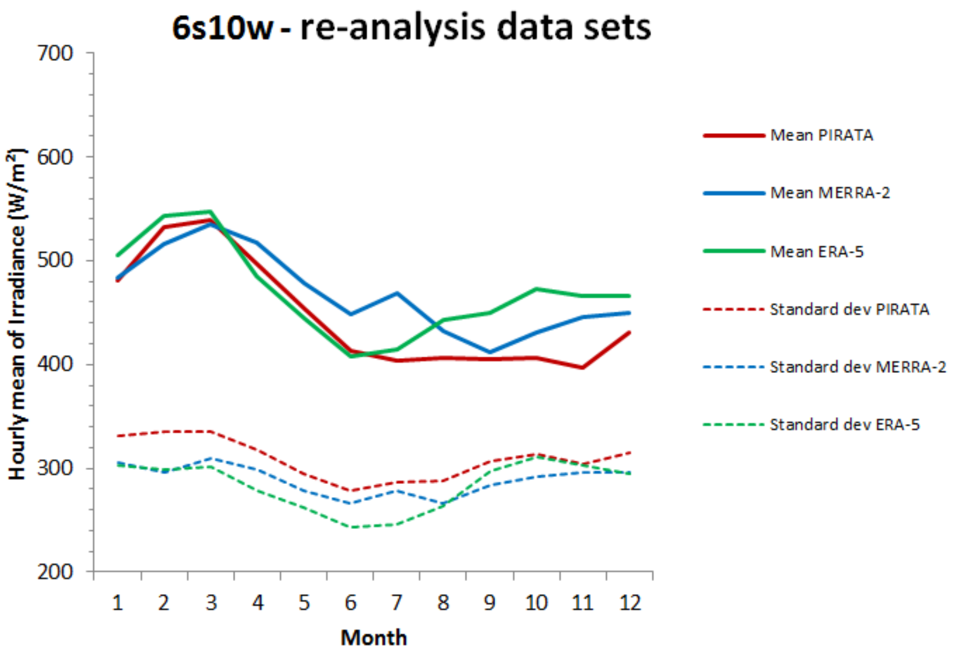


Figure 8: Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 6s10w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

3.2.2 Daily values

Table 6 shows the correlation coefficients, the biases and the standard deviations at each buoy for the two re-analyses data sets, for daily means of E and KT (correlation coefficient only) in the period 2012-2013. These quantities are similar to those for 2010 and 2011 (not presented).

Buoy	Correlation coefficient MERRA-2	Correlation coefficient ERA5	Bias MERRA-2	Bias ERA5	Standard deviation MERRA-2	Standard deviation ERA5
0n0e	0.379 *(0.355)*	0.647 *(0.599)*	-20 (-9 %)	13 (5 %)	55 (25 %)	40 (18 %)
0n10w	0.328 *(0.360)*	0.547 *(0.563)*	-8 (-3 %)	12 (5 %)	52 (22 %)	38 (16 %)
0n23w	0.350 *(0.386)*	0.513 *(0.544)*	-5 (-2 %)	-3 (-1 %)	41 (16 %)	31 (11 %)
6s10w	0.485 *(0.465)*	0.479 *(0.368)*	6 (2 %)	7 (3 %)	41 (17 %)	41 (17 %)
19s34w	0.809 *(0.584)*	0.873 *(0.705)*	11 (4 %)	-5 (-2 %)	40 (16 %)	33 (13 %)

Table 6: For each PIRATA buoy and each re-analysis data set: Correlation coefficient between measurements and estimates from re-analysis for irradiance and clearness index *(in italics)*, bias and standard deviation (W m^{-2}) between measurements and estimates for irradiance with relative values (in brackets) for daily means in the period 2012-2013.

Expectedly and similarly to the case of the satellite-derived data sets, the numbers in this Table are consistent between hourly and daily values for a given data set and a given PIRATA site: the relative biases are the same for the hourly and daily means, and the standard deviations of errors are less for daily values than for hourly values because of the averaging over the day (Table 5, 6). The correlation coefficients for the daily DSIS are low: from 0.33 to 0.49 (0.81 for 19s34w) for MERRA-2, and from 0.48 to 0.65 (0.87 for 19s34w) for ERA5. They are less than those for hourly *E*. The correlation coefficients are also low for the daily *KT*: from 0.36 to 0.47 (0.58 for 19s34w) for MERRA-2 and from 0.37 to 0.60 (0.71 for 19s34w) for ERA5. At most 22 % (34 % for 19s34w) of the variance contained in the measured daily *KT* is explained by the estimates in MERRA-2; it is more for ERA5 with 36 % (50 % for 19s34w). There is no regional trend of the correlation coefficients for any of the two data sets. The 2D histograms (not shown) have shapes which are not elongated and look more like discs. They exhibit very large scattering, in full agreement with the low correlation coefficients. For both data sets, *E*, respectively *KT*, is sometimes overestimated but more frequently underestimated, especially for the greatest values.

3.2.3 Discussion on the two re-analyses data sets

One striking feature observed in both data sets is that both re-analyses often report cloud-free conditions while actual conditions are cloudy and reciprocally, actual cloud-free conditions as cloudy, at both hourly and daily time scales. Such observations were reported for daily DSIS over the Atlantic Ocean by Boilley and Wald (2015) for the MERRA and ERA-Interim re-analyses.

The bias in hourly *E* ranges between -42 and 23 W m^{-2} for MERRA-2 and shows a tendency to increase (from negative to positive) from East to West. The dependency of the bias with the solar zenithal angle and other variables is weak, except for the cloud type whose influence is prominent. Fig. 9 shows the dependency of the bias (left) and of the correlation coefficient (right) as a function of the cloud type from the CAMS cloud classification at each site for MERRA-2. In a given graph, a column, respectively row, would be uniformly coloured in case of no dependency of the bias or correlation coefficient with the type of cloud, respectively with the site.

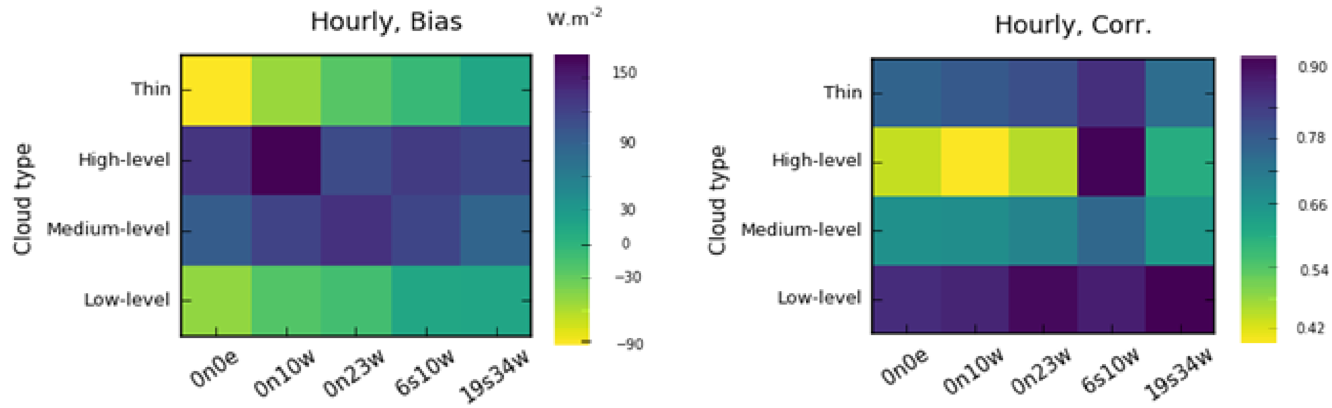


Figure 9: Bias (left, in W m^{-2}) and correlation coefficient (right) as a function of the cloud type at each PIRATA buoy for MERRA-2.

One observes greenish tones in the lowest row in Fig. 9 (left) for the ‘low-level’ type (water cloud at low altitude). This means biases close to 0 for this type at each site. There is a slight change in the green tones over the row, meaning that the bias does not vary much from one site to another for this type of cloud though the tendency to increase (from negative to positive) from East to West (from left to right) is noticeable. The tones are blueish for the ‘medium-level’ (water cloud at medium altitude) and ‘high-level’ (deep cloud of large vertical extent from low altitude to medium altitude), meaning large positive bias, with changes from site to site without any clear trend. The ‘thin’ type (thin ice cloud) exhibits yellow to greenish tones, with negative bias. One may note the increase of the bias (from negative to positive) from East to West.

Fig. 9 (right) shows that the correlation coefficients are similar or very close for all stations (each row is fairly uniformly coloured), except for ‘high-level’ clouds for which the coefficients vary very much with the site without a clear trend. This type of cloud exhibits the lowest correlation coefficients while the ‘low-level’ and ‘thin’ types offer the greatest ones. The fact that the ‘medium-level’ and ‘high-level’ clouds exhibit more bias and less correlation than the ‘low-level’ clouds is consistent with the recent and preliminary findings of Doddy et al. (2017) who looked at the differences between measurements of daily *E* performed at terrestrial stations in Ireland and MERRA-2 outputs and suggested a systematic link between prevailing cloud structures and errors.

The bias in hourly *E* ranges between -10 and 25 W m^{-2} for ERA5. It exhibits a regional tendency to decrease in absolute values and to tend to underestimation with increasing mean *KT* (Table 5). However, such complex behaviours can only be speculated given the small number of sites. The dependency of the bias with the solar zenithal angle and other variables is weak, except for the cloud type which is discussed now. Fig. 10 shows the dependency of the bias (left) and of the correlation coefficient (right) as a function of the cloud type from the CAMS cloud classification at each site for ERA5.

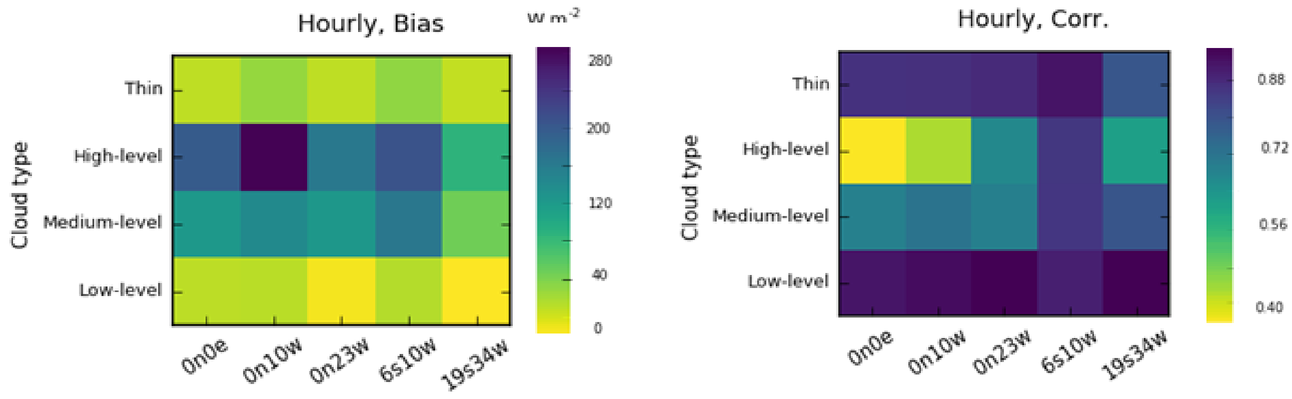


Figure 10: Bias (left, in W m^{-2}) and correlation coefficient (right) as a function of the cloud type at each PIRATA buoy for ERA5.

One observes yellowish and greenish tones in the lowest row in Fig. 10 (left) for the 'low-level' type for ERA5. The bias is positive and small. It is the lowest at 0n23w and 19s34w which are the sites with the greatest mean *KT*. The bias is also small and positive for the 'thin' clouds with a weak dependency upon the sites. Biases are greater for the two other types and exhibit greater changes from site to site, though the buoys 0n23w and 19s34w show lower biases than the other sites. One may note that the bias in Fig. 10 (left) are positive for every row, i.e. every type, while the bias for all conditions is either negative or positive (Table 5). This is explained by the fact that these sites exhibit a large number of cloud-free conditions during which the DSIS tends to be underestimated by ERA5 as discussed above. Similarly to MERRA-2, the correlation coefficients (Fig. 10 right) show large values that are fairly the same between sites for the 'low-level' and 'thin' types, and the correlation is weaker for the two other types with coefficients varying from site to site without a clear trend.

MERRA-2 exhibits large changes in biases from one site to another with a trend from East to West as previously reported; the amplitude of change in bias between sites is 65 W m^{-2} for hourly DSIS. The high variability of bias in space indicates an artificial spatial distortion of the field of DSIS. In an initial evaluation of the climate in MERRA-2, Koster (2015) reported significant regional biases between the yearly means of the DSIS of MERRA-2 and those from the CERES (Clouds and the Earth's Radiant Energy System) EBAF (Energy Balanced and Filled) satellite-based observational data set, the latter being taken by Koster as a reference. In his Figure 4.6, one may see a noticeable difference between both data sets. It ranges from -20 W m^{-2} in the Gulf of Guinea to 20 W m^{-2} along the Brazilian coast. The findings reported in the present work give some flesh to the work of Koster as the magnitude of this difference and its spatial structure are supported by the present work which is built on in situ measurements.

The spatial distortion in MERRA-2 combines with very large standard deviations of errors, about 130 W m^{-2}, and this may indicate that the spatial distortion occurs at various temporal scales. This is supported by the findings of Bengulescu et al. (2017). These authors performed a comparison between several data sets, among which HC3v5 and MERRA-2, and in situ measurements made at two terrestrial stations: Vienna (Austria) and Kishinev (Moldova). They reported a very high correlation coefficient between MERRA-2 and in situ measurements (0.97 and 0.97 respectively) and showed that this high correlation was mostly due to the very high correlation coefficients between MERRA-2 and in situ measurements at the yearly period (0.99 at both stations), i.e. MERRA-2 reproduces well the seasonal variability. For any period less than 1 year, the correlation coefficient is less than 0.8, which means that less than 64 % of the variance of the measurements is explained by MERRA-2, i.e. MERRA-2 does not reproduce the variability observed in measurements for these periods shorter than 1 year.

ERA5 exhibits biases and standard deviations of errors that are much greater (in absolute values for the bias) than those for the satellite-derived data sets. The difference in bias for hourly DSIS from one site to another is 35 W m^{-2} for ERA5, slightly less than that of SARAH-2: 44 W m^{-2}. This change in bias in space means that the spatial features of E are partly artificially distorted. One notes that the standard deviation of error is fairly similar from one site to another: the amplitude of change is only 29 W m^{-2}, much less than those from the satellite-derived data sets. However, it is very large of order of 130 W m^{-2}. Though the correlation coefficients are close to 0.90, the large scattering of errors combined with the changes in bias hampers the use of ERA5 in detailed studies of the spatial and temporal variability of the DSIS.

3.3 About the differences in spatial support of the buoy and the grid cell of the data sets

One may object that the size of the grid cell is inappropriate for the comparison with a single buoy because surface measurements are for a single point in space, whereas the estimated irradiances are for the area of a pixel (typically 5 km) or a grid cell (typically 50 km). Cloud properties may vary within the grid cell and large random errors are unavoidable at hourly time steps. Using monthly averages is a means to reduce the errors caused by the problem (see e.g. Zhao et al., 2013). One may believe that this mismatch in spatial support of information may explain the performances of the re-analyses presented here. However, it can be argued that there is no orographic effect in the Atlantic Ocean and there is no strong systematic gradient in irradiance over short distances corresponding to the hourly time step. Hence, the irradiance field is fairly homogeneous at sub-meso-scale and this should mitigate the effects of the differences in spatial support of the buoy and the grid cell. In addition, one may note that the drawbacks reported above are also observed at daily scale. Finally, the work of Boilley and Wald (2015) can be mentioned. These authors compared the satellite-derived HelioClim-1 data set to PIRATA measurements. HelioClim-1 is fairly similar to the re-analyses with regard to the spatial support of information because it is made of estimates of the DSIS made on 5 km pixels spaced by 25 km in both latitude and longitude (Lefèvre et al., 2007, 2014), and a spatial bi-linear interpolation was performed to create the time-series at PIRATA locations. Though the period is not the same than in the present study (HelioClim-1 covers the period 1985-2005), one may compare the correlation coefficients reported by these authors that range between 0.82 and 0.88 for daily E and from 0.79 to 0.88 for daily KT for HelioClim-1, and are much greater than those obtained for the re-analyses both in the work of Boilley and Wald and here (Table 2). These findings of Boilley and Wald support the argument that differences in spatial support of information cannot be the only reason for the bad performances of the re-analyses.

Conclusions

This work brings new information on the capabilities of five data sets about assessing the magnitude of the DSIS and its variability in space and time at both hourly and daily time step in the tropical Atlantic Ocean for a more informed usage of these data in ocean sciences. Five buoys within the PIRATA network are offering enough data of high quality to perform an assessment of the two meteorological analyses MERRA-2 and ERA5 and the three satellite-derived data sets HC3v5, SARAH-2 and CRS.

It was found that the re-analyses MERRA-2 and ERA5 often report cloud-free conditions while actual conditions are cloudy, yielding an underestimation of surface irradiance, and reciprocally, actual cloud-free conditions as cloudy, yielding an overestimation. These alternating underestimations and overestimations compensate each other with a small bias as a result masking some deficiencies in properly modelling cloud properties. These conclusions are similar to those already reported regarding meteorological re-analyses as a whole (Wild, 2008). The estimates from MERRA-2 or ERA5 poorly correlate with the clearness indices at buoys: the correlation coefficient ranges between 0.49 and 0.56 for MERRA-2, and between 0.57 and 0.64 for ERA5. Hence, a large part of the variability in the optical state of the atmosphere is not captured by the MERRA-2 or ERA5 re-analyses. It is recommended not to use them in studies of the variability in time of the surface irradiance in the tropical Atlantic Ocean when it is necessary to reproduce actual measurements.

The buoys experience a large amount of cloud-free conditions under which ERA5 tends to underestimate the irradiance. The bias for the two re-analyses depend on the cloud type: they exhibit more bias in irradiance in cases of medium and high level clouds than for low level clouds.

The bias varies noticeably with the calendar month, which means that MERRA-2 or ERA5 cannot be used confidently at a monthly scale. The re-analyses exhibit small biases when compared to PIRATA measurements over one or more years though the study was limited to 4 years for 2 buoys. It can be speculated that one may use them to follow changes in yearly values of irradiance at one location.

Another striking feature is the variability of the bias and other performance indicators for both MERRA-2 and ERA5 within this ocean area which is fairly homogeneous for the irradiance and clearness index. Accordingly, an additional recommendation on re-analyses may be not to use them to study the spatial field of irradiance at whatever time scale: the performances strongly vary from one location to another, especially for MERRA-2, which means that the field of surface irradiance is spatially distorted, even on a yearly scale.

The present results bring new evidence on the qualities and limitations of MERRA-2 and ERA5 which have been little studied for the irradiance at surface. These re-analyses may be used in studies of the tropical Atlantic Ocean with proper understanding of the limitations and uncertainties. Zhao et al. (2013) proposed an empirical relationship for correcting the bias observed between MERRA estimates and measurements of monthly averages of irradiance performed at several sites in North America taking into account the dependency between the bias and KT and surface elevation. The bias and the root mean square error were reduced but at the expenses of an increase in standard deviation of errors. Jones et al. (2017) have tested several methods using HC3v5 for decreasing the bias of the ERA-Interim daily estimates of *E*. They found that when compared to measurements of daily irradiance performed at 55 terrestrial stations in Europe, the bias was reduced at 10 stations and similar for the others and that the other indicators (standard deviation of errors, root mean square error, correlation coefficient, median of errors, etc.) were unchanged. Though the works were performed for MERRA or ERA-Interim, it is speculated that similar conclusions would be reached when applied to MERRA-2 or ERA5, given the similarities between these re-analyses.

The three satellite-derived data sets exhibit similar performances between them and have better performance indicators than the two re-analyses, except for the bias. The correlation coefficients are greater than 0.95 for irradiance and 0.80 for clearness index in most cases and at hourly and daily time scales. Each data set reproduces well the dynamics of the irradiance at both time scales though amplitudes of variation in time may be hindered by the large standard deviation of the errors which amounts to approximately 80 W m^{-2} for hourly DSIS and 20 W m^{-2} for daily DSIS. The same conclusion applies to the clearness index.

The three satellite-derived data sets exhibit overestimation, with the lowest biases reached by CRS and ranging between 10 and 36 W m^{-2} depending on site. The bias for both SARAH-2 and CRS shows a tendency to decrease as the mean clearness index increases. The bias of HelioClim-3v5 is almost the same for the five sites and the irradiance pattern may not be noticeably distorted by artefacts induced by spatial changes in bias. For this reason, one should prefer HC3v5 when the study of the irradiance field and of its spatial features is at stake.

It can be concluded that the three satellite-derived data sets are appropriate to study the dynamics of the downward solar irradiance at the surface of the tropical Atlantic Ocean and that their performances are fairly similar. Assuming that pyranometer measurements of the PIRATA buoys achieve the “moderate quality” defined by WMO (2008, rev. 2012), one may ask if the estimates from the satellite-derived data sets are compliant with “moderate quality”, assuming that the bias may be removed. The relative uncertainty is defined as the 95 % probability (P95) and should not exceed 20 % to meet the “moderate quality” (WMO, 2008, rev. 2012). The total uncertainty takes into account the uncertainty of PIRATA and the uncertainty of the estimates. It can be expressed in a first approximation as the quadratic sum of both uncertainties. As a consequence, the total relative uncertainty should not exceed 28 % (P95), or 14 % (P66) if the estimates were of “moderate” quality. The standard deviations (P66) for each data set reported in Table 3 are below 14 %. It can be concluded that to a first approximation, the three satellite-derived data sets can be considered of moderate quality if bias can be removed.

One may note several similarities in performances between HC3v5 and SARAH-2. It is speculated that this is partly due to the fact that they exploit the same method, Heliosat-2, though the implementation differs.

Other data sets are available that cover the tropical Atlantic Ocean and may be assessed against the PIRATA measurements to gain knowledge on their limitations and confidence in their use. Examples are the satellite-derived OSI-SAF (www.osi-saf.org) or the Japanese 55-year re-analysis (JRA 55, Kang et al., 2015; Kobayashi et al., 2015).

The findings reported here are consistent with the very few works already published for terrestrial stations. This demonstrates a posteriori that the PIRATA measurements may be used for the validation of models and data sets. However, some uncertainties remain. Potential biases in the PIRATA time series are an issue and difficult to estimate. It complicates validation of the satellite-derived data sets. While the different levels and variability of surface irradiance at buoys might impact the quality of the satellite-based data sets, a reduced data quality of the buoy data (despite the quality control applied) might also have an impact on the presented evaluation. Studies like these when multiple data records are considered can help to identify problems in surface reference measurements (Urraca et al., 2017). The PIRATA network is a unique and valuable means to study and monitor the surface irradiance in the tropical Atlantic Ocean and deserves support for operations to further enrich the data records.

Data availability

PIRATA measurements performed every 2 min were downloaded from the web site (www.pmel.noaa.gov/tao/drupal/disdel/) of the National Oceanic and Atmospheric Administration (NOAA) of the U.S.A. The authors acknowledge the help of the GTMBA Project Office of NOAA/PMEL in getting the data and the PIRATA team for servicing the network and freely providing the data.

Time-series of HelioClim-3v5 data were downloaded from the SoDa Service web site (www.soda-pro.com) managed by the company Transvalor. Data are available to anyone for free for years 2004-2006 as a GEOSS Data-CORE (GEOSS Data Collection of Open Resources for Everyone) and for a charge for the most recent years with the amount depending on requests and requester. The time-series used in this article are available for free in CSV format by request to Mireille Lefèvre.

Time-series of SARAH-2 data were extracted from the gridded data sets available at https://doi.org/10.5676/EUM_SAF_CM/SARAH/V002.

Time-series of CAMS Radiation Service data were downloaded from the SoDa Service web site (www.soda-pro.com).

Time-series of cloud classification were downloaded from the SoDa Service web site (www.soda-pro.com).

MERRA-2 times-series were extracted from the gridded data sets available at https://goldsmr4.gesdisc.eosdis.nasa.gov/data/MERRA2/.

ERA5 times-series were extracted from the gridded data sets available at http://apps.ecmwf.int/data-catalogues/era5/?class=ea&stream=enda&expver=1.

Author contribution

All authors contributed equally to this work.

Competing interests

The authors declare no competing interests.

Disclaimer

N/A

Acknowledgements

The research leading to these results has partly received funding from the Copernicus Atmosphere Monitoring Service, a program being operated by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Union. The authors thank the French company Transvalor, which takes care of the SoDa Service for the common good, thus providing an efficient access to the HelioClim databases. The authors thank especially Gregory Foltz for his helpful advice on the PIRATA measurements, the two anonymous reviewers and Yehia Eissa for their kind and useful comments which greatly help in improving the clarity of this article.

References

Bengulescu, M., Blanc, P., Boilley, A., and Wald, L.: Do modelled or satellite-based estimates of surface solar irradiance accurately describe its temporal variability?, *Adv Sci Res*, 14, 35-48, doi:10.5194/asr-14-35-2017, 2017.

Blanc, P., and Wald L.: The SG2 algorithm for a fast and accurate computation of the position of the Sun, *Sol Energy*, 86, 3072-3083, doi: 10.1016/j.solener.2012.07.018, 2012.

Boilley, A., and Wald, L.: Comparison between meteorological re-analyses from ERA-Interim and MERRA and measurements of daily solar irradiation at surface, *Renew Energ*, 75, 135-143, doi: 10.1016/j.renene.2014.09.042, 2015.

Bourlès, B., Lumpkin, R., McPhaden, M. J., Hernandez, F., Nobre, P., Campos, E., Yu, L., Planton, S., Busalacchi, A., Moura, A. D., Servain, J., and Trotte, J.: The Pirata Program: History, accomplishments, and future directions, *B Am Meteorol Soc*, 89, 1111–1125, doi:10.1175/2008BAMS2462.1, 2008.

Budyko, M. I.: The effect of solar radiation variations on the climate of the Earth, *Tellus*, 21, 611-619, doi:10.1111/j.2153-3490.1969.tb00466.x, 1969.

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

Doddy, E., Sweeney, C., McDermott, F.: An investigation of systematic errors in solar radiation from reanalysis datasets. EMS Annual Meeting 2017, Dublin, Ireland, 4-8 September 2017. Abstract EMS2017-675.

Eissa, Y., Munawwar, S., Oumbe, A., Blanc, P., Ghedira, H., Wald, L., Bru, H., and Goffe, D.: Validating surface downwelling solar irradiances estimated by the McClear model under cloud-free skies in the United Arab Emirates, *Sol Energy*, 114, 17-31, doi:10.1016/j.solener.2015.01.017, 2015.

Foltz, G. R., Evan, A. T., Freitag, H. P., Brown, S., and McPhaden, M. J.: Dust accumulation biases in PIRATA shortwave radiation records, *J Atmos Ocean Technol*, 30, 1414–1432, doi: 10.1175/JTECH-D-12-00169.1, 2013.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2), *J Climate*, 30, 5419–5454, doi: 10.1175/JCLI-D-16-0758.1, 2017.

Gschwind, B., Ménard, L., Albuisson, M., and Wald, L.: Converting a successful research project into a sustainable service: the case of the SoDa Web service, *Environ Modell Softw*, 21, 1555-1561, doi:10.1016/j.envsoft.2006.05.002, 2006.

Hersbach, H., and Dee, D.: ERA5 reanalysis is in production, ECMWF Newsletter No. 147, p. 7, 2016.

ISO Guide to the Expression of Uncertainty in Measurement: first edition, International Organization for Standardization, Geneva, Switzerland, 1995.

Jones, P., Harpham, C., Troccoli, A., Gschwind, B., Ranchin, T., Wald, L., Goodess, C., and Dorling, S.: Using ERA-Interim reanalysis for creating datasets of energy-relevant climate variables, *Earth Syst Sci Data*, 9, 471-495, 10.5194/essd-9-471-2017, 2017.

Kang, S., and Ahn, JB.: Global energy and water balances in the latest reanalyses, *Asia-Pacific J Atmos Sci*, 51, 293-302, doi: 10.1007/s13143-015-0079-0, 2015.

Katsaros, K. B., and DeVault, J. E.: On irradiance measurement errors at sea due to tilt of pyranometers, *J Atmos Ocean Tech*, 3, 740-745, doi: 10.1175/1520-0426(1986)003<0740:OIMEAS>2.0.CO;2, 1986.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General specifications and basic characteristics, *J Meteor Soc Japan*, 93, 5-48, doi:10.2151/jmsj.2015-001, 2015.

Korany, M., Boraiy, M., Eissa, Y., Aoun, Y., Abdel Wahab, M. M., Alfaro, S. C., Blanc, P., El-Metwally, M., Ghedira, H., Hungershoefer, K., and Wald, L.: A database of multi-year (2004-2010) quality-assured surface solar hourly irradiation measurements for the Egyptian territory, *Earth Syst Sci Data*, 8, 105–113, doi:10.5194/essd-8-105-2016, 2016.

Koster, R. (editor): Technical Report Series on Global Modeling and Data Assimilation, Volume 43, National Aeronautics and Space Administration (NASA), 139 p., 2015.

Lean, J., and Rind, D.: Climate forcing by changing solar radiation, *J Climate*, 11, 3069–3094, doi: 10.1175/1520-0442(1998)011<3069:CFBCSR>2.0.CO;2, 1998.

Lefèvre, M., and Wald, L.: Validation of the McClear clear-sky model in desert conditions with three stations in Israel, *Adv Sci Res*, 13, 21-26, doi:10.5194/asr-13-21-2016, 2016.

Lefèvre, M., Diabaté, L., and Wald, L.: Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance, *Sol Energy*, 81, 240-253, doi:10.1016/j.solener.2006.03.008, 2007.

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., and Morcrette, J.-J.: McClear: a new model estimating downwelling solar radiation at ground level in clear-sky condition, *Atmos Meas Tech*, 6, 2403-2418, doi: 10.5194/amt-6-2403-2013, 2013.

Lefèvre, M., Blanc, P., Espinar, B., Gschwind, B., Ménard, L., Ranchin, T., Wald, L., Saboret, L., Thomas, C., and Wey, E.: The HelioClim-1 database of daily solar radiation at Earth surface: an example of the benefits of GEOSS Data-CORE, *IEEE JSTARS*, 7, 1745-1753, doi:10.1109/JSTARS.2013.2283791, 2014.

Liu, C., Allan, R. P., Mayer, M., Hyder, P., Loeb, N. G., Roberts, C. D., Valdivieso, M., Edwards, J. M., and Vidale, P.-L.: Evaluation of satellite and reanalysis based global net surface energy flux and uncertainty estimates, *J Geophys Res Atmos*, 122, doi:10.1002/2017JD026616, 2017.

Long, C. N., Bucholtz, A., Jonsson, H., Schmid, B., Vogelmann, A., and Wood, J.: A method of correcting for tilt from horizontal in downwelling shortwave irradiance measurements on moving platforms, The Open Atmospheric Science Journal, 4, 78-87, doi: 10.2174/1874282301004010078, 2010.

MacWhorter, M. A., and Weller, R. A.: Error in measurements of incoming shortwave radiation made from ships and buoys, *J Atmos Ocean Tech*, 8, 108-117, doi: 10.1175/1520-0426(1991)008<0108:EIMOIS>2.0.CO;2, 1991.

Manabe, S.: Climate and the ocean circulation, *Mon Wea Rev*, 97, 739-774, doi:10.1175/1520-0493(1969)097<0739:CATOC>2.3.CO;2, 1969.

Marchand, M., Al-Azri, N., Ombe-Ndeffotsing, A., Wey, E., and Wald, L.: Evaluating meso-scale change in performance of several databases of hourly surface irradiation in South-eastern Arabic Peninsula, *Adv Sci Res*, 14, 7-15, doi:10.5194/asr-14-7-2017, 2017.

Müller, R., Pfeifroth, U., Traeger-Chatterjee, C., Trentmann, J. and Cremer, R.: Digging the METEOSAT treasure-3 decades of solar surface radiation, *Remote Sens-Basel*, 7, 8067-8101, doi:10.3390/rs70608067, 2015.

Muneer, T., and Fairooz, F.: Quality control of solar radiation and sunshine measurements - lessons learnt from processing worldwide databases, *Build Serv Eng Res T*, 23, 151-166, doi: 10.1191/0143624402bt038oa, 2002.

Pfeifroth, U., Kothe, S., Müller, R., Trentmann, J., Hollmann, R., Fuchs, P., and Werscheck M.: Surface Radiation Data Set - Heliosat (SARAH) - Edition 2. Satellite Application Facility on Climate Monitoring. doi: 10.5676/EUM_SAF_CM/SARAH/V002, 2017a.

Pfeifroth, U., Sanchez-Lorenzo, A., Manara, V., Trentmann, J., and Hollmann, R.: Trends and variability of surface solar radiation in Europe based on surface- and satellite-based data records, *J Geophys Res*, under review, 2017b.

Qu, Z., Gschwind, B., Lefèvre, M., and Wald, L.: Improving HelioClim-3 estimates of surface solar irradiance using the McClear clear-sky model and recent advances in atmosphere composition, *Atmos Meas Tech*, 7, 3927–3933, doi:10.5194/amt-73927-2014, 2014.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald L.: Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method, *Meteorol Z*, 26, 33-57, doi:10.1127/metz/2016/0781, 2017.

Reynolds, R. M.: Correcting global shortwave irradiance measurements for platform tilt, Internal Report, RMR Company, available at www.rmrco.com/docs/m0703_psp_lowlevel_correction.pdf, last accessed: 2017-07-28.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA’s modern-era retrospective analysis for research and applications, *J Climate*, 24, 3624-3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Rigollier, C., Lefèvre, M., and Wald, L.: The method Heliosat-2 for deriving shortwave solar radiation from satellite images, *Sol Energy*, 77, 159-169, doi 10.1016/j.solener.2004.04.017, 2004.

Siegel, D. A., Ohlmann, J. C., Washburn, L., Bidigare, R. R., Nosse, C. T., Fields, E., and Zhou, Y.: Solar radiation, phytoplankton pigments and the radiant heating of the equatorial Pacific warm pool, *J Geophys Res*, 100(C3), 4885–4891, doi:10.1029/94JC03128, 1995.

Thomas, C., Wey, E., Blanc, P., and Wald, L.: Validation of three satellite-derived databases of surface solar radiation using measurements performed at 42 stations in Brazil, *Adv Sci Res*, 13, 81-86, doi:10.5194/asr-13-81-2016, 2016.

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

Wald L.: Solar radiation energy (fundamentals). In Solar Energy Conversion and Photoenergy Systems, edited by Julian Blanco and Sixto Malato, in Encyclopedia of Life Support Systems (EOLSS), Developed under the Auspices of the UNESCO, Eolss Publishers, Oxford ,UK, [http://www.eolss.net], 2007.

Wild, M.: Short-wave and long-wave surface radiation budgets in GCMs: a review based on the IPCC-AR4/CMIP3 models, *Tellus*, 60A, 932–945, doi: 10.1111/j.1600-0870.2008.00342.x, 2008.

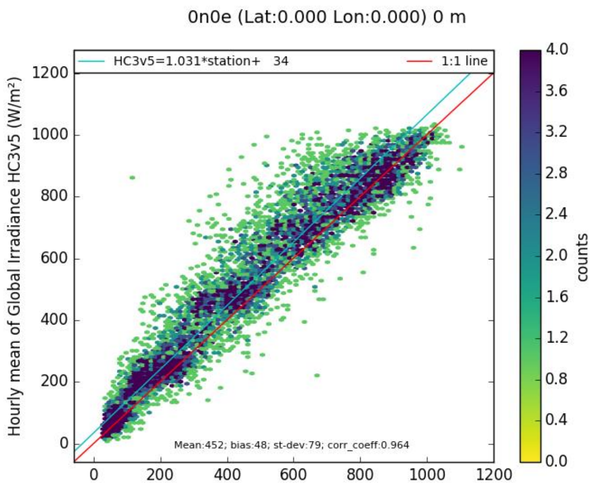
WMO: Guide to meteorological instruments and methods of observation, WMO-No 8, 2008 edition updated in 2010, World Meteorological Organization, Geneva, Switzerland, 2012.

Zhao, L., Lee, X., and Liu, S.: Correcting surface solar radiation of two data assimilation systems against FLUXNET observations in North America, *J Geophys Res Atmos*, 118, 9552–9564, doi:10.1002/jgrd.50697, 2013.

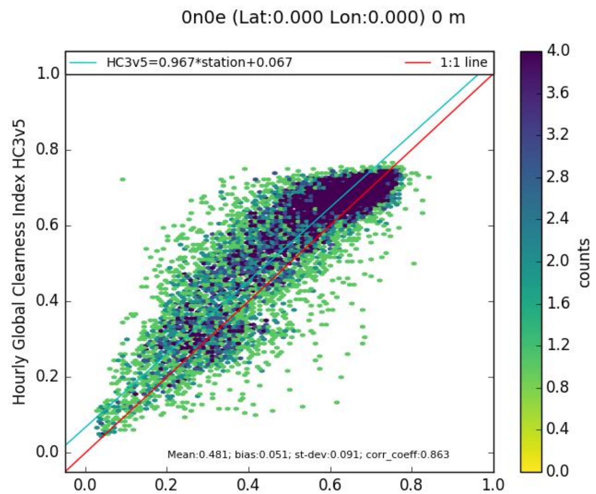
Appendix

0n0e

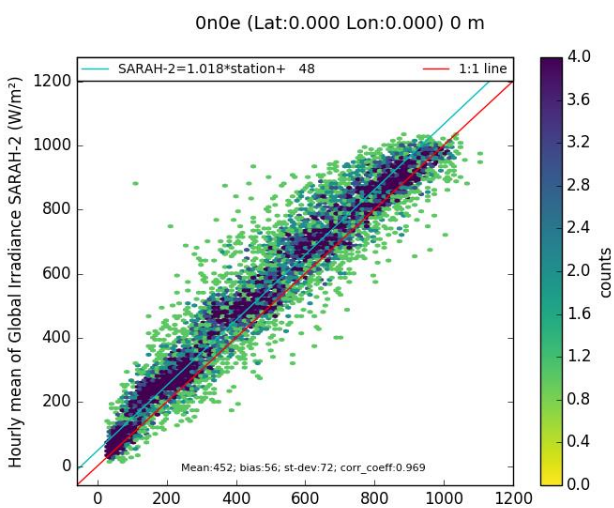
Satellite-derived data sets



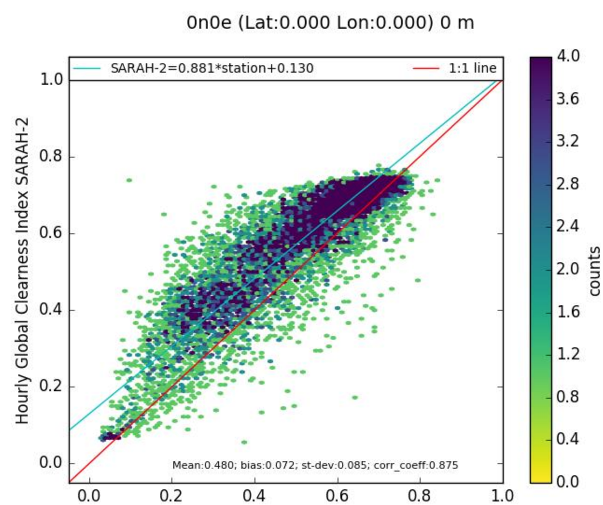
(a)



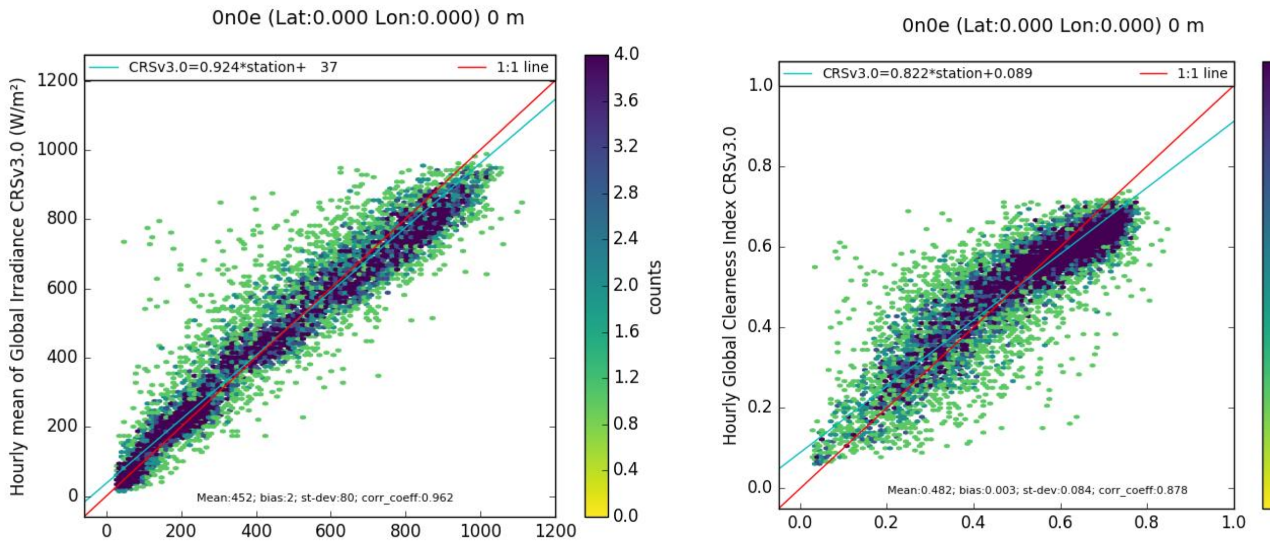
(b)



(c)

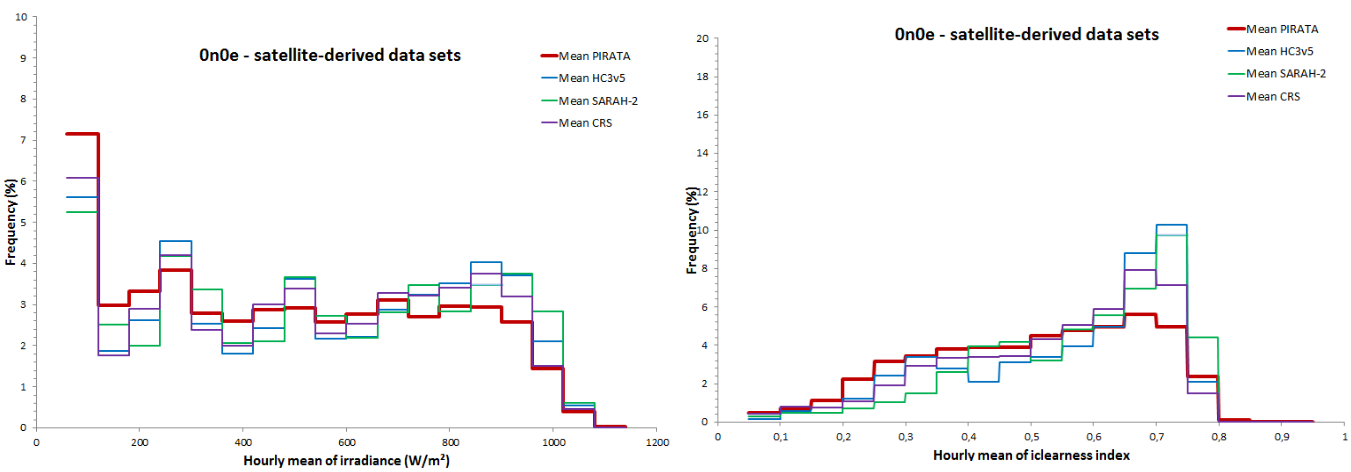


(d)

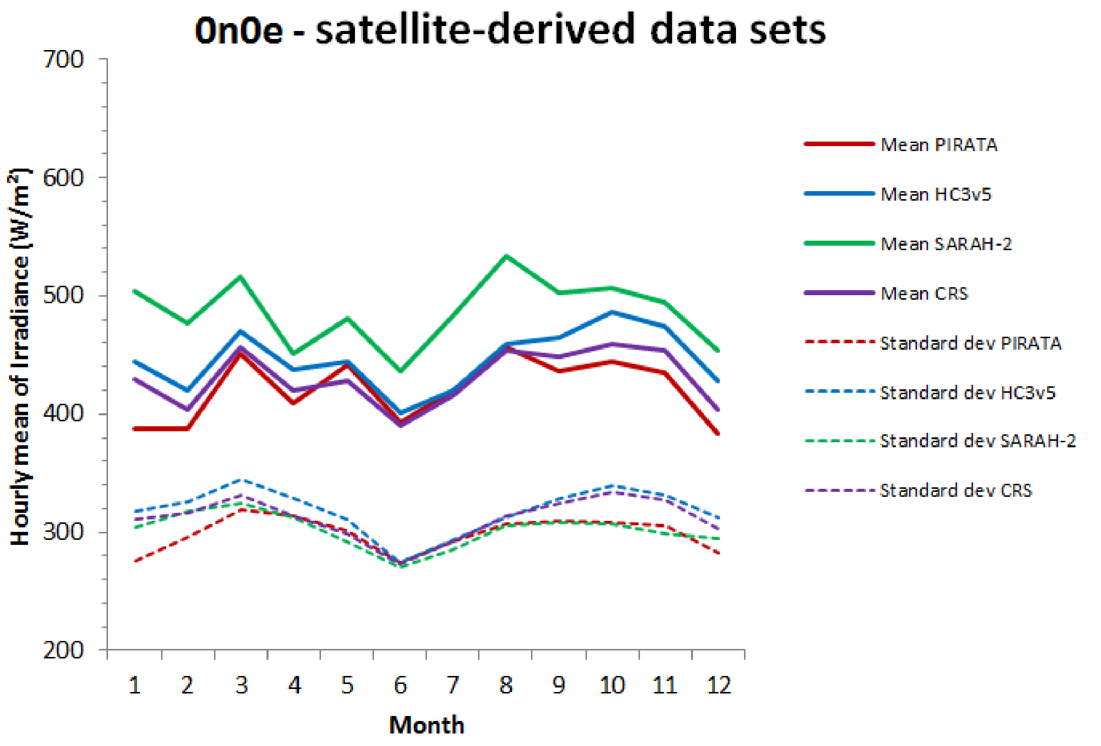


(e) (f)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) for the station 0n0e for E (left) and *KT* (right). HC3v5: (a), (b); SARAH 2: (c), (d); CRS: (e), (f). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should overlay the red line.

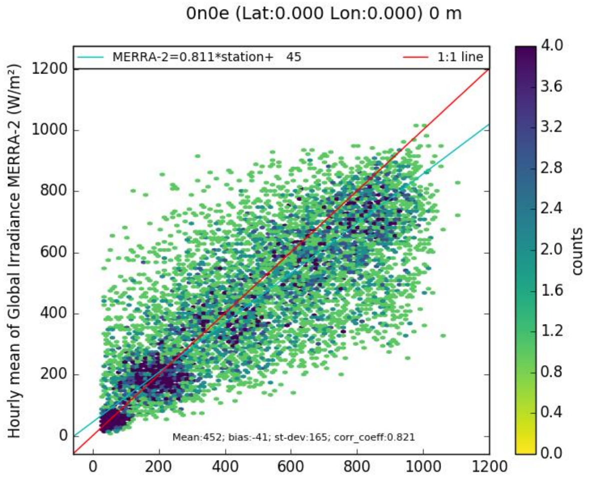


Frequency distributions of PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) for the station 0n0e for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

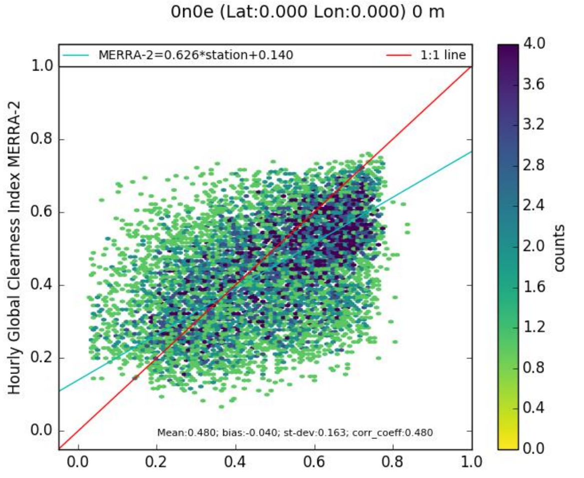


Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) at 0n0e. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

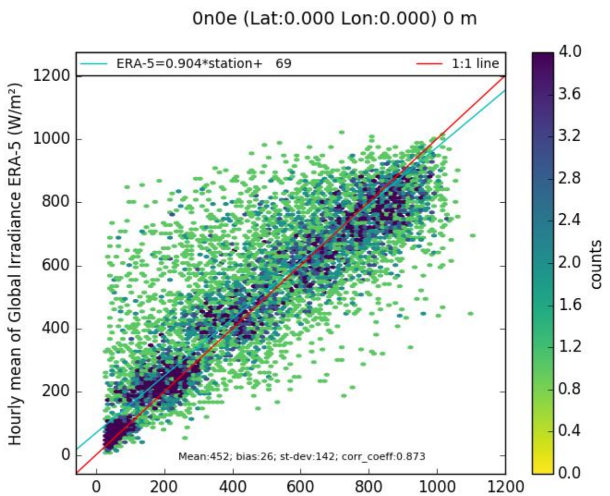
Re-analysis data sets



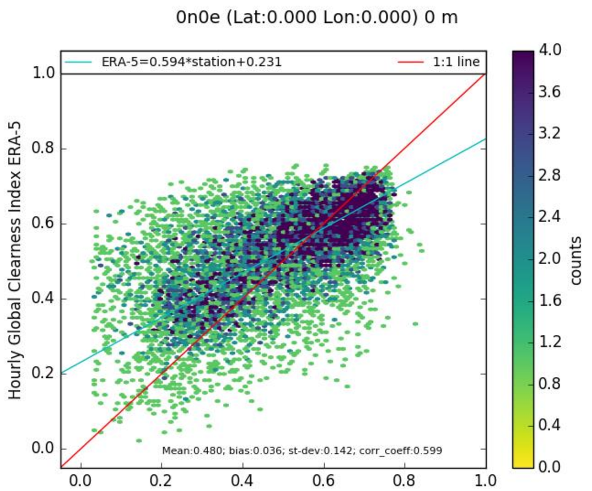
(a)



(b)

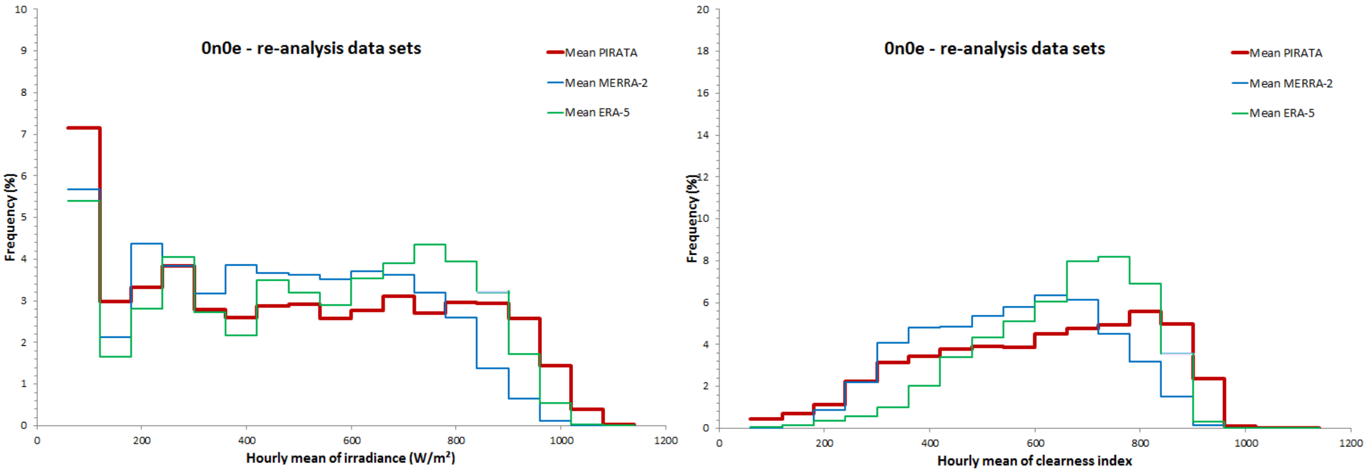


(c)

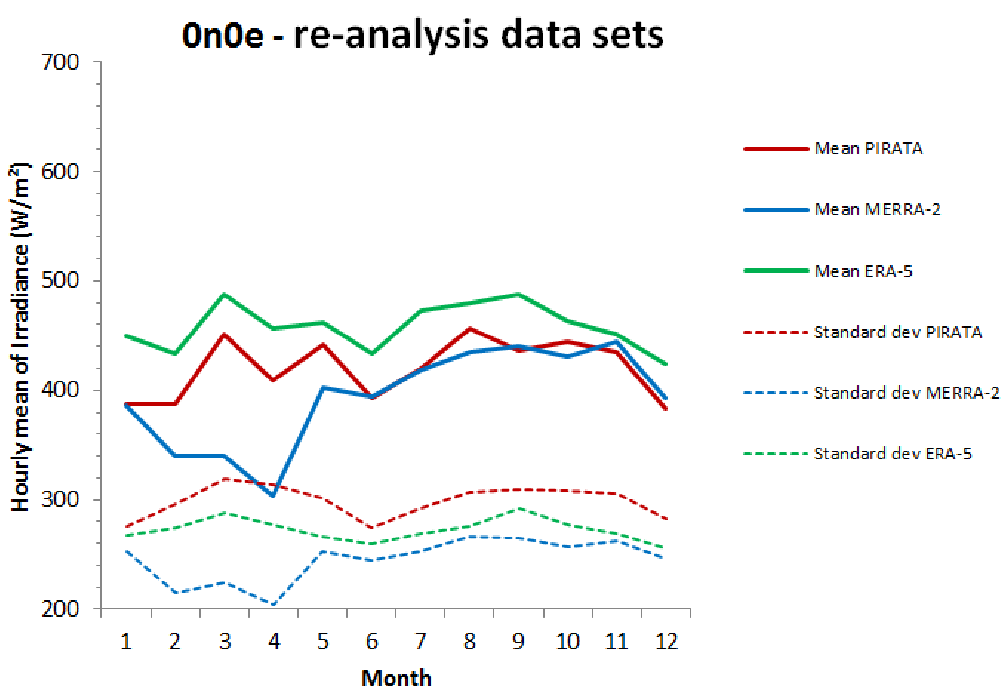


(d)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at buoy 0n0e for *E* and *KT*. MERRA-2: (a) and (b); ERA5: (c) and (d). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.



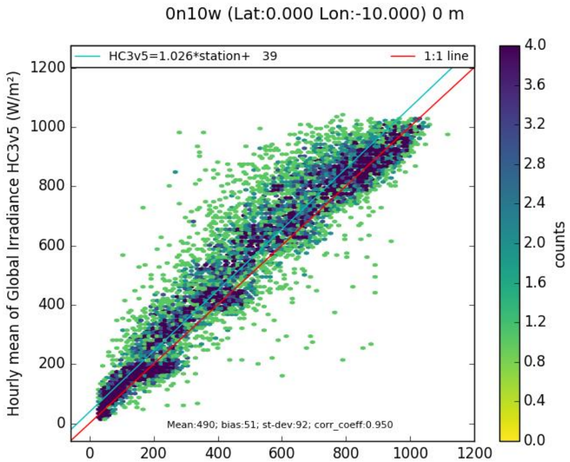
Frequency distributions of PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) for the station 0n0e for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.



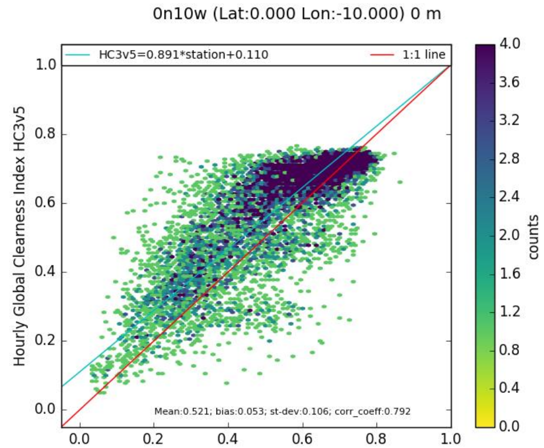
Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 0n0e. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

0n10w

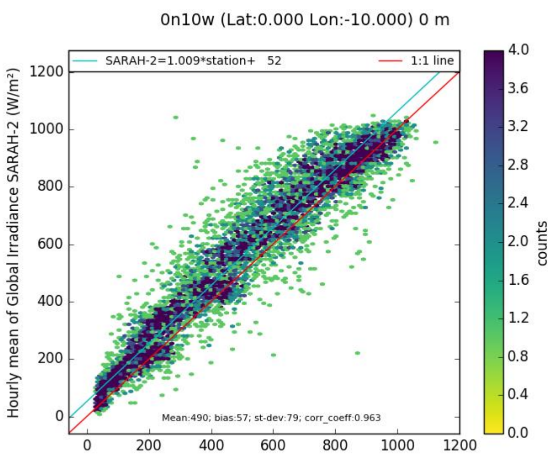
Satellite-derived data sets



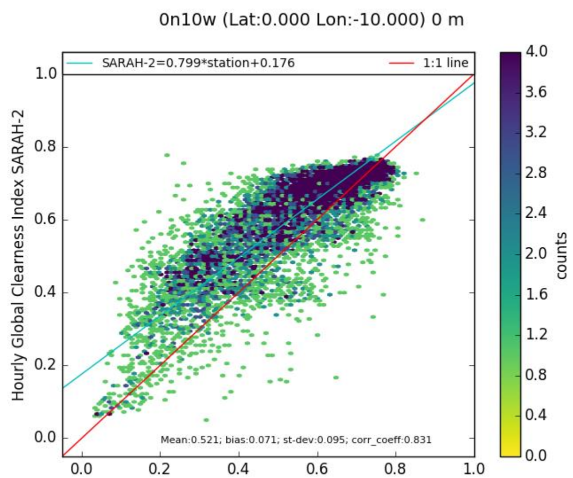
(a)



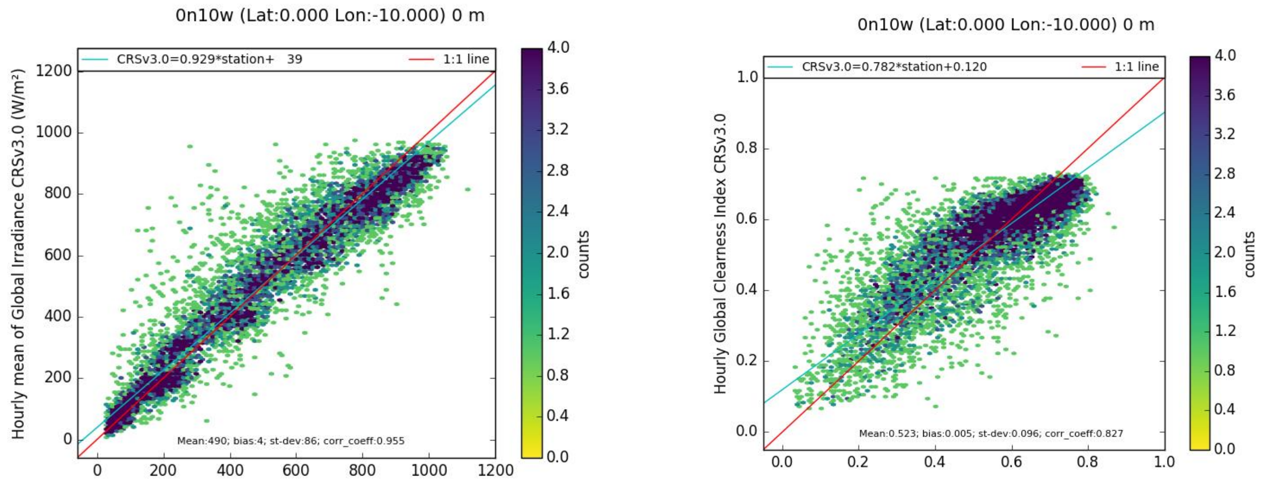
(b)



(c)

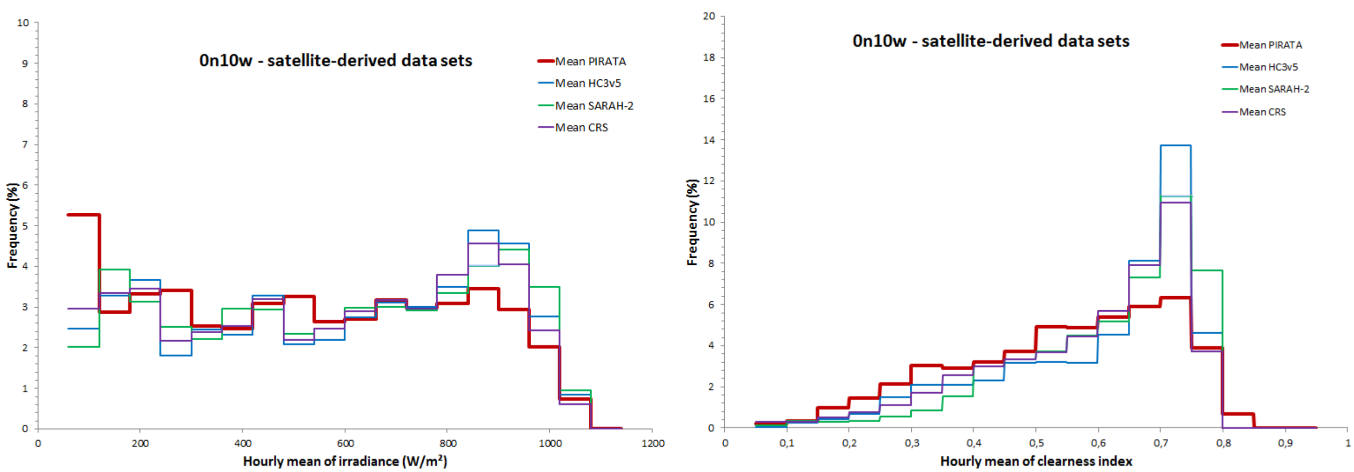


(d)

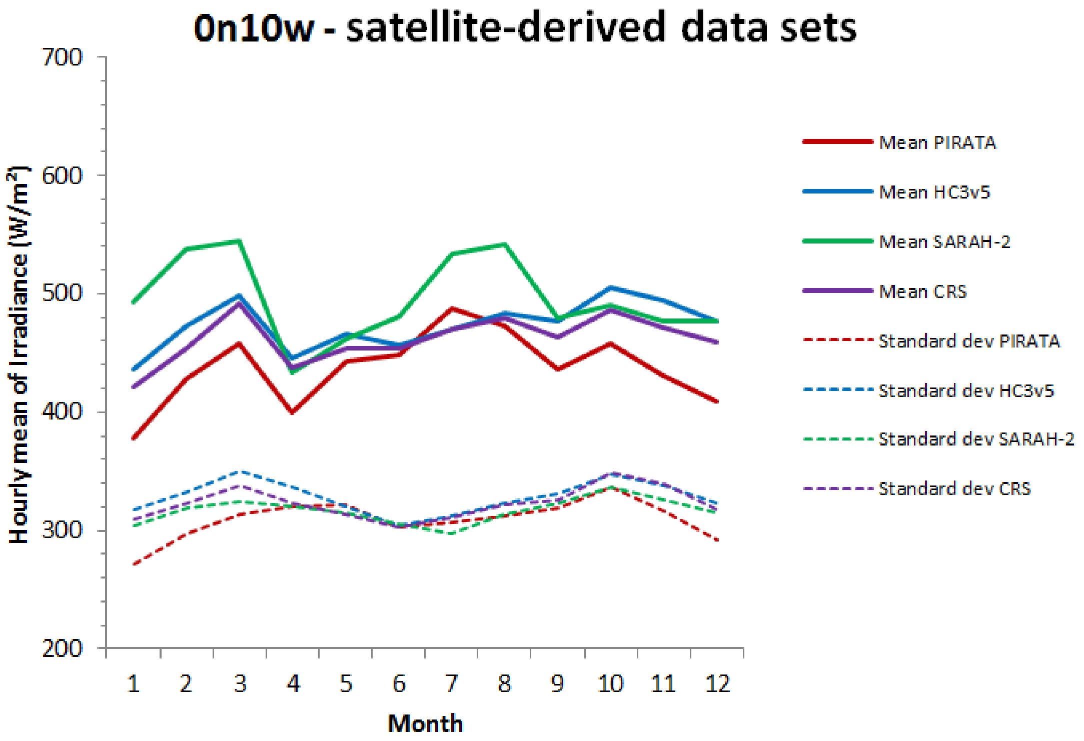


(e) (f)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) for the station 0n10w for E (left) and *KT* (right). HC3v5: (a), (b); SARAH 2: (c), (d); CRS: (e), (f). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should overlay the red line.

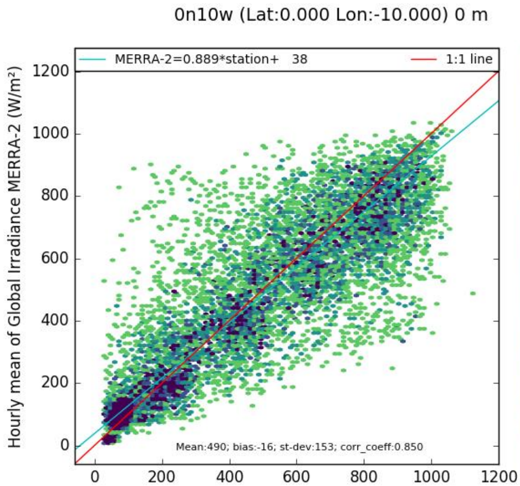


Frequency distributions of PIRATA measurements (red) and data sets (HC3v5: blue; SARAH 2: green; CRS: purple) for the station 0n10w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

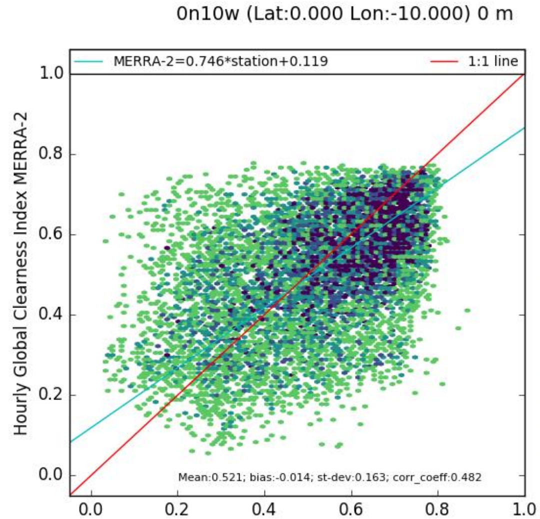


Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) at 0n10w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

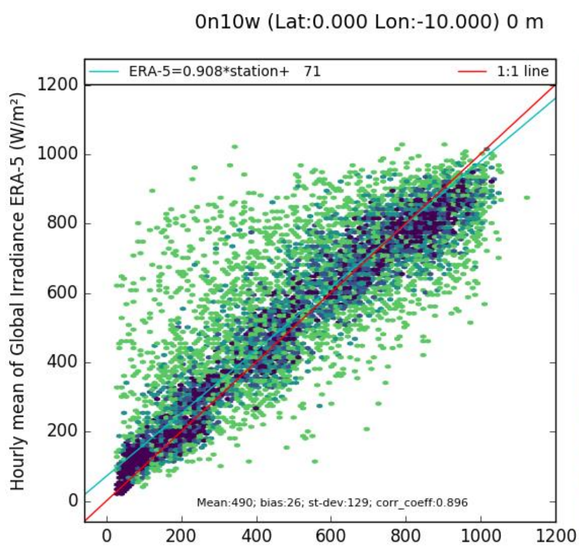
Re-analysis data sets



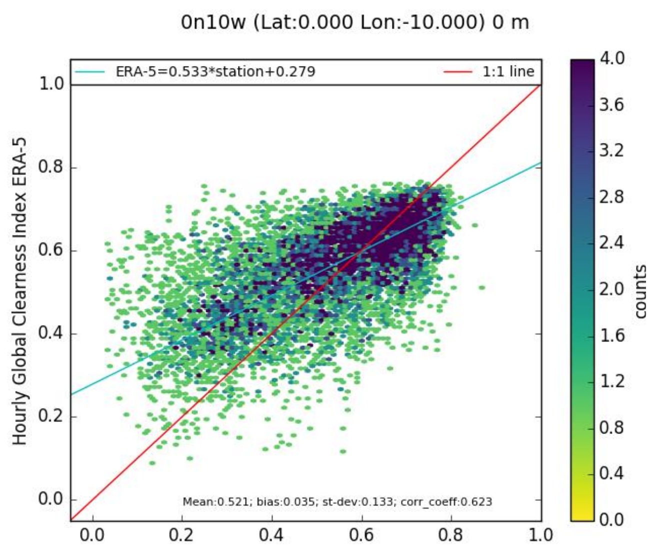
(a)



(b)

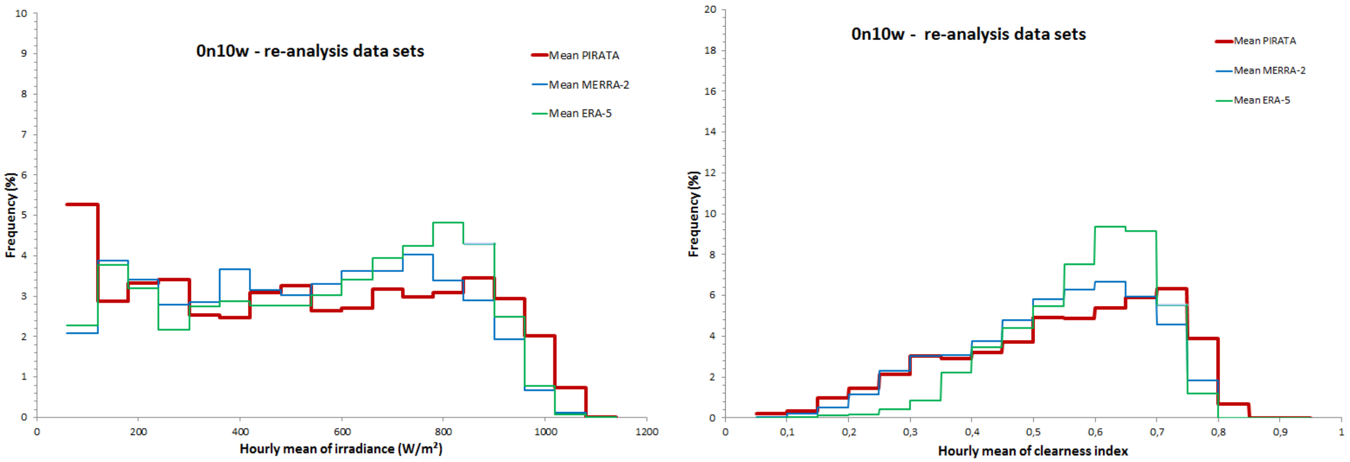


(c)

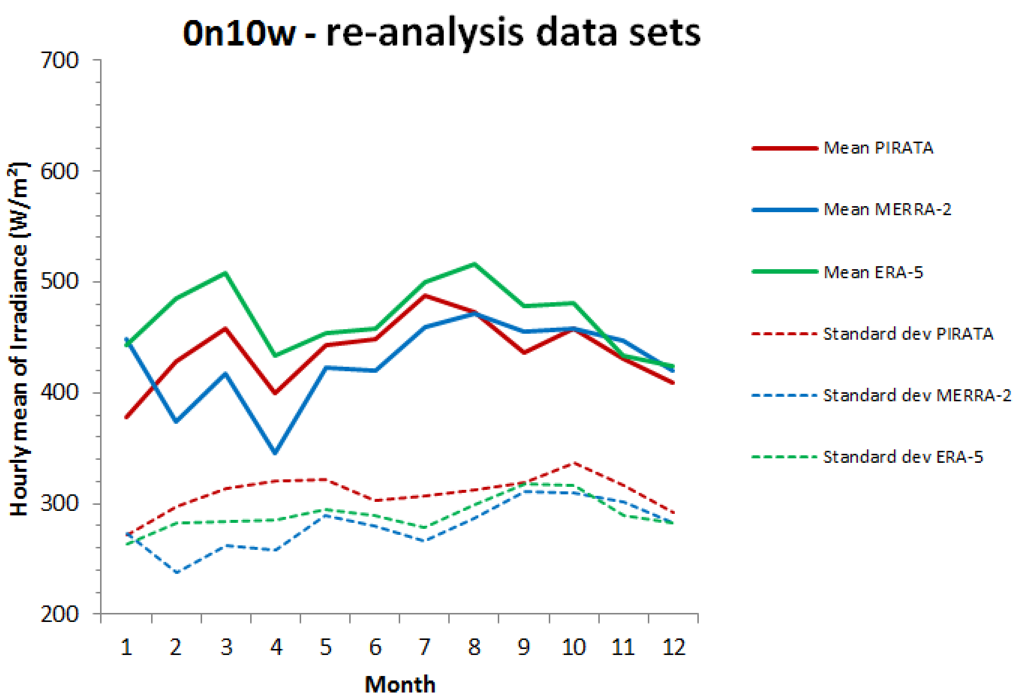


(d)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at buoy 0n10w for *E* and *KT*. MERRA-2: (a) and (b); ERA5: (c) and (d). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.



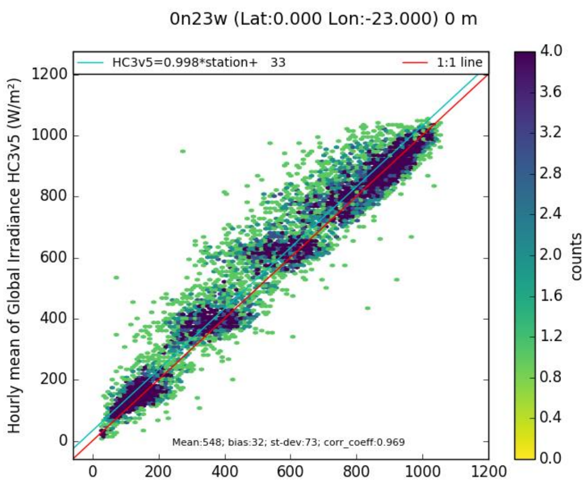
Frequency distributions of PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) for the station 0n10w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.



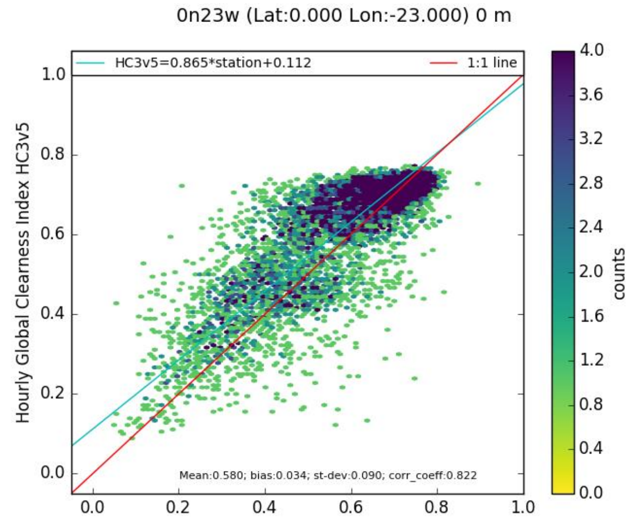
Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 0n10w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

0n23w

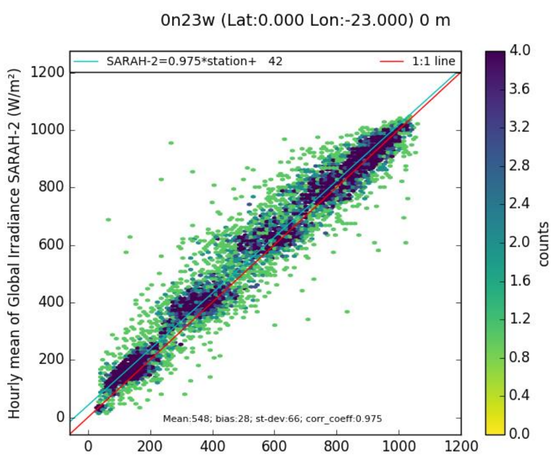
Satellite-derived data sets



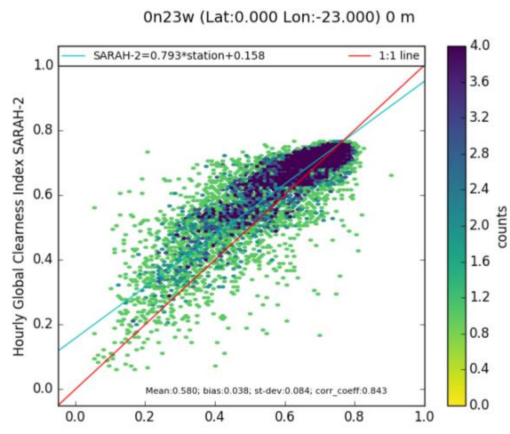
(a)



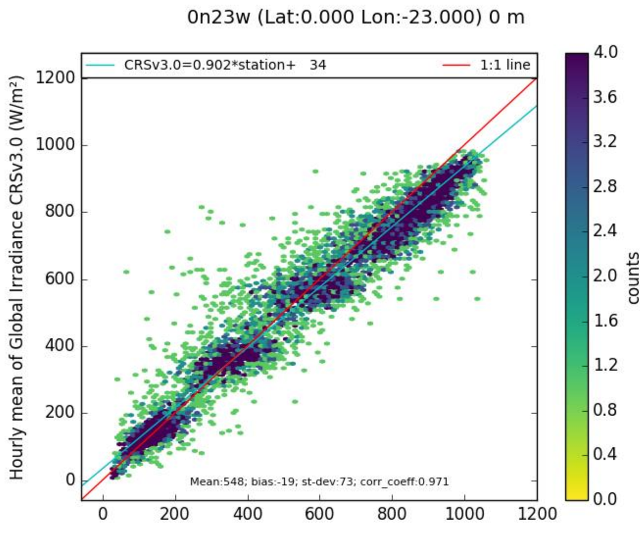
(b)

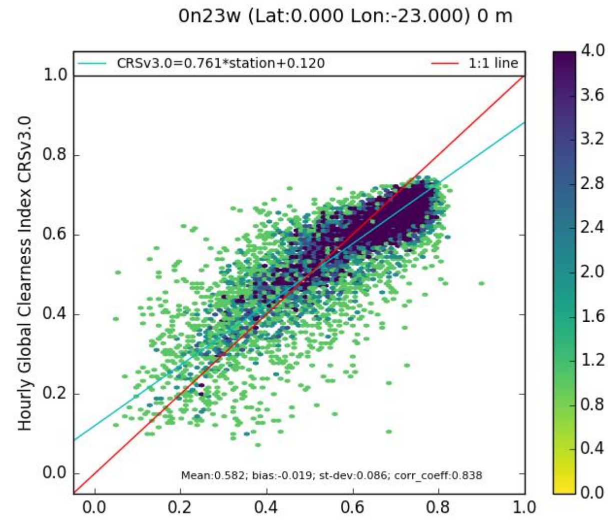


(c)



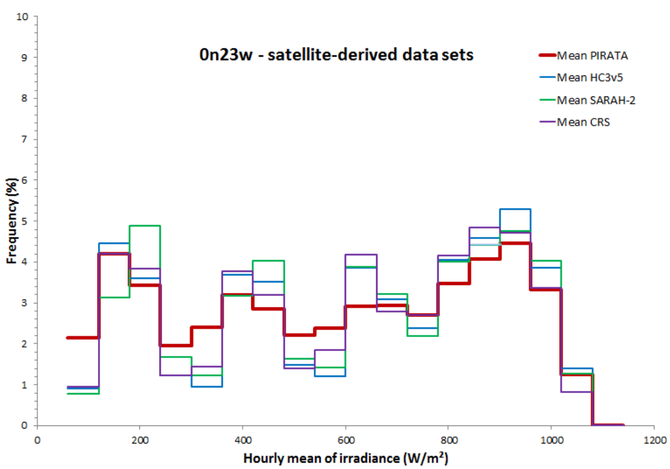
(d)

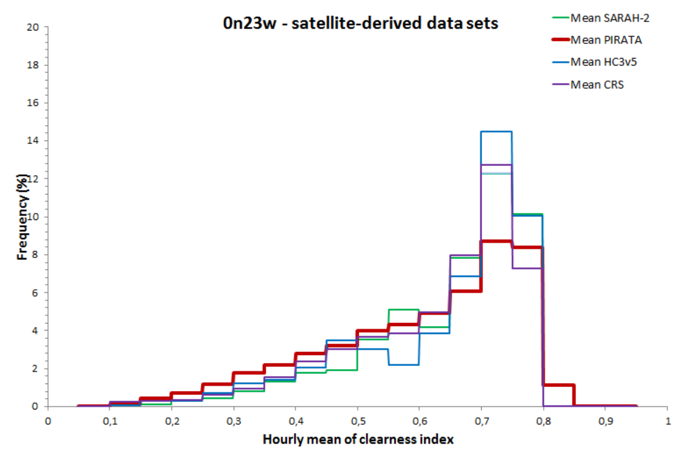




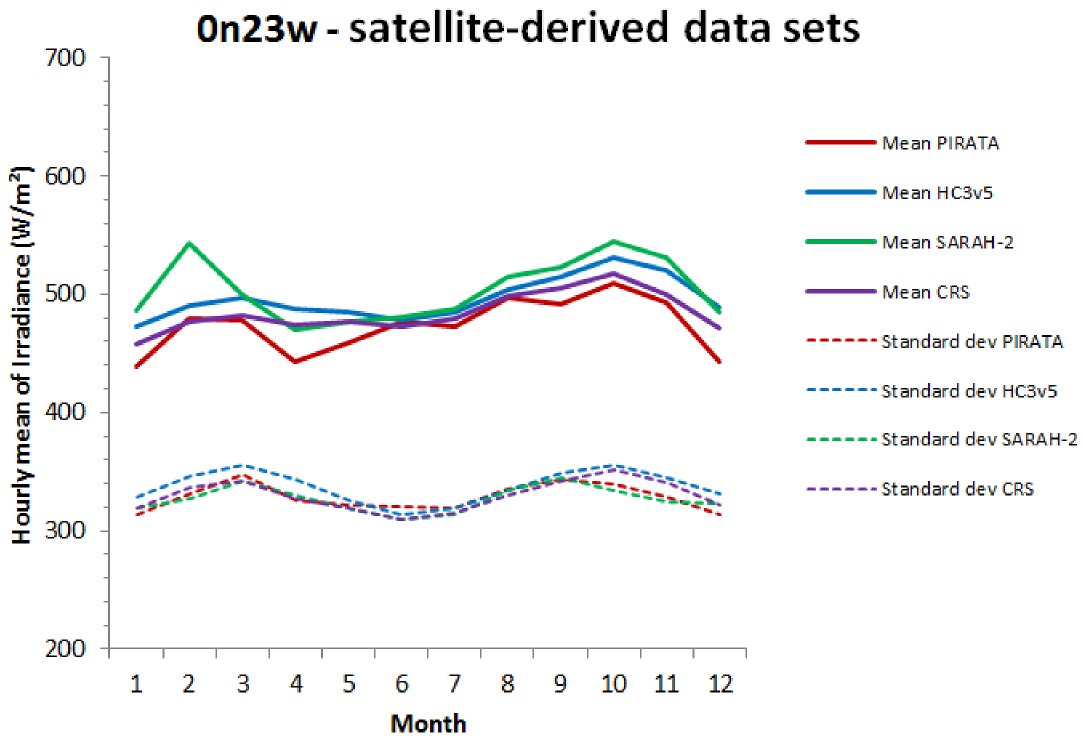
(e) (f)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) for the station 0n23w for E (left) and *KT* (right). HC3v5: (a), (b); SARAH 2: (c), (d); CRS: (e), (f). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should overlay the red line.



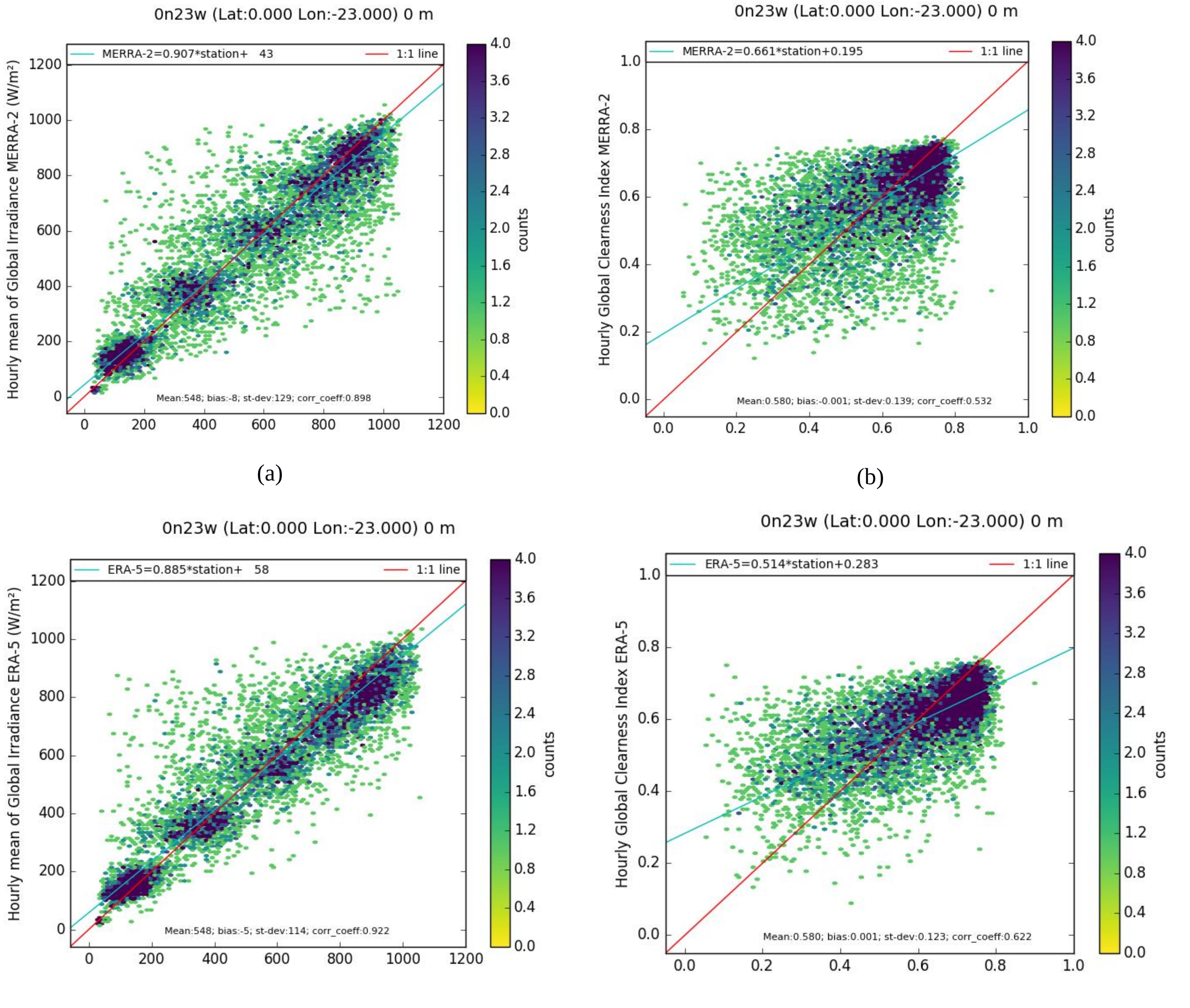


Frequency distributions of PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) for the station 0n23w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

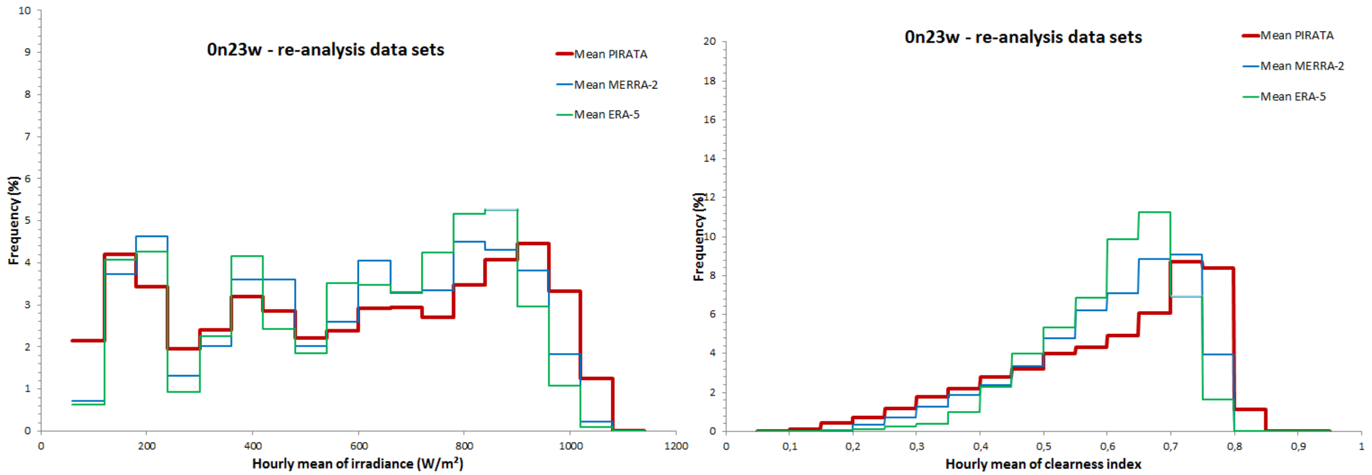


Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) at 0n23w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

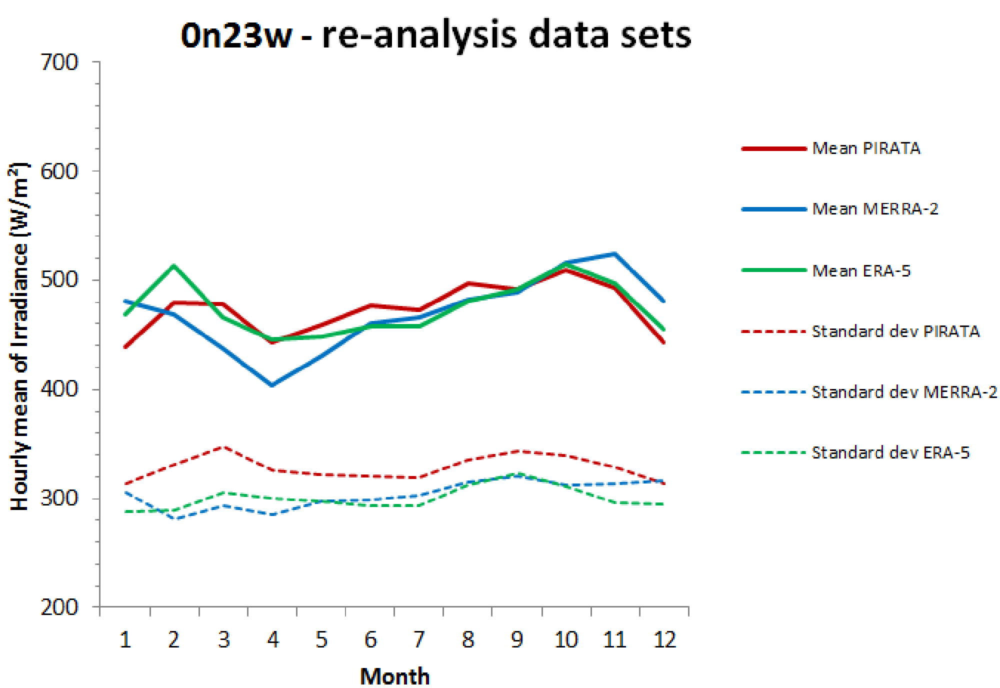
Re-analysis data sets



2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at buoy 0n23w for *E* and *KT*. MERRA-2: (a) and (b); ERA5: (c) and (d). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.



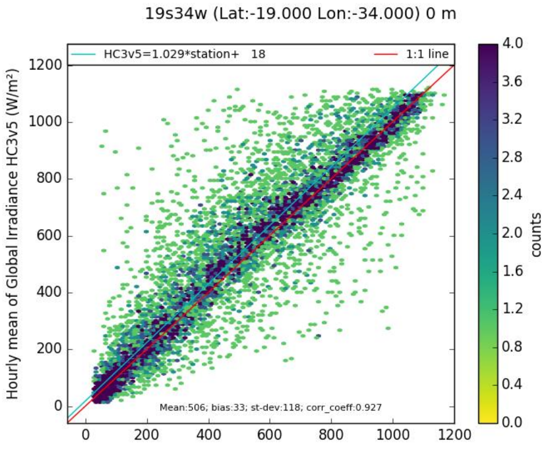
Frequency distributions of PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) for the station 0n23w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.



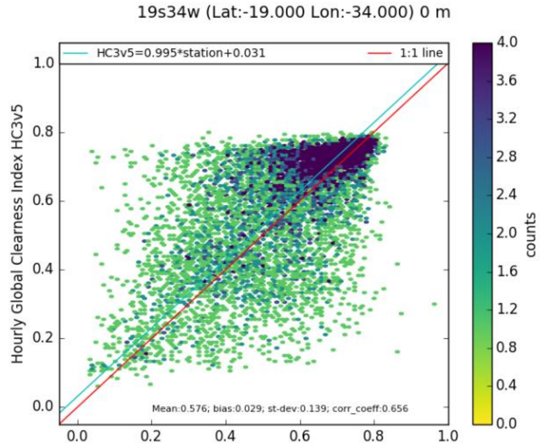
Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 0n23w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

19s34w

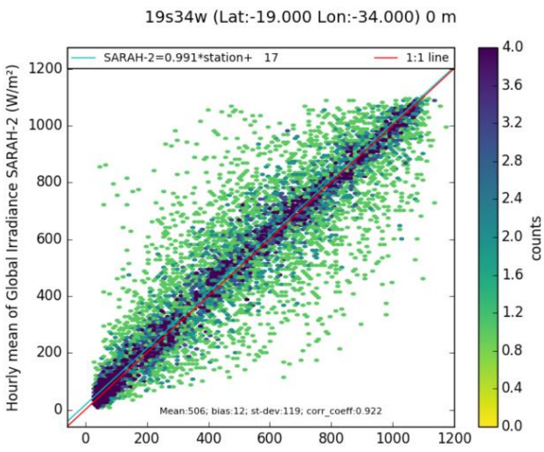
Satellite-derived data sets



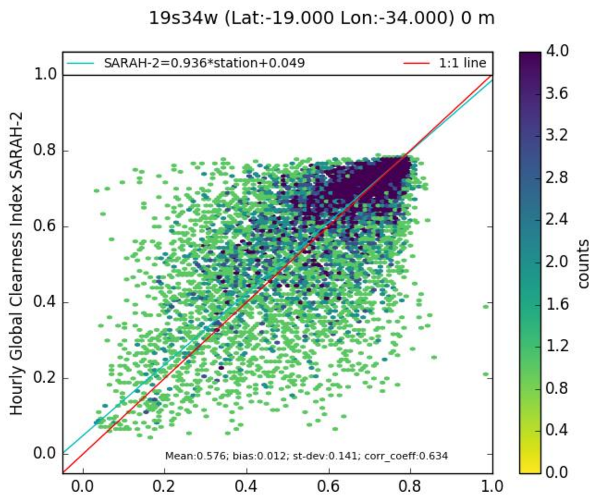
(a)



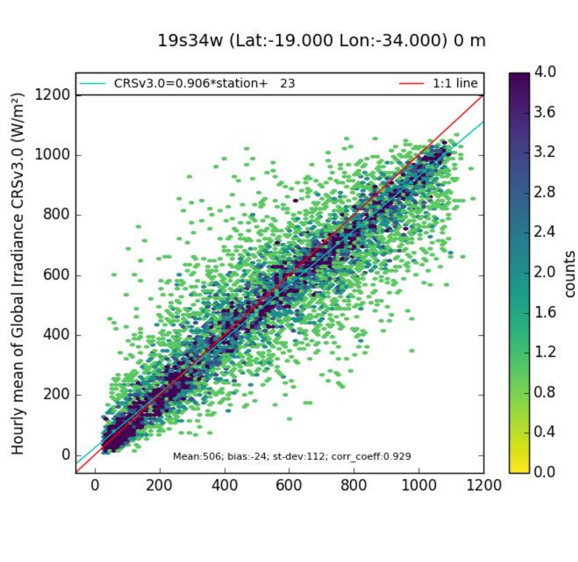
(b)



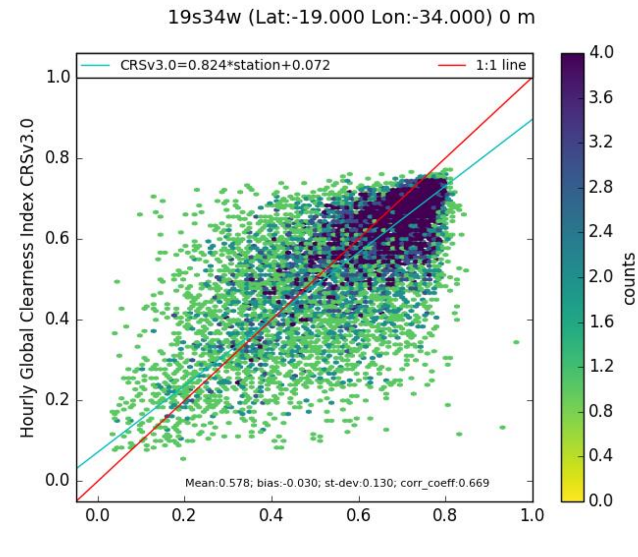
(c)



(d)

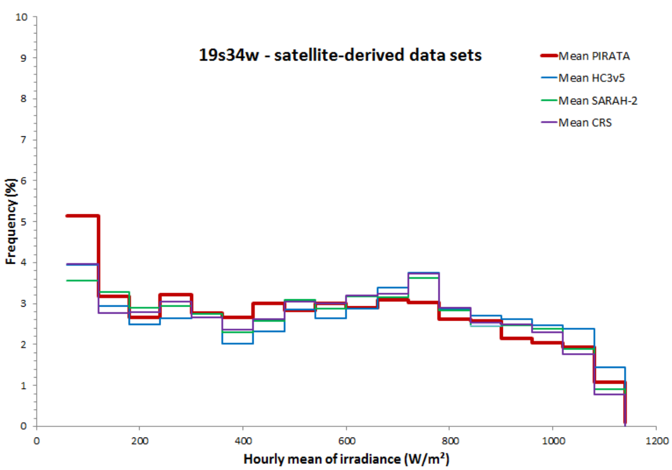


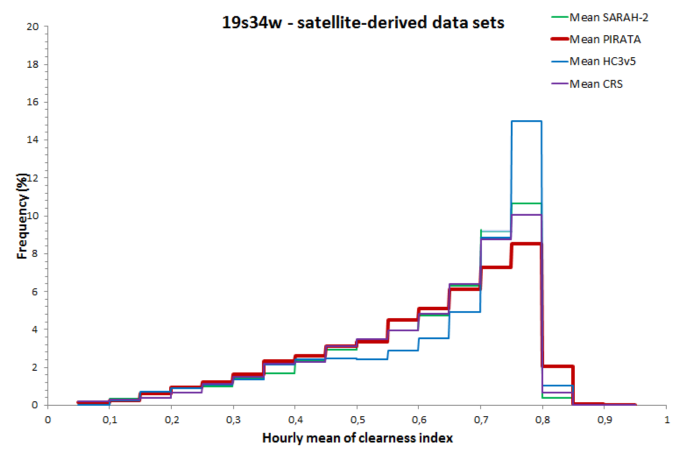
(e)



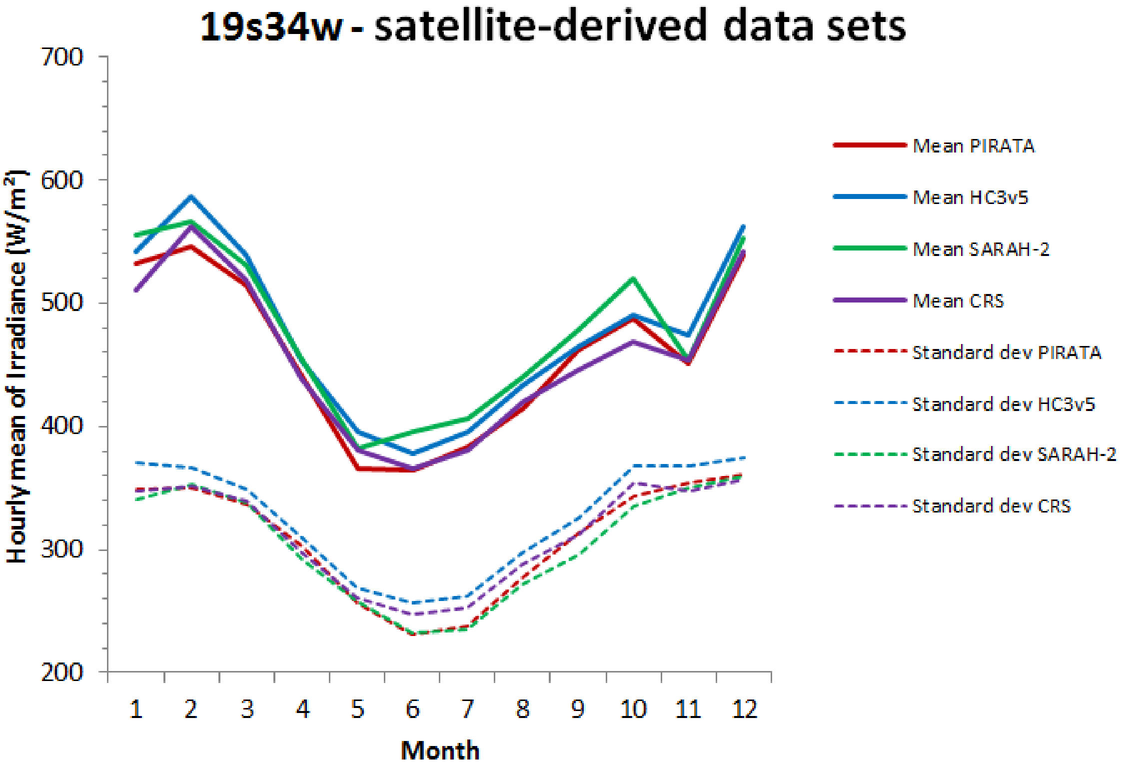
(f)

2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) for the station 19s34w for E (left) and *KT* (right). HC3v5: (a), (b); SARAH 2: (c), (d); CRS: (e), (f). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should overlay the red line.



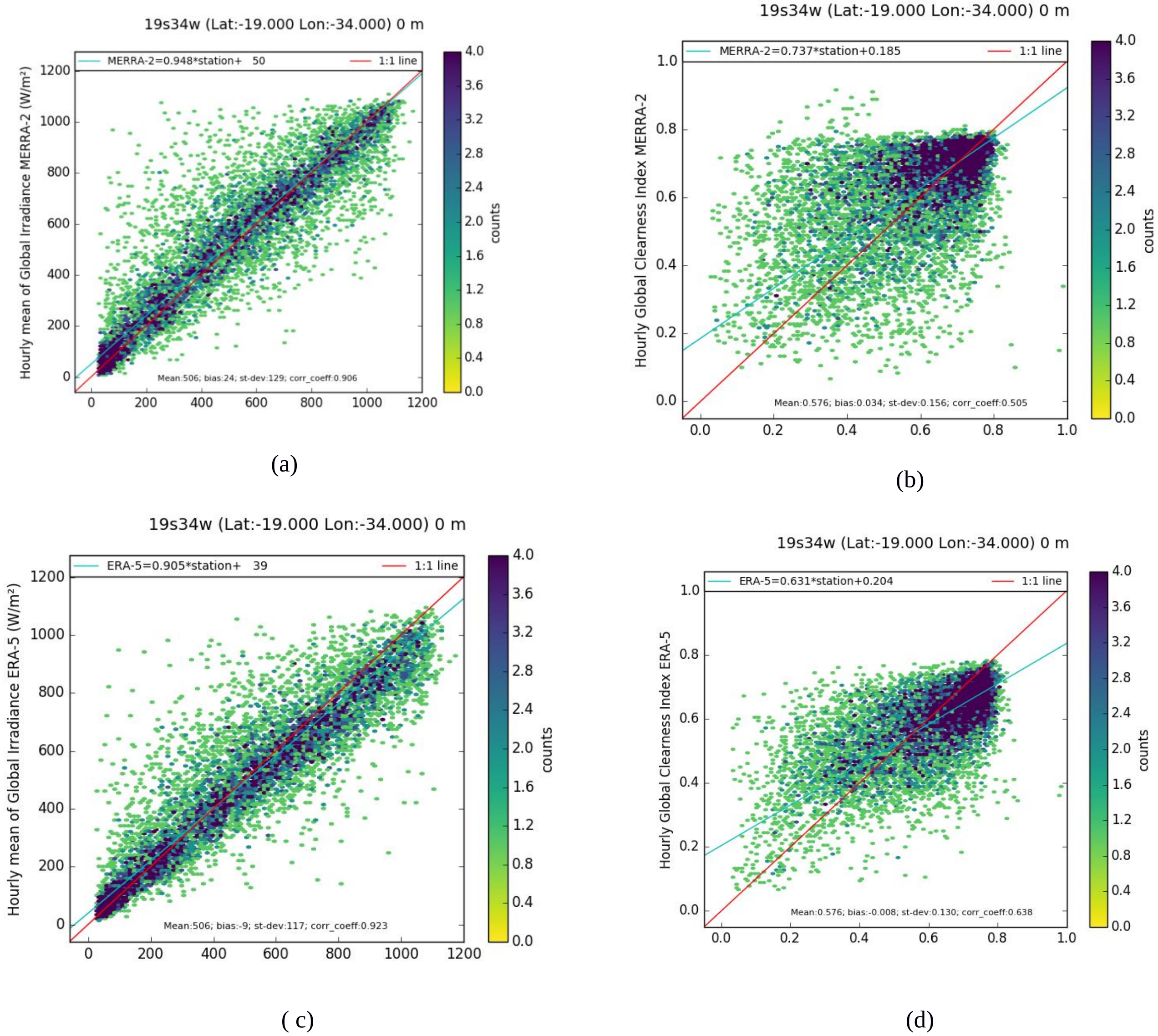


Frequency distributions of PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) for the station 19s34w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.

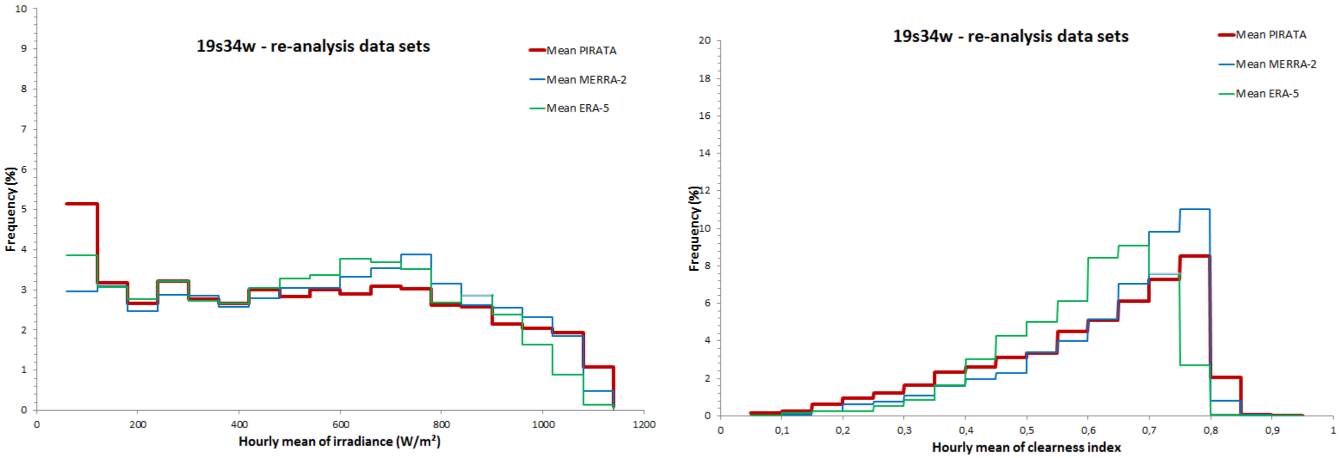


Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (HC3v5: blue ; SARAH 2: green; CRS: purple) at 19s34w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.

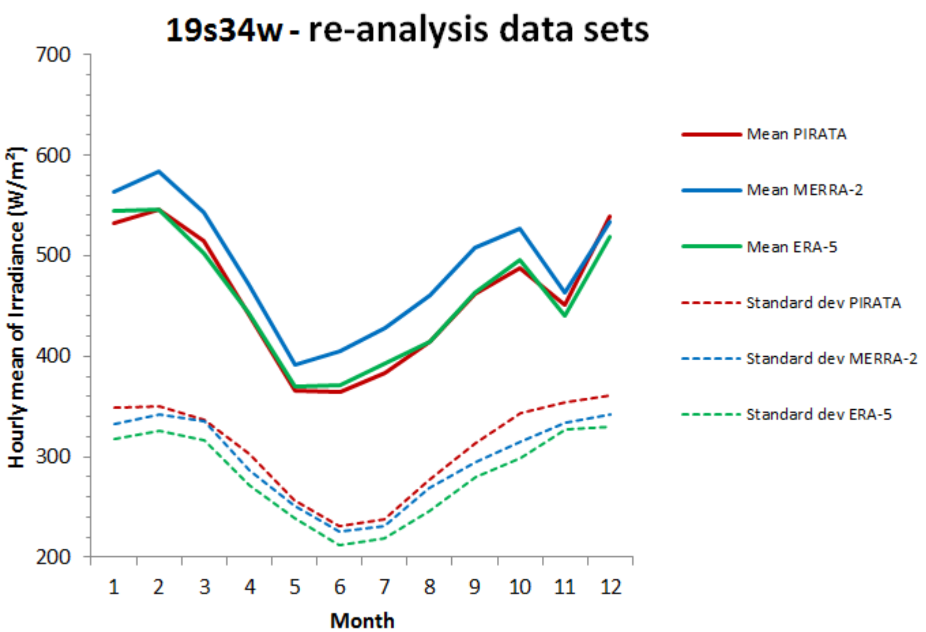
Re-analysis data sets



2D histogram of PIRATA measurements (horizontal axis) and data sets (vertical axis) at buoy 19s34w for *E* and *KT*. MERRA-2: (a) and (b); ERA5: (c) and (d). Ideally, the dots should lie along the red line (1:1 line). The blue line is the affine function fitted over the points and should ideally overlay the red line.



Frequency distributions of PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) for the station 19s34w for E (left) and KT (right). If the coloured line is above, respectively below, the red one for a given sub-range of values, it means that the data set produces these values too frequently, respectively too rarely with respect to the PIRATA measurements.



Monthly means (line) and standard deviations (dotted line) of hourly DSIS, in W m^{-2}, from PIRATA measurements (red) and data sets (MERRA-2: blue, ERA5: green) at 19s34w. A difference between red line (measurements) and coloured line (data set) for a given month denotes a systematic error for this month: underestimation if the coloured line is below the red line, overestimation otherwise. For a given month, a coloured dotted line above the red one means that the data set produces too much variability for this month; in the opposite case, the data set does not contain enough variability.