Remote Sensing of Environment 118 (2012) 186–198

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Remote sensing of solar surface radiation for climate monitoring — the CM-SAF retrieval in international comparison

R. Posselt [b], R.W. Mueller [a,*], R. Stöckli [b], J. Trentmann [a]

[a] Climate Monitoring Satellite Application Facility, German Meteorological Service, Germany
[b] Climate Monitoring Satellite Application Facility, MeteoSwiss, Switzerland

ARTICLE INFO

Article history:
Received 6 May 2011
Received in revised form 16 November 2011
Accepted 17 November 2011
Available online 22 December 2011

Keywords:
Solar surface irradiance
Satellite based climate monitoring
Radiative transfer modeling
Interactions with atmosphere (clouds aerosols, water vapor) and land/sea surface
Remote sensing
Reanalysis

ABSTRACT

Solar surface irradiance (SIS) and direct (SID) irradiance as well as effective cloud albedo (CAL) climate data records (CDR) derived from the Meteosat first generation satellites (Meteosats 2 to 7, 1983–2005) are presented. The CDRs are available free of charge for all purposes from wui.cmsaf.eu at monthly, daily and hourly means at a spatial resolution of 0.03°.

The processing employed a climate version of the Heliosat algorithm combined with a clear sky model using an eigenvector look-up table method. Modifications to the Heliosat method include a self-calibration algorithm as well as a running mean based clear sky retrieval algorithm.

The datasets are validated using ground based observations from the Baseline Surface Radiation Network (BSRN) as a reference. The validation threshold for the mean absolute bias between satellite-derived and surface-measured radiation is given by the target accuracy for solar irradiance fields defined by the Global Climate Observing system (GCOS) and a measurement uncertainty for the surface data. The results demonstrate that the target accuracy is achieved for monthly and daily means. Furthermore, an intercomparison with similar datasets reveal a better performance and climate monitoring potential of the CM SAF SIS CDR at most BSRN sites compared to established data sets like e.g. ERA-reanalysis, GEWEX (Global Energy and Water Cycle Experiment) and ISCCP (International Satellite Cloud Climatology Project). Lastly, the realistic representation of both seasonal and inter-annual variability guarantees the applicability of the satellite-based climate data sets for climate monitoring and analysis of extremes.

No trends in the normalized bias between the CM SAF and the BSRN datasets are detectable, which demonstrates the stability and homogeneity of the global and direct irradiance for the period covered by BSRN measurements. However, inconsistencies are detectable at few satellite transition dates for certain regions in earlier years.

1. Introduction

The radiation budget at the Earth's surface is a key parameter for climate monitoring and analysis. Satellite data allow the retrieval of the surface radiation budget with high spatial and temporal resolution and a large areal coverage (up to global). Through radiometric measurements satellite sensors provide information on the interaction of solar radiative fluxes with the atmosphere and the Earth's surface. This information is the basis for the retrieval of surface radiation datasets.

A 23 year long (1983–2005) continuous and validated climate data record (CDR) of surface solar irradiance (SIS), direct irradiance (SID) and effective cloud albedo (*CAL*) has recently been generated by the Satellite Application Facility on Climate Monitoring (CM SAF). CM SAF is part of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facilities (SAFs) network. Within CM SAF, special emphasis is placed on the generation of satellite-derived data records for climate monitoring (Schulz et al., 2009).

The presented SIS, SID and *CAL* CDR is based on the visible channel (0.45 − 1 μm) of the MVIRI (METEOSAT Visible and Infrared Imager) instruments on-board the Meteosat First Generation (MFG) satellites. The processing employed a climate version of the Heliosat algorithm (Beyer et al., 1996; Cano et al., 1986) which includes a self-calibration method and an improved algorithm for the determination of the clear-sky reflectivity. Both modifications to the original Heliosat method are presented in Section 2.2 below. The purpose of the self-calibration method is to automatically account for the degradation of the individual satellite instruments during their lifetime and the discontinuities induced by changes between different satellite instruments in the generation of the climate data record.

The CM SAF SIS, SID and *CAL* datasets are presented in Fig. 1 as seasonal means on the full disk. Within the annual cycle the datasets show the correct patterns with the highest radiation values in regions with highest sun elevation and lowest values in the winter hemispheres (lowest sun elevation). Furthermore, the shadowing effect of clouds on radiation is very well depicted (especially for SID) in the stratocumulus region close to the western, south African coast and in the tropics with the large amount of cumulus clouds.

* Corresponding author.
E-mail address: richard.mueller@dwd.de (R.W. Mueller).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2011.11.016

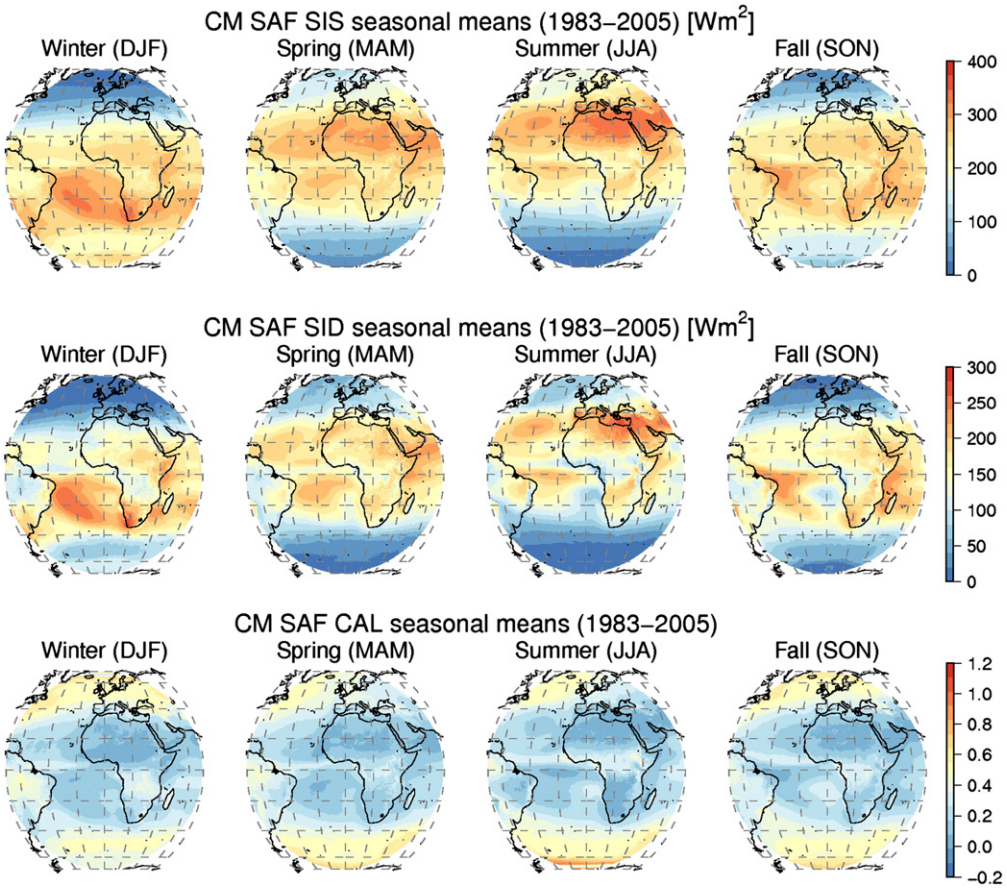

Fig. 1. Seasonal means of SIS (upper row), SID (middle row) and *CAL* (lower row) for the whole CDR (1983–2005).

For the solar surface irradiance SIS a number of similar datasets exist. The HelioClim project (Cros et al., 2004; Lefèvre et al., 2007) also applies a Heliosat algorithm to the MFG satellite data to obtain SIS on the Meteosat disk from 1985 to 2005. Furthermore, the flux dataset from the International Satellite Cloud Climatology project (Rossow & Dueñas, 2004) and the surface radiation budget dataset from the Global Energy and Water Cycle Experiment (Gupta et al., 2006) provide global SIS in a time range from 1983 to 2005. Last but not least, the model-based reanalysis dataset ERA-Interim from the European Center for Medium-Range Weather Forecasts (Berrisford et al., 2009) does also provide global SIS fields starting in 1989. These four datasets are used as intercomparison datasets for the CM SAF SIS dataset. They are, thus, further described in Section 2.4 and the results of the intercomparisons are presented in Section 4.1.

The following Section 2 describes the Heliosat algorithm used to derive the surface radiation from the satellite measurements including the applied modifications. Section 3 presents the influences of the modifications applied to Heliosat in comparison to the unmodified Heliosat. Section 4 presents the validation results by directly comparing the results to surface measurements from the Baseline Surface Radiation Network (BSRN, Ohmura et al., 1998) and to four intercomparison datasets. The last Section 5 summarizes the results and concludes this publication.

2. Methods and data

2.1. Heliosat

The solar surface irradiance is derived from the geostationary satellite measurements in a two-step approach. First, the Heliosat algorithm uses the reflectance or radiance measurements to determine the effective cloud albedo *CAL* (also denoted as cloud index *n* in previous literature) (Beyer et al., 1996; Cano et al., 1986; Hammer et al., 2003). In a second step a clear sky model is used to calculate the clear sky solar surface (or global) and clear sky direct irradiance (SIS_{cs} and SID_{cs}). SIS and SID are finally determined by combining the clear sky irradiances with the retrieved effective cloud albedo.

Traditionally, the Heliosat algorithm uses the digital count *D* of the visible satellite channel. It neither depends on any calibration information nor on information from other channels. The normalized counts ρ are derived from *D* accounting for the dark offset, the solar zenith angle and the Sun–Earth distance. The effective cloud albedo is defined by the relation of the current satellite measurement and the corresponding observations under clear-sky conditions for the same satellite pixel. The brighter the pixel the more or thicker clouds are present.

The effective cloud albedo is derived from the normalized counts ρ, the clear sky normalized counts ρ_{cs} and the normalized counts of a compact (not convective) cloud deck as an estimation for the maximum possible normalized count ρ_{max} by

$$CAL = \frac{\rho - \rho_{cs}}{\rho_{max} - \rho_{cs}}. \quad (1)$$

In general, *CAL* takes values between 0 (clear-sky) and 1 (overcast). However, certain combinations of ρ, ρ_{cs} and ρ_{max} result in *CAL* values outside of this range. $CAL < 0$ might occur if either ρ or ρ_{max} is smaller than ρ_{cs}. The first case can for instance occur in areas with melting snow where the evolution of ρ_{cs} lags the actual state of ρ_{cs}. This effect is a pure artifact of our processing scheme and can be minimized by a fast adapting scheme for ρ_{cs} in case of a darkening of the pixel (see Section 2.2). The second case can occur over bright surfaces

like snow where ρ_{cs} can become larger than ρ_{max}. If so, a *CAL* value smaller than 0 can be understood to account for higher radiation values due to reflections from the surface. In cases of very bright convective clouds it might occur that $\rho > \rho_{max}$ so that $CAL > 1$. This implies that this clouds block the radiation more effectively than the "standard" cloud defined by ρ_{max}. This is possible as the clouds used to determine ρ_{max} are more stratiform (see Section 2.2) but this effect is bound to a maximum *CAL* value of 1.1 (see also Eq. 3).

The clear sky normalized count is determined for every pixel separately as the minimum value of normalized counts ρ during a certain time period (e.g., 1 month). The "maximum" normalized count can be determined in a similar fashion by choosing the maximum ρ per pixel and time span (Beyer et al., 1996) or by using a single value for the full disk that is dependent on the radiometer of the different satellites (Hammer et al., 2003). Here, the maximum normalized count is derived by the new self-calibration method presented in Section 2.2. Also, a revised method to determine the minimum normalized count is used in the present paper as outlined in Section 2.2.

The clear sky irradiance SIS_{cs} is calculated using an eigenvector look-up table method (Mueller et al., 2009). It is based on the libRadtran radiative transfer model (Mayer & Kylling, 2005) (http://www.libradtran.org) and enables the use of extended information about the atmospheric state. Accurate analysis of the interaction between the atmosphere, surface albedo, atmospheric transmission and the top of atmosphere albedo has been the basis for the new method, characterized by a combination of parameterizations and "eigenvector" look-up tables. The source code of the method (Mesoscale Atmospheric Global Irradiance Code — MAGIC) is available under the gnu-public license at http://sourceforge.net/projects/gnu-magic/.

SIS is derived by the combination of effective cloud albedo *CAL* and clear sky irradiance SIS_{cs}.

$$SIS = k(CAL)SIS_{cs} \tag{2}$$

Thereby, $k(CAL)$ denotes the clear sky index which relates to *CAL* through the following relation (Hammer et al., 2003):

$$k(CAL) = \begin{cases} 1.2 & \text{for } CAL \leq -0.2 \\ 1 - CAL & \text{for } -0.2 < CAL \leq 0.8 \\ 2.0667 - 3.667 \cdot CAL + 1.6667 \cdot CAL^2 & \text{for } 0.8 < CAL \leq 1.1 \\ 0.05 & \text{for } CAL > 1.1 \end{cases} \tag{3}$$

2.2. Heliosat modifications

2.2.1. Self-calibration

The self-calibration algorithm is based on the determination of the "maximum" normalized count ρ_{max} that serves as the self-calibration parameter. In analogy to Rigollier et al. (2002) a histogram of all observed normalized counts is generated. However, instead of using an upper and a lower bound as in Rigollier et al. (2002), only the 95%-percentile in a specific cloudy target region is used as the self-calibration parameter and set to ρ_{max}. The method of Rigollier et al. (2002) requires an absolute calibration of one image as a reference for the initialization of the autocalibration and the spectral response function (sensitivity) of the Meteosat sensors. Additionally, there is a need to derive the statistical quantities for the complete disk, which is computationally expensive. The use of the 95%-percentile derived in a specific cloudy target region overcomes these limitations and enables a flexible autocalibration without the need for any external calibration information or for the sensor spectral response functions. The method can also be applied to Meteosat-East and Meteosat Second Generation (SEVIRI). In general, the derived solar surface irradiance shows a remarkably good consistency, demonstrating the flexible functionality of the method (Posselt et al., 2011; T. Huld, JRC, personal communication). Our empirical analysis has shown that the use of the 95% percentile for ρ_{max} results in the highest stability of the self-calibration parameter. The use of a percentile value and not the maximum of the distribution excludes saturated pixels and convective clouds which would introduce unrealistic month-to-month variability.

The analysis of the full disk is computationally too expensive and does not improve the accuracy, hence, a regional subset of the full disk was selected. This region is located in the southern Atlantic between 15°W to 0°W and 58°S to 48°S. It features a frequent appearance of frontal systems with high cloud amounts but hardly any deep convective clouds. The self-calibration algorithm is applied to the 1300 UTC slot, which accounts for the slight westward shift of the region. The histograms are generated each month with a month long input dataset. The resulting ρ_{max} is then applied to calculate the effective cloud albedo *CAL* for that month following Eq. (1).

By definition the Heliosat method retrieves the effective cloud albedo. This circumvents the requirement to use land surface targets to calibrate the satellite radiances. In fact, by using a land surface target in the self-calibration method errors could be introduced as the aging detected by observation of clear sky radiances is different than that detected with an appropriate cloud target. This effect is due to spectral dependency of the aging of optical devices, which leads to different rates of aging for different land surface types, due to different spectral responses of the targets. In our self-calibration method the aging effects of the "clear sky" observations for different surface types is automatically considered by the retrieval of the clear sky normalized digital counts ρ_{cs}.

2.2.2. Retrieval of clear sky reflection

The retrieval of the clear sky reflection (ρ_{cs}) is one of the key requirements for the application of Heliosat when applied to MVIRI data. A recently developed algorithm for the retrieval of the clear sky reflection is used to calculate ρ_{cs} (Dürr & Zelenka, 2009; Zelenka, 2001). Instead of using a monthly field of the clear sky normalized counts ρ_{cs}, a seven day running mean of ρ_{cs} is used. Thus, changes in ρ_{cs}, e.g., due to changes in the vegetation and snow cover, are captured and represented faster than with the standard method. In addition, unrealistic steps in the effective cloud albedo between different months are avoided. Further, the presented method allows the detection of snow and, thus, provides the opportunity to correct for some of the radiative properties of snow.

The algorithm to determine ρ_{cs} is based on a time series approach by Zelenka (2001). Two tests are applied to the normalized counts (ρ) in order to determine the temporal evolution of ρ_{cs}. If ρ exceeds ρ_{cs} by up to ϵ_{up}, "Test '1" (see Eq. 4) applies a slow adaptation of ρ_{cs} to the current minimum counts. This is the case if ρ is to a certain extend larger than ρ_{cs}. An adaptation time of 7 days is assumed which corresponds to the time scale of involved processes (e.g., changes in vegetation, soil properties). Additionally, a sudden strong decrease in ρ below $\rho_{cs} - \epsilon_{low}$ also leads to a slow change of ρ_{cs}. This should prevent the contamination of ρ_{cs} with cloud shadows. If ρ falls below ρ_{cs} by up to ϵ_{low}, "Test 2" (see Eq. 5) executes a fast adaptation of ρ_{cs}. This latter scheme applies to events with relatively fast processes as, e.g., melting of snow.

If both tests fail, i.e., if ρ is rather large as occurring in case of clouds and snow, ρ_{cs} is not changed. If this happens for a time span of $t_s = 28$ subsequent days (i.e., the pixel is very bright at all times) then this pixel is assigned to be snow covered. This algorithm is based on the low temporal variability of snow compared to clouds. Thus, continuously detecting very high reflectivities in one time slot over a certain number of days indicate snow. This strategy can fail in the cases of persistent cloud decks (e.g., marine stratocumulus or fog). In case of a positive snow detection (i.e., t_s is reached) ρ_{cs} subsequently evolves according to Eq. (5) so that ρ_{cs} increases to the range

of the current minimum normalized counts and the basic scheme (i.e., Test 1 and Test 2) can be applied in the subsequent time steps.

$$\text{Test 1}: \rho_{cs,t} \leq \rho_t \leq \rho_{cs,t} + \epsilon_{up} \text{ or } \rho_t < \rho_{cs,t} - \epsilon_{low} \rightarrow \rho_{cs,t+1} = \frac{6}{7}\rho_{cs,t} + \frac{1}{7}\rho_t (\text{slow evolution}) \quad (4)$$

$$\text{Test 2}: \rho_{cs,t} - \epsilon_{low} \leq \rho_t < \rho_{cs,t} \rightarrow \rho_{cs,t+1} = \frac{1}{2}\rho_{cs_t} + \frac{1}{2}\rho_t (\text{fast evolution}) \quad (5)$$

$$\text{else}: \rho_{cs,t+1} = \rho_{cs,t} \quad (6)$$

The subscript t denotes the current and $t+1$ the subsequent time step. We use variable bandwidths ϵ_{up} and ϵ_{low} depending on ρ_{max} and ρ_{cs} given in Dürr and Zelenka (2009) instead of the fixed ones used in Zelenka (2001).

$$\epsilon_{up} = 0.125\rho_{max} + \frac{8(\rho_{cs} - 0.15\rho_{max})}{0.25\rho_{max}} \quad (7)$$

$$\epsilon_{low} = 0.0875\rho_{max} + \frac{6(\rho_{cs} - 0.15\rho_{max})}{0.25\rho_{max}} \quad (8)$$

In order to better account for snow events, a reduced $t_s = 7$ is applied in case the considered pixel already experienced a snow event in the past (Zelenka, 2001). The shorter time range is chosen to still avoid that persistent fog is mistaken for snow and that elevated sites had the time to adapt ρ_{cs} after a sudden snow fall. A past snow event is, on one hand, determined by a large amplitude of ρ_{cs} (given by $\Delta\rho_{cs} = \rho_{cs,max} - \rho_{cs,min} > 0.55\rho_{max}$) reflecting the range of the clear sky reflectivity between snow and soil/vegetation. Further, a higher probability of snow events is also given if $\rho_{cs} > \rho_{snow}$ with ρ_{snow} being an empirically derived snow threshold. It includes a term $\rho_{snow,min}$ that is high enough to exclude oceans or other dark surfaces and it includes a term $\rho_{snow,range}$ that accounts for the higher probability of a former snow event for pixels with a high amplitude ρ_{cs}.

$$\rho_{snow,min} = 0.48\rho_{max} \quad (9)$$

$$\rho_{snow,range} = max(0.15\rho_{max}, 0.6\rho_{cs}) \quad (10)$$

$$\rho_{snow} = max\left(\rho_{snow,min}, \rho_{cs,min} + \rho_{snow,range}\right) \quad (11)$$

The calculation of the effective cloud albedo *CAL* for non-snow pixels follows the standard Heliosat method (see Eq. 1). In case snow is detected the calculation of the cloud albedo *CAL* is modified according to Dürr and Zelenka (2009) in which ρ_{cs} is substituted by $\rho_{snow,min}$ and ρ_{max} is increased by a factor of 1.41 in order to artificially increase the count range and, thus, enhance the contrast. This empirically developed modified cloud albedo formulation generally gives lower cloud albedo values in order to account for the radiative properties of snow. This includes mainly the reflectivity of the bright snow surface which leads to higher surface irradiance values.

$$CAL_s = \frac{\rho - 0.48\rho_{max}}{1.41\rho_{max} - 0.48\rho_{max}} = \frac{\rho - 0.48\rho_{max}}{0.93\rho_{max}} \quad (12)$$

2.3. Satellite data

Data from EUMETSAT's geostationary Meteosat satellites of the First Generation (Meteosats 2–7) are used. They were in operation from 1982 to 2005 at a location directly over the equator at a longitude of 0° at an altitude of about 36,000 km. The resulting visible disk reaches up to 80°N/S and 80°E/W, respectively. Meteosat 7 is still operational above the Indian Ocean.

The satellites of Meteosat's First Generation carried MVIRI, a radiometer with 3 spectral bands in the visible band (VISSN — 0.45 — 1 μm), in the water vapor band (WV — 5.7 — 7.1 μm) and infrared band (IR — 10.5 — 12.5 μm). One scan of the visible disk was accomplished within 30 min at a horizontal resolution of around 2.5 km at the sub-satellite point.

The Meteosat data were obtained from Eumetsat's UMARF archive in OpenMTP. They were converted into a binary format with one file per slot to serve as input for the Heliosat algorithm. The output data (global and direct radiation and cloud index) were regridded to a regular lon–lat grid with a 0.03° spacing using a triangulation technique.

The data from the first two of the used MFG satellites (Meteosats 2 and 3) needed a special treatment before they could serve as input to the Heliosat algorithm. The visible picture of the MFG satellites consisted of two half-pictures from two sensors that measured concurrently two adjacent lines. Due to limited transmission bandwidth, Meteosats 2 and 3 transferred the full visible image only every second slot. For every other slot only one of the half-images was transmitted. This leads to missing lines in the visible image. These were filled by copying the adjacent existing line onto it to generate a full visible image for the Heliosat input. Secondly, the data of Meteosats 2 and 3 were encoded with 6-bits instead of 8-bits as the later satellites. EUMETSAT already converted this data to 8-bit by multiplying it with 4. An uncertainty estimation of these two measures will be presented in Section 4.

2.4. Validation and intercomparison data

The validation of the new data sets for the surface incoming solar radiation (SIS) and the surface incoming direct solar radiation (SID) is performed by comparison with high-quality ground based measurements from the Baseline Surface Radiation Network (Ohmura et al., 1998). The BSRN stations used for the validation are listed in Table 1 and their geographical positions are shown in Fig. 2. Thereby, only those stations were used that have an overlap of at least 12 months with the satellite data. The selected 12 stations are located mainly on the northern hemisphere but they cover the main climatic regions and they span a substantial part (1992–2005) of the satellite time period. The cloud albedo (*CAL*) as a pure satellite product cannot be validated by comparison with ground based measurements directly. As the cloud albedo is the underlying driver for the calculation of SIS, the accuracy evaluated for SIS can be used to indirectly infer the accuracy of the cloud albedo.

To derive monthly and daily mean values from the surface measurements, the hourly means of 1 month were calculated first to reduce the impact of missing data values on the averaging. These hourly mean values are then averaged to derive the monthly and daily mean radiation data to be used as the reference value for the validation.

Table 1
List of used BSRN stations for the validation.

Station	Country	Code	Latitude [°N]	Longitude [°E]	Elevation [m]	Data since
Bermuda	Bermuda	ber	32.27	−64.67	8	1.1.1992
Camborne	UK	cam	50.22	−5.32	88	1.1.2001
Carpentras	France	car	44.05	5.03	100	1.8.1996
De Aar	South Africa	daa	−30.67	23.99	1287	1.5.2000
Florianopolis	Brasil	flo	−27.53	−48.52	11	1.6.1994
Lerwick	UK	ler	60.13	−1.18	84	1.1.2001
Lindenberg	Germany	lin	52.21	14.12	125	1.9.1994
Payerne	Switzerland	pay	46.81	6.94	491	1.9.1992
Sede Boger	Israel	sbo	30.9	34.78	500	1.1.2003
Solar Village	Saudi Arabia	sov	24.91	46.41	650	1.8.1998
Tamanrasset	Algeria	tam	22.78	5.51	1385	1.3.2000
Toravere	Estonia	tor	58.25	26.46	70	1.1.1999

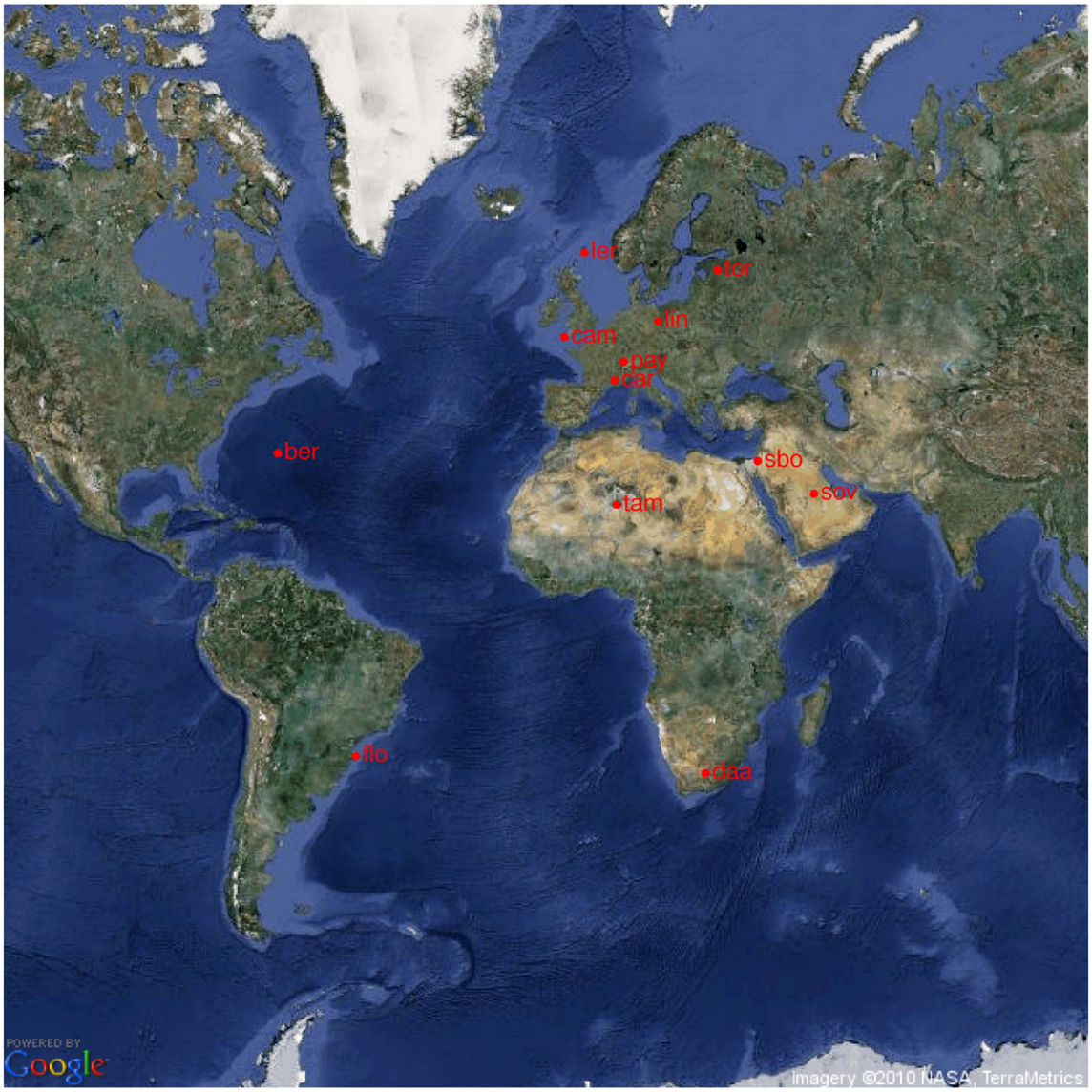

Fig. 2. Map of used BSRN stations for the validation.

CM-SAF products and data sets are reviewed by external scientists. A successful review is the pre-requisite for the release of the data under the Eumetsat SAF umbrella. The review aims to guarantee the completeness and quality of the documentation and an appropriate accuracy of the data sets. For the latter purpose certain accuracy limits (referring to the Mean Absolute Bias — MAB) are employed for SIS and SID as well as for daily and monthly means which are summarized in Table 2. They are based on the target accuracy defined in the CM SAF CDOP Product Requirements Document (www.cmsaf.eu). At least the threshold accuracy has to be reached for a successful release. Furthermore, the validation accounts for the non-systematic error of the BSRN data of 5 W m^{-2} for solar irradiance measurements (Ohmura et al., 1998).

Table 2
Validation threshold (Th), target (Ta) and optimal (Op) accuracy (all referring to Mean Absolute Bias — MAB) for monthly mean as well as for daily mean SIS, SID and *CAL*.

	SIS [W m^{-2}]			SID [W m^{-2}]			*CAL* [−]		
	Th	Ta	Op	Th	Ta	Op	Th	Ta	Op
Monthly	15	10	8	20	15	12	0.15	0.1	0.05
Daily	25	20	15	30	25	20	0.2	0.15	0.1

In addition to the validation with surface measurements, the quality of the CM SAF surface radiation SIS CDR is evaluated with already available datasets. Unfortunately, to the best of our knowledge, no comparable data sets for SID are available.

These datasets include data from the HelioClim-Project (http://www.soda-is.com; Cros et al., 2004; Lefèvre et al., 2007). HelioClim also employs a Heliosat algorithm to the data from the visible channel of the MFG satellites. In combination with the clear-sky models from the European Solar Radiation Atlas (Rigollier et al., 2000; Scharmer & Greif, 2000), SIS is obtained. This HC-1 database covers the time of the MFG satellites (1985–2005) at a reduced resolution following the ISCCP-B2 (Schiffer & Rossow, 1985) resolution on the Meteosat disk.

The flux dataset from the International Satellite Cloud Climatology Project ISCCP FD (http://isccp.giss.nasa.gov/products/products.html; Rossow & Dueñas, 2004) and the surface radiation budget dataset from the Global Energy and Water Cycle Experiment (Gupta et al., 2006) are used for the intercomparison. Both datasets are based on the same cloud cover data retrieved within ISCCP from geostationary and polar orbiting satellites which allow a global coverage. Differences between the datasets originate from the application of different

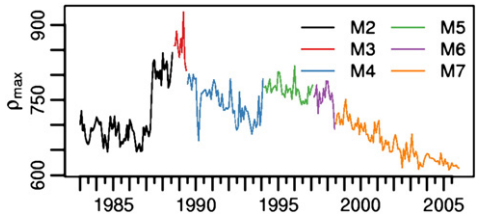

Fig. 3. Evolution of ρ_{max} in time, colored according to the operational satellite.

algorithms to derive the clear sky solar surface irradiance — ISCCP used the radiative transfer model from the NASA GISS climate model (Zhang et al., 2004) whereas GEWEX used an algorithm developed by Pinker and Laszlo (1992). The spatial resolution also differs — with 2.5° in the case of the ISCCP FD dataset and 1° for the GEWEX SRB dataset. Daily means of the data were used for the comparison in the time range from 1983 to 2005.

The model-based reanalysis dataset ERA-Interim from the European Center for Medium-Range Weather Forecasts (ECMWF http://www.ecmwf.int; Berrisford et al., 2009) also provide SIS. It spans the time range between 1989 and the present. The resolution of the data is 1° and daily means covering the time range from 1989 to 2005 were generated from the 6-hourly model output available from the ECMWF archive. Monthly means for each data set were calculated by averaging the corresponding daily means.

For the intercomparison, the four datasets were statistically analyzed in the same way as the CM SAF data set. Here, the same monthly/daily mean values from the surface observations were used, i.e., only those data are considered that are spatially and temporally covered by the BSRN and CM SAF dataset.

3. Influence of the Heliosat modifications

3.1. Self-calibration

Fig. 3 shows the temporal evolution of ρ_{max} for the MVIRI period from 1983 to 2005 for the different MVIRI satellites. It can be seen that ρ_{max} decreases substantially with the operation time of a satellite (e.g., Meteosat 7) which reflects the degradation of the instrument with time. Gain changes in the instrument are also visible as seen for Meteosat 2. However, the differences of ρ_{max} for the times of satellite changes are comparably small. This shows that the self-calibration is able to capture the sensitivity changes (i.e., calibration) within the lifetime of the used instrument and, thus, ρ_{max} can serve as parameter to calibrate the time series toward homogeneity.

3.2. Clear sky retrieval

The changes in the effective cloud albedo due to the modifications in the derivation of the clear sky normalized counts plus the changed definition of *CAL* for snow covered regions are shown in Fig. 4. Thereby, SNOW refers to the processing with the modified Heliosat whereas ORIG applies the standard/original Heliosat algorithm. The modification affects only ρ_{cs} whereas ρ and ρ_{max} are the same in both cases for each time step and pixel. Thus, the changes in *CAL* are only caused by changes in ρ_{cs} (see Eq. 1) In general, a decrease in the effective cloud albedo is observable. This implies that, due to the modifications, the difference $\rho - \rho_{cs}$ is smaller and, thus, ρ_{cs} is larger than in the original Heliosat.

The original Heliosat performs a monthly minimum value composite for ρ_{cs}.The modifications use a 7 (clear sky) day running mean as value for ρ_{cs}. This value is in general delimited at its lower bound by the minimum value determined in the original Heliosat algorithm. However, higher ρ_{cs} values are possible due to changes in the surface albedo and/or cloud contamination within the regarded month.

This effect is especially strong over snow and ice as well as in cloudy regions close to the border of the disk (northern and southern Atlantic). In those cloudy regions the contamination of ρ_{cs} with clouds is more likely in SNOW than in ORIG leading to a lower amount of clouds (lower *CAL*) in these regions. Furthermore, the altered calculation of *CAL* over snowy areas additionally lowers the effective cloud albedo *CAL*. This leads to higher radiation in these areas for SNOW than for ORIG.

A lower decrease of ρ_{cs} or even increase in effective cloud albedo due to the modifications is visible during JJA over the southern hemisphere. This effect might be caused by a trapping of ρ_{cs} at too low values. The satellite data can contain significant noise in dark areas due to low digital quantification. If the count D is within the magnitude of the dark offset D_0, which occurs for dark surfaces, every value of ρ which is lower than the preceding ρ_{cs} would automatically

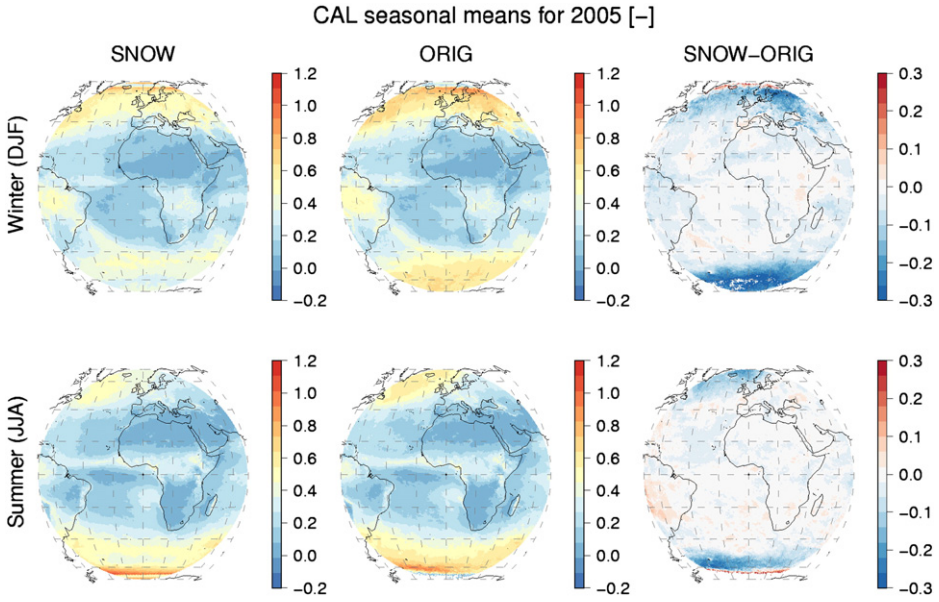

Fig. 4. Changes in the effective cloud albedo *CAL* due to the modifications (SNOW) in comparison to the original Heliosat method (ORIG).

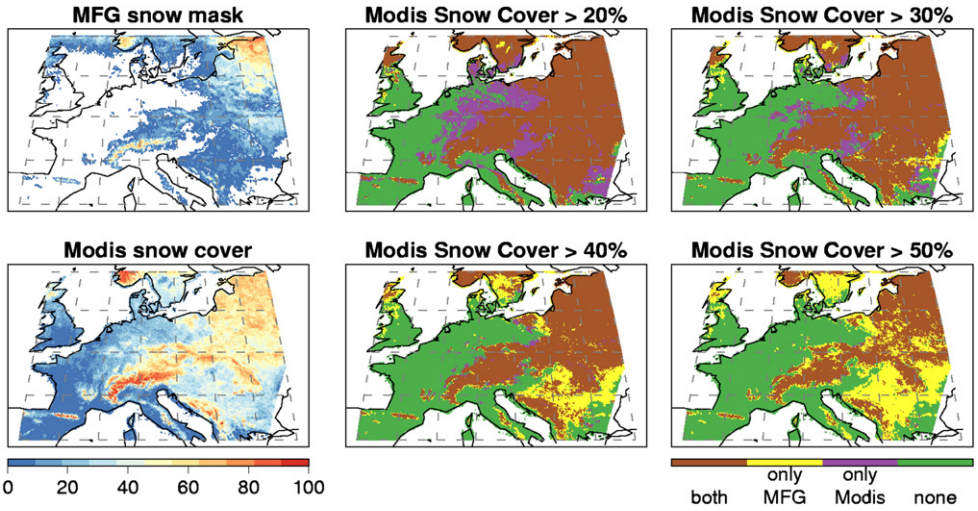

Fig. 5. Comparison of monthly mean Heliosat snow mask with monthly mean fractional snow cover derived from MODIS for DJF 2005.

fulfill the bandwidth criteria. This is not problematic if the observed normalized counts are correct but would lead to a trap for artificial low normalized counts introduced by noise close to the dark offset. Especially, as the decrease of ρ_{cs} is fast whereas the increase is slow. Hence, this could lead to artificial low clear sky normalized counts and, in turn, to an artificial overestimation of cloud albedo.

However, the modifications in the determination of ρ_{cs} with the application of a 7 (clear sky) day running mean prevent unphysical jumps of ρ_{cs} and thus *CAL* at the month's borders which serves the homogeneity of the dataset. In comparison to the original algorithm our modifications generate a continuous seasonal cycle of ρ_{cs}. However, validations are difficult due to missing measurements of this parameter. ρ_{cs} is not the surface albedo but the planetary clear sky reflection given in normalized counts and not in radiances.

3.3. Snow mask

As part of the modifications to Heliosat a snow mask is determined which influences the determination of the effective cloud albedo. The snow mask detected by the modified clear sky albedo algorithm is compared to the fractional snow cover derived from MODIS (Moderate Resolution Imaging Spectroradiometer Hall et al., 2002) observations. For the comparison one has to keep in mind that the Heliosat snow mask is only binary stating 0 for “no snow” and 1 for “snow” whereas MODIS has fractional cloud cover ranging from 0 (no snow) to 100% (all snow covered).

Fig. 5 shows the monthly mean snow mask/cover fields as well as the comparison between Heliosat and MODIS for various MODIS snow cover thresholds over Europe. Furthermore, for the comparison between MODIS Snow Cover and the Meteosat snow mask the following scores and skill scores were used: the proportion correct (PC), the Bias and the Hansen–Kuipers-Discriminant (HK). Descriptions of these scores and skill scores are given in Appendix B and the actual values are listed in Table 3.

Table 3
Scores and skill scores for the comparison between MFG snow mask and different MODIS snow cover thresholds.

	PC	BIAS	HK
MODIS snow cover > 20%	0.83	0.80	0.71
MODIS snow cover > 30%	0.87	0.98	0.75
MODIS snow cover > 40%	0.83	1.32	0.70
MODIS snow cover > 50%	0.74	1.90	0.62

From Fig. 5 it can be seen that the overall snow cover pattern determined by MODIS is reproduced by the Heliosat snow detection algorithm. However, low snow amounts which correspond to low MODIS snow cover values are not detected by Heliosat. The monthly mean Heliosat snow mask corresponds best to a MODIS snow cover field starting at 30% snow cover as for this threshold the largest PC and HK are obtained as well as a Bias *B* close to 1 (see Table 3).

4. Validation of the dataset

The validation employs several statistical measures and scores to evaluate the quality of the SIS, SID an *CAL* CDR. Beside the commonly used bias, mean absolute bias (MAB) and standard deviation (SD), we also use the correlation of the anomalies (AC) derived from the surface measurements and the CM SAF dataset. For each dataset we further provide the fraction of time steps (months or days) that exceed the accuracy threshold ($Frac_{time}$) to characterize the quality of the data sets. The applied quality measures are defined in the Appendix C.

4.1. Solar surface irradiance–global irradiance

The results of the validation of the monthly mean SIS are summarized in Table 4 for the overall performance of the CM SAF CDR at all BSRN stations. It shows that the MAB of the dataset is well below the requested accuracy threshold of 15 W m^{-2} and even fulfills the optimal accuracy requirement of 8 W m^{-2}. In total only 11.35% of the monthly mean data exceed the target accuracy. The dataset is also able to reproduce the monthly mean anomalies of SIS that were measured at the surface, which is shown by the high anomaly correlation of 0.89.

Table 5 provides the validation result for the daily means of the CM SAF SIS CDR. As expected, the mean bias values are very

Table 4
Statistics for the comparison of monthly mean SIS between the mean of all BSRN stations and CM SAF as well as HelioClim, ERA-Interim, GEWEX and ISCCP.

SIS (monthly mean)	n_{mon}	Bias [W m^{-2}]	MAB [W m^{-2}]	SD [W m^{-2}]	AC	$Frac_{mon}$ >15 W m^{-2} [%]
CM SAF–BSRN	855	4.40	7.99	8.14	0.89	11.35
HelioClim–BSRN	855	−17.90	22.76	16.87	0.74	58.83
ERAint–BSRN	855	5.71	10.61	12.15	0.80	25.26
GEWEX–BSRN	855	−2.56	12.26	11.12	0.82	32.75
ISCCP–BSRN	855	−0.03	11.79	11.37	0.78	29.94

Table 5
Statistics for the comparison of daily mean SIS between the mean of all BSRN stations and CM SAF as well as ERA-Interim, GEWEX and ISCCP.

SIS (daily mean)	n_{day}	Bias [W m^{-2}]	MAB [W m^{-2}]	SD [W m^{-2}]	AC	Frac$_{day}$ >25 W m^{-2} [%]
CM SAF–BSRN	28,674	4.58	15.52	23.73	0.92	17.08
HelioClim–BSRN	28,674	−17.94	34.89	43.13	0.74	50.09
ERAint–BSRN	28,674	5.64	26.86	39.14	0.74	36.62
GEWEX–BSRN	28,674	−2.81	22.90	31.31	0.85	32.89
ISCCP–BSRN	28,674	0.40	26.62	37.22	0.76	37.65

comparable to of the ones previously shown for the monthly means while the MAB values for the daily means are about twice as high as those for the monthly means. Still, the MAB of the CM SAF SIS daily mean data set (i.e., 15.52 W m^{-2}) is well below the limit of 25 W m^{-2} and very close to the optimal accuracy of 15 W m^{-2}. Nearly 83% of the individual daily absolute bias values meet the target accuracy.

Also included in these tables are the corresponding values for the four intercomparison datasets. It is visible that for nearly all quality measures the CM SAF SIS CDR has the highest quality among the evaluated data sets. Especially the spread (SD) and the fraction of month above the target accuracy of the intercomparison datasets are considerably larger, resulting in higher uncertainties. However, the higher spatial resolution of the CM SAF dataset improves the colocation with the BSRN stations. This in turn is partly responsible for the better performance in comparison to the other datasets.

A detailed illustration of the absolute bias at each considered BSRN station and for all stations (last column) is shown in Fig. 6. The box–whisker plots represent the range between the 25% and 75% percentiles (1st and 3rd quartiles — Q_1 and Q_3) with the colored boxes and the whiskers extend to 1.5 times the interquartile range $IQR = Q_3 - Q_1$ (see also Wilks, 2006). The IQR is a measure for the spread of the corresponding absolute bias distribution. Additionally, the median (2nd quartile — Q_2) and the mean value which corresponds to the MAB are depicted.

The CM SAF dataset has the lowest MAB of all four monthly mean datasets (Fig. 6 top panel). Furthermore, the spread of the absolute bias is also very small. HelioClim shows for some stations extremely large deviations from the surface measurements as well as from the other intercomparison datasets. This could be due to the reduced spatial resolution of the freely available Helioclim data. The high resolution data might perform better at these stations. ISCCP FD and GEWEX SRB strongly underestimate the incoming solar radiation at the desert stations of Sede Boqer (sbo), Solar Village (svo) and Tamanrasset (tam) resulting in large absolute biases. The CM SAF SIS CDR and ERA-Interim perform much better at these stations. At the station of Lerwick (ler) the opposite can be observed. ISCCP FD and GEWEX SRB yield much better SIS values than CM SAF. This might be due to the position of the station, which is located far north and thus, close to the border of the satellite's visible disk. Additional information from polar orbiting satellites that are used in ISCCP FD and GEWEX SRB datasets are likely to help to constrain SIS in such areas.

The daily mean absolute bias for the individual BSRN stations are shown in Fig. 6 (bottom). Generally, the CM SAF SIS CDR shows the best performance with lowest MAB and smallest spread. At nearly all stations the target accuracy is reached. The only exception is Florianopolis, where the MAB marginally misses the target accuracy of 25 W m^{-2}. However, the intercomparison datasets perform even worse at Florianopolis with extremely large spreads. Problems with the surface measurements (that cancel out when calculating monthly averages) could be responsible for this anomalous behavior in all datasets.

The intercomparison between the various datasets for the seasonal means on the full disk is shown in Fig. 7. HelioClim is not included here because the retrieval via the SoDa service allows only point extraction of the data. The overall, seasonally induced patterns are equally well represented in all four datasets. Differences occur especially in the tropics along the inner-tropical convergence zone

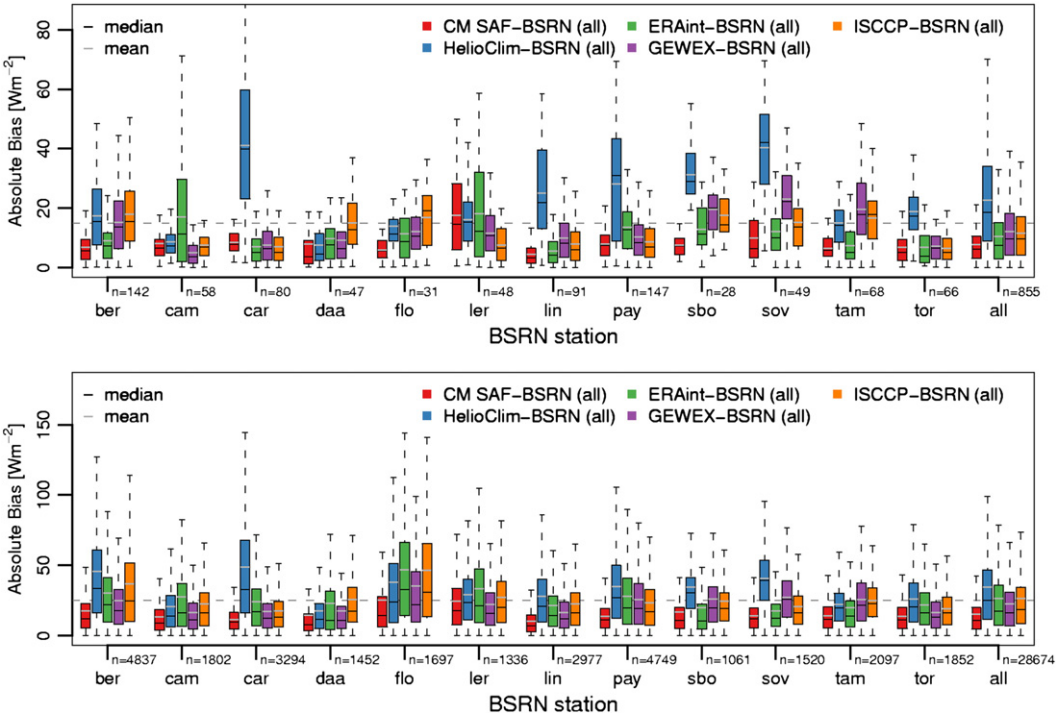

Fig. 6. Absolute bias [W m^{-2}] of SIS for the monthly (top) and daily (bottom) difference between BSRN surface measurements and CM SAF, ERAint, GEWEX SRB, Helioclim and ISCCP FD, respectively.

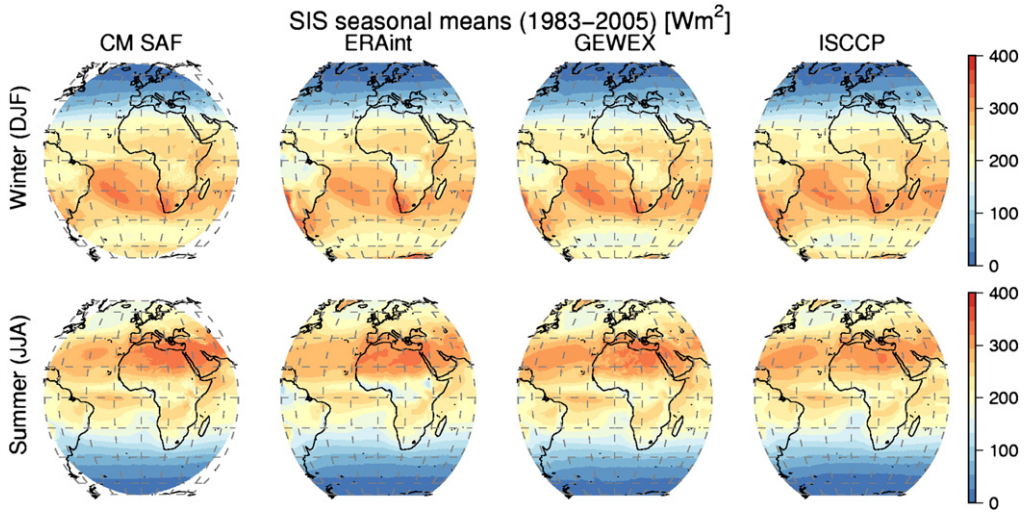

Fig. 7. Seasonal means for the CM SAF dataset compared to the datasets from ERAint, GEWEX SRB and ISCCP FD.

(ITCZ) and in the stratocumulus region off the coast of South-West Africa. Unfortunately, no surface stations applicable for validation are available in these regions. Hence, no conclusion about which dataset represents the actual state better can be drawn.

4.2. Direct radiation

Table 6 shows the validation results of the CM SAF SID CDR. The MAB is 11.0 W m^{-2} and hence, well below the required validation threshold of 20 W m^{-2} and also below the optimal accuracy of 12 W m^{-2}. Thus, the accuracy requirement is fulfilled. The standard deviation and, thus, the spread is also larger for SID than for SIS (12.69 W m^{-2} compared to 8.14 W m^{-2}). The fraction of months that show differences outside of the accuracy requirement is comparable to the corresponding value for SIS. The anomaly correlation is still very good with a value of 0.83.

The validation results for the daily means of the CM SAF SID CDR are shown in Table 7. The MAB is slightly larger than for the daily mean SIS CDR (20.73 W m^{-2} compared to 15.52 W m^{-2}), but it is well below the required validation threshold of 30 W m^{-2} and close to the optimal accuracy of 20 W m^{-2}. Thus, the accuracy requirement is fulfilled. As for SIS, also the daily mean SID shows a larger spread than the monthly means.

The results for the individual BSRN stations are shown in Fig. 8. For nearly all stations the accuracy requirement is met for the monthly as well as for the daily means. The MAB is well within the requested 20 W m^{-2} for monthly means (top panel) and 30 W m^{-2} for daily means (bottom panel), respectively.

In comparison to SIS, SID MAB is within the limits at Lerwick. However, the spread is still substantial. At Lerwick the validation thresholds are exceeded during several months. The possible causes outlined for SIS should also hold for SID.

Larger issues are present at the station of Tamanrasset where SID is substantially underestimated by the CM SAF CDR. Also the two other desert stations, Sede Boqer and Solar Village, show relatively high MABs and spreads. At all three stations, however, the relative error is small as these stations experience high direct solar radiation. The cloudless conditions result in a higher sensitivity of SID on the direct clear sky radiation, which depends on the prescribed atmospheric conditions. Especially uncertainties in the optical properties of aerosol particles (e.g., in dust storms) result in substantial uncertainties in the direct clear sky radiation. For example the fact that the SIS is replicated with little error, while SID is not, suggests that there might be a problem with aerosols in Tamanrasset.

Table 6
Statistics for the comparison of monthly mean SID between the mean of all BSRN stations and CM SAF.

SID (monthly mean)	n_{mon}	Bias [W m^{-2}]	MAB [W m^{-2}]	SD [W m^{-2}]	AC	Frac$_{mon}$ >20 W m^{-2} [%]
CM SAF–BSRN	805	0.89	11.00	12.69	0.83	15.40

Table 7
Statistics for the comparison of daily mean SID between the mean of all BSRN stations and CM SAF.

SID (daily mean)	n_{day}	Bias [W m^{-2}]	MAB [W m^{-2}]	SD [W m^{-2}]	AC	Frac$_{day}$ >30 W m^{-2} [%]
CM SAF–BSRN	26,614	0.74	20.73	30.34	0.89	23.42

4.3. Homogeneity

The homogeneity of a climate data set is particularly necessary in order to enable a meaningful statistical trend analysis. Artificial steps and/or temporal trends in the dataset, e.g., due to changes in the satellite instrument, would result in unrealistic changes and trends, which do not represent changes or trends of the climate. Special attention is therefore given to the times when the satellite instruments changed. Table 8 gives an overview over the major operational periods (longer than 3 months) of the individual Meteosat satellites. Switches between satellites for a few days due to the decontamination procedure are not listed here. For a complete listing of Meteosat operational periods see Rigollier et al. (2002).

A simple method to test for homogeneity is to analyze the anomalies with respect to any obvious steps. Changes in the mean state from one satellite to the other would be visible as an increase or decrease in positive or negative anomalies. Another, more objective way to test the homogeneity of a time series is to analyze the temporal derivative of the time series. Inhomogeneities and change points then appear as systematic (i.e., for a wide range of pixels) positive or negative values for one time step. Fig. 9 shows the Hovmoeller diagram of the temporal derivative of the monthly mean anomalies of SIS. The dashed vertical lines mark the satellite switches listed in Table 8. The zonal means were calculated for the longitude band from 10°W to 30°E (contains Europe, the Sahara and the South Atlantic). The time range contains all satellites starting with Meteosat 2 in

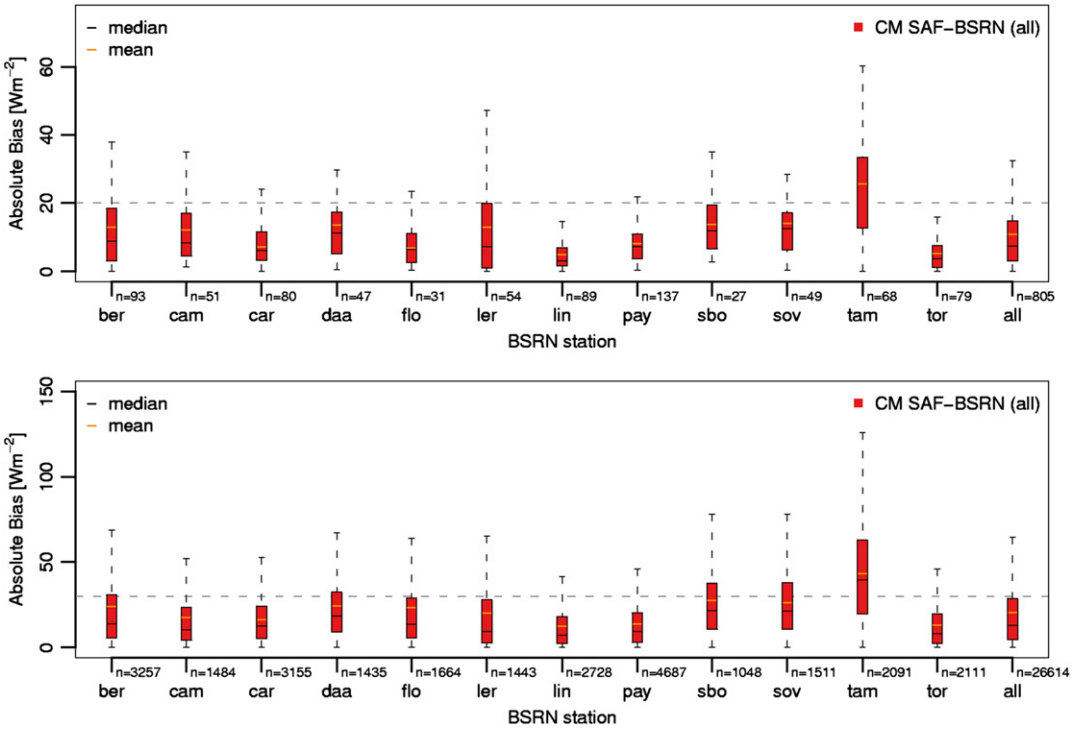

Fig. 8. Absolute bias [W m^{-2}] of SID for the monthly (top) and daily (bottom) difference between BSRN surface measurements and CM SAF.

Table 8
Major operational periods for the used Meteosat satellites.

Satellite	From	To
Meteosat 2	16 Aug 1981	11 Aug 1988
Meteosat 3	11 Aug 1988	19 Jun 1989
Meteosat 4	19 Jun 1989	24 Jan 1990
Meteosat 3	24 Jan 1990	19 Apr 1990
Meteosat 4	19 Apr 1990	4 Feb 1994
Meteosat 5	4 Feb 1994	13 Feb 1997
Meteosat 6	13 Feb 1997	3 Jun 1998
Meteosat 7	3 Jun 1998	31 Dec 2005

1983 until Meteosat 7 in 2005. The Hovmoeller diagram for the SID CDR resembles the ones shown in Fig. 9 and is therefore not shown here.

The temporal derivative of the anomaly shows a small jump from Meteosat 3 to Meteosat 4 in April 1990. This feature is mostly prominent from the Equator to 40°N, thus, covering the region of the Sahara and North Africa. Also, there seems to be a slight inconsistency during the switch from Meteosat 5 to 6 in February 1997, which is however, limited to the region between the Equator and 20°N. For the mentioned regions and time spans, the CM SAF SIS and SID CDR should be used with care.

Inhomogeneities and artificial jumps might be induced by the self-calibration method, discussed in Section 3.1. The stability and performance of the self-calibration has been evaluated against ground measurements from the Baseline Surface Radiation Network (BSRN) (Ohmura et al., 1998) as these data constitutes a reliable reference. The BSRN measurements have quite different start and end points, leading to different weights of local bias values. In order to avoid corruption of the analysis induced by local bias values the local bias values are corrected, leading to a bias of zero for each station. However, trends induced by temporal changes in the monthly differences are not affected by this normalization procedure. The differences between the monthly means of the BSRN and satellite data are then averaged over all available stations, whenever monthly means from at least 3 stations are available. This leads to an overall time series of "normalized" differences of monthly means, referred to as normalized

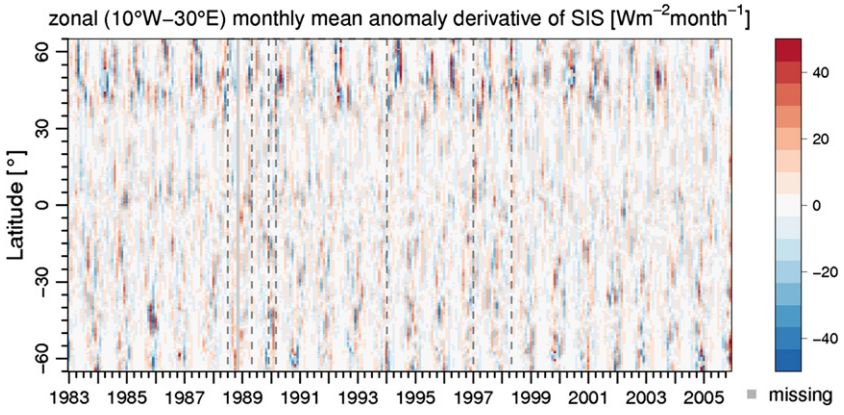

Fig. 9. Hovmoeller diagram of monthly mean anomaly derivation averaged over 10°W to 30°E.

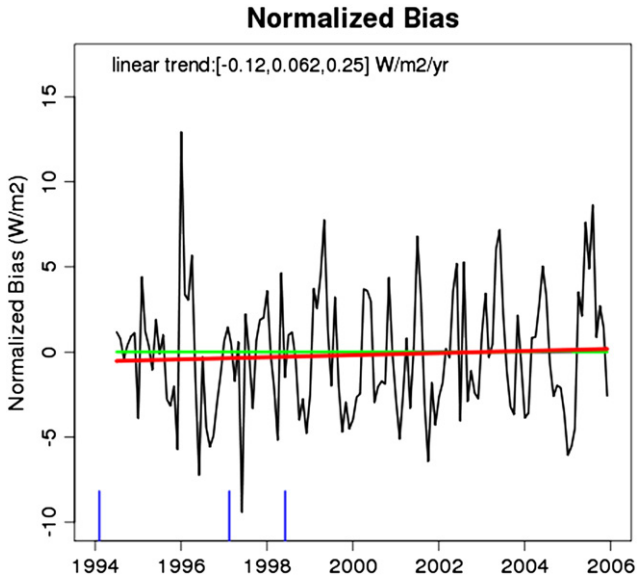

Fig. 10. Temporal evolution of the normalized bias between the CM SAF data set and the BSRN data including the trend colored in red. The green line is the neutral axis. The blue lines at the bottom of the plot indicate the change of satellites (Meteosat-5, Meteosat-6, Meteosat-7). The slight slope of the red line is statistically not significant, hence no significant trend is apparent.

bias. Fig. 10 shows the temporal evolution of the normalized bias between the CM SAF data set and the BSRN data.

As can be seen in Fig. 10 there is no detectable trend in the temporal evolution of the normalized bias (monthly differences), which proves evidence of the homogeneity of the SSI data for the analyzed time period. In contrast, the satellite raw counts show a significant and large trend, a decrease of the sensibility due to aging (see Fig. 3). In contrast to CM-SAF, a significant trend in the normalized bias is apparent in the GEWEX and ISCCP data set (not shown here). A more complete homogeneity analysis will be carried out in a follow-up study.

5. Summary and conclusion

The satellite-derived data sets of the surface global and direct irradiance (SIS and SID) as well as the effective cloud albedo from CM SAF have been presented in this publication. The dataset employed data from visible channel of the Meteosat first generation satellites and, thus, spans a time range from 1983 to 2005. The dataset is freely available for all purposes under wui.cmsaf.eu as monthly, daily and hourly means at a spatial resolution of 0.03°.

The applied climate version of the Heliosat algorithm included various modifications to ensure the generation of a homogeneous climate data record. These modifications include a self-calibration algorithm and changes in the derivation of the clear sky normalized counts. Within the self-calibration the degradation of the satellites as well as gain changes are captured. The modifications in the clear sky algorithm lead mainly to lower effective cloud albedos in the strongly cloud covered region at the north and south rims of the Meteosat disk.

However, an advantage compared to the original algorithm is the application of a running mean preventing unrealistic jumps of ρ_{cs} at the month borders. Still, improvements to the method applied to retrieve the clear sky reflection are necessary and will be subject of future CM SAF projects. For instance, the MVIRI thermal channel could be concurrently used together with the VIS channel to retrieve clear sky pixels.

The SIS and SID datasets are validated by comparison with observations from 12 high-quality ground based stations of the BSRN network. The applied validation thresholds combine the target accuracy, which is based on the GCOS accuracy requirement for the variables of the surface radiation budget, and the systematic error of the BSRN surface measurements.

The MABs of the monthly means of SIS and SID are well below the target accuracy of 10 W m^{-2} and 15 W m^{-2} for the Mean Absolute Bias. The validation target is also reached for most considered stations, only at stations close to the border of the visible disk (SIS and SID) and at desert stations with high solar insulation (SID) the absolute bias exceeds the target accuracy. At Tamanrasset SID shows quite large bias values, but this could be also due to inaccurate SID ground measurement. However, about 90% (SIS) and 85% (SID) of the monthly absolute bias values are below the respective target accuracy.

The MAB for the daily means is also below the target accuracy of 20 and 25 W m^{-2} for SIS and SID for the complete data set and also for most individual stations. The problematic stations are the same as for the monthly means. The target accuracy is fulfilled at about 85% (SIS) and 75% (SID) of all considered individual days at all considered stations.

The evaluation of the SIS CDR with the SIS datasets from Helio-Clim, ERA-Interim, GEWEX SRB and ISCCP FD demonstrated that the monthly and daily mean CM SAF SIS CDR has a higher quality than the inter-comparison data sets. The MAB of the CM SAF SIS dataset is substantially lower and the differences to the BSRN measurements have a much smaller spread than the three evaluation datasets. Moreover the amount of absolute bias values above the target of 10 W m^{-2} (plus uncertainty of ground measurements) is much lower for the CM-SAF climate data record. This is especially important for a reliable monitoring of extremes.

Overall, it was shown that the target accuracy is achieved for monthly and daily means of the global and direct surface solar irradiance in the CM SAF CDR. The effective cloud albedo is a central input quantity for the calculation of SIS and SID. The high accuracy of these data sets demonstrates in turn the high quality of the effective cloud albedo.

In general, the validation demonstrates the outstanding quality of the CM-SAF data set for climate monitoring and analysis, also in intercomparison with other data sets. As a consequence of the high spatial resolution the CM-SAF data can be also applied to regional climate monitoring and analysis. Moreover, the successful retrieval of high quality data from the non-calibrated Meteosat First Generation satellite enables the generation of long term data sets. Heliosat in combination with MAGIC can be also applied to the Meteosat Second Generation satellites, which enables the prolongation and continuation of this high quality data set. Thus, the analysis of climate trends and extremes becomes possible. The self-calibration approach enables the application of Heliosat to other geostationary satellites, e.g. GOES, as well as to the Meteosat-East satellites. Within this scope Heliosat has been implemented at the Joint Research Center for the processing of Meteosat East data. The first validation results show a similar accuracy (T. Huld, personal communication). The application of the method to other geostationary satellites enables the extension of the data set toward a geostationary ring, which increases the attractivity of the data set for further applications, e.g. the evaluation of global climate models.

Appendix A.

Nomenclature

D	digital count (range 0–1023)
D_0	dark offset in digital count
ρ	normalized count
ρ_{cs}	clear sky normalized count
ρ_{max}	maximum normalized count that serves as satellite calibration parameter

CAL	effective cloud albedo (formerly also known as cloud index)
SIS	solar surface irradiance = global radiation
SIS_{cs}	clear sky solar surface irradiance
SID	direct irradiance = direct radiation
SID_{cs}	clear sky direct irradiance = direct radiation

Appendix B. Applied scores and skill scores

The definitions of the scores and skill scores are taken from Wilks (2006). The applied contingency table for "forecasted" and "observed" snow events is shown in Table B.1.

Table B.1
Contingency table for comparison between MODIS snow cover and Meteosat snow mask.

		MODIS (=observation)	
		Snow	No snow
MFG (=forecast)	Snow	*a*	*b*
	No snow	*c*	*d*

Accuracy reflects the correspondence between pairs of forecasts and observations. A perfectly accurate forecast would result in $b=c=0$. One accuracy measure is the proportion correct (PC) given by

$$PC = \frac{a+d}{a+b+c+d}. \tag{B.1}$$

The worst PC to obtain is 0 and the best is 1.

Bias is the comparison of the average forecast to the average observation. It is defined as the ratio of the number of forecasted to the number of observed events.

$$B = \frac{a+b}{a+c} \tag{B.2}$$

Thereby, unbiased forecasts yield $B=1$. For $B<1$ and $B>1$ the events are under- and overforecasted, respectively. The bias provides no information about the correspondence of the events.

Hansen–Kuipers-Discriminant or Peirce skill score is a skill score based on PC. Skill scores quantify the improvement of the actual forecast in comparison to a reference forecast (in case of HK it is a random (unbiased) forecast). The HK can also be understood as the difference of two conditional probabilities, the "probability of detection" (POD) and the "probability of false detection" (POFD).

$$\mathrm{HK} = \mathrm{POD} - \mathrm{POFD} = \frac{a}{a+c} - \frac{b}{b+d} = \frac{ad-bc}{(a+c)(b+d)} \tag{B.3}$$

The forecast has skill in comparison to the reference forecast if $HK>1$. For $HK\leq 1$ the forecast has equal or less skill than the reference forecast.

Appendix C. Statistical measures for the validation

In the following, the applied quality measures are described. The definitions of the statistical measures are taken from Wilks (2006). Thereby, the variable y describes the dataset to be validated (e.g., CM SAF) and o denotes the reference dataset (i.e., BSRN). The individual time step is marked with k and n is the total number of time steps.

Bias

The bias (or mean error) is simply the mean difference between the two considered datasets. It indicates whether the dataset on average over- or underestimates the reference dataset.

$$\mathrm{Bias} = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k) = \bar{y} - \bar{o} \tag{C.1}$$

Mean absolute bias (MAB)

In contrast to the bias, the mean absolute bias (MAB) is the average of the absolute values of the differences between each member of the time series.

$$\mathrm{MAB} = \frac{1}{n}\sum_{k=1}^{n}|y_k - o_k| \tag{C.2}$$

Standard deviation (SD)

The standard deviation SD is a measure for the spread around the mean value of the distribution formed by the differences between the generated and the reference dataset.

$$\mathrm{SD} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}((y_k - o_k) - (\bar{y} - \bar{o}))^2} \tag{C.3}$$

Anomaly correlation (AC)

The anomaly correlation AC describes to which extent the anomalies of the two considered time series correspond to each other without the influence of a possibly existing bias. The correlation of anomalies retrieved from satellite data and derived from surface measurements allows the estimation of the potential to determine anomalies from satellite observations.

$$\mathrm{AC} = \frac{\sum_{k=1}^{n}(y_k - \bar{y})(o_k - \bar{o})}{\sqrt{\sum_{k=1}^{n}(y_k - \bar{y})^2}\sqrt{\sum_{k=1}^{n}(o_k - \bar{o})^2}} \tag{C.4}$$

Here, for each station the mean annual cycle $\bar{y}$ and $\bar{o}$ were derived separately from the satellite and surface data, respectively. The monthly/daily anomalies were then calculated using the corresponding mean annual cycle as the reference.

Fraction of time steps above the validation threshold (Frac)

A measure for the uncertainty of the derived dataset is the fraction of the time steps that are outside the requested thresholds Th.

$$\mathrm{Frac} = 100\frac{\sum_{k=1}^{n} f_k}{n} \text{ with } \begin{cases} f_k = 1 & \text{if } y_k > \mathrm{Th} \\ f_k = 0 & \text{otherwise} \end{cases} \tag{C.5}$$

References

Berrisford, P., Dee, D., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., & Uppala, S. (2009). *The ERA-Interim archive.* ERA report series 1*ECMWF*. Shinfield Park: Reading.

Beyer, H. G., Costanzo, C., & Heinemann, D. (1996). Modifications of the heliosat procedure for irradiance estimates from satellite images. *Solar Energy*, *56*, 207–212.

Cano, D., Monget, J. M., Albuisson, M., Guillard, H., Regas, N., & Wald, L. (1986). A method for the determination of the global solar-radiation from meteorological satellite data. *Solar Energy*, *37*, 31–39.

Cros, S., Albuisson, M., Lefevre, M., Rigollier, C., & Wald, L. (2004). *Helioclim: a long-term database on solar radiation for Europe and Africa. Proceedings of Eurosun 2004*, *3*. (pp. 916–920) Freiburg, Germany: PSE GmbH.

Dürr, B., & Zelenka, A. (2009). Deriving surface global irradiance over the Alpine region from METEOSAT Second Generation data by supplementing the HELIOSAT method. *International Journal of Remote Sensing*, *30*(22), 5821–5841, doi:10.1080/01431160902744829.

Gupta, S., Stackhouse, P., Jr., Cox, S., Mikovitz, J., & Zhang, T. (2006). Surface radiation budget project completes 22-year data set. *GEWEX News*, *16*(4), 12–13.

Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., & Bayr, K. J. (2002). Modis snow-cover products. *Remote Sensing of Environment*, *83*(1–2), 181–194, doi: 10.1016/S0034-4257(02)00095-0.

Hammer, A., Heinemann, D., Hoyer, C., Kuhlemann, R., Lorenz, E., Müller, R., & Beyer, H. (2003). Solar energy assessment using remote sensing technologies. *Remote Sensing of Environment*, *86*(3), 423–432.

Lefèvre, M., Wald, L., & Diabaté, L. (2007). Using reduced data sets ISCCP-B2 from the Meteosat satellites to assess surface solar irradiance. *Solar Energy*, *81*(2), 240–253.

Mayer, B., & Kylling, A. (2005). Technical note: the libRadtran software package for radiative transfer calculations — description and examples of use. *Atmospheric Chemistry and Physics*, *5*, 1855–1877 doi:1680-7324/acp/2005-5-1855.

Mueller, R., Matsoukas, C., Gratzki, A., Behr, H., & Hollmann, R. (2009). The CM-SAF operational scheme for the satellite based retrieval of solar surface irradiance — a LUT based eigenvector hybrid approach. *Remote Sensing of Environment*, *113*(5), 1012–1024, doi:10.1016/j.rse.2009.01.012.

Ohmura, A., et al. (1998). Baseline Surface Radiation Network (BSRN/WCRP): new precision radiometry for climate research. *Bulletin of the American Meteorological Society*, *79*(10), 2115–2136.

Pinker, R., & Laszlo, I. (1992). Modeling surface solar irradiance for satellite applications on a global scale. *Journal of Applied Meteorology*, *31*(2), 194–211.

Posselt, R., Mueller, R., Stöckli, R., & Trentmann, J. (2011). Spatial and temporal homogeneity of solar surface irradiance across satellite generations. *Remote Sensing*, *3*(5), 1029–1046, doi:10.3390/rs3051029.

Rigollier, C., Bauer, O., & Wald, L. (2000). On the clear sky model of the ESRA – European Solar Radiation Atlas – with respect to the Heliosat method. *Solar Energy*, *68*(1), 33–48.

Rigollier, C., Lefevre, M., Blanc, P., & Wald, L. (2002). The operational calibration of images taken in the visible channel of the meteosat series of satellites. *Journal of Atmospheric and Oceanic Technology*, *19*(9), 1285–1293.

Rossow, W. B., & Dueñas, E. N. (2004). The international satellite cloud climatology project (isccp) web site: an online resource for research. *Bulletin of the American Meteorological Society*, *85*(2), 167–172, doi:10.1175/BAMS-85-2-167.

Scharmer, K., & Greif, J. (2000). *The European solar radiation atlas.* TRANSVALOR Presses des MINES.

Schiffer, R., & Rossow, W. (1985). ISCCP global radiance data set: a new resource for climate research. *Bulletin of the American Meteorological Society*, *66*, 1498–1505.

Schulz, J., et al. (2009). Operational climate monitoring from space: the EUMETSAT satellite application facility on climate monitoring (CM-SAF). *Atmospheric Chemistry and Physics*, *9*, 1687–1709.

Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences* International Geophysics Series (2nd ed.). . : Academic Press.

Zelenka, A. (2001). *Estimating insolation over snow covered mountains with Meteosat VIS-channel: a time series approach. Proceedings of the Eumetsat Meteorological Satellite Data Users Conference, Eumetsat Publ. EUM P, Vol. 33.* (pp. 346–352).

Zhang, Y., Rossow, W., Lacis, A., Oinas, V., & Mishchenko, M. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on isccp and other global data sets: refinements of the radiative transfer model and the input data. *Journal of Geophysical Research*, *109*(D19), D19, doi:10.1029/2003JD004457 [105].